

CASE STUDIES ON INSTITUTIONAL REPOSITORY DEVELOPMENT:  
CREATING NARRATIVES FOR PROJECT MANAGEMENT AND ASSESSMENT

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE  
UNIVERSITY OF HAWAI'I IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

COMMUNICATION & INFORMATION SCIENCES

DECEMBER 2008

by  
Jennifer Campbell-Meier

Dissertation Committee:  
Rebecca Knuth, Chairperson  
Diane Nahl  
Péter Jacsó  
Dan Wedemeyer  
Marie Iding

We certify that we have read this dissertation and that, in our opinion, it is satisfactory in scope and quality as a dissertation for the degree of Doctor of Philosophy in Communication and Information Sciences.

DISSERTATION COMMITTEE

Rebecca Knuth  
Chairperson

Marie Iding

Jane Lott

Shirley

Donna

© 2008

by

Jennifer Campbell-Meier

All rights reserved.

## **Acknowledgements**

I would like to thank everyone who assisted me while completing this study, especially the librarians who gave me their time and insight during this process. It was inspiring to discuss IR development with such a dedicated group of people. Special thanks to the librarians at each campus who facilitated my campus visits.

I would also like to thank the members of my dissertation committee, especially Rebecca Knuth, for all their comments and encouragement. I could not have completed this without their support and willingness to work with me at a distance.

Mahalo to the Library and Information Science Program faculty at the University of Hawai'i for their support and encouragement.

Thanks also go out to Susan Johnson, Viil Lid and Alex Bergo for all their help.

I could not have gotten this far without the support of my husband, Steven and my parents, Thomas and Evelyn Campbell.

I dedicate this dissertation to my father, Thomas Patrick Campbell, who always knew I would finish. Every time we spoke, my dad would ask about my progress – how many pages did I write and when I was going to finish. It is with great sadness that I did not get the opportunity to defend my dissertation before he passed away, September 15, 2008.

## **Abstract**

The development of an institutional repository (IR) is one of the more complex projects that librarians may undertake. While many librarians have managed large information system projects, IR projects involve a larger stakeholder group and require support from technical services, public services and administration to succeed. A significant increase in the development of repositories is expected with technology and process improvements for digital collection development. This study investigated the development of repositories at doctoral institutions, identifying factors that influence development and best practices using a comparative case study analysis approach to gather and analyze data. A detailed account and analysis of academic institutional repositories was formed providing knowledge of individual IR development as well as a cross case comparison of developmental factors including adoption, motivating factors and perceived benefits. The use of a narrative, project management practices beyond technical development, and the inclusion of the campus community are identified as key factors in development. Best practices and recommendations for future developers, such as early involvement of stakeholder groups and the need to educate both librarians and faculty about open access collections are also discussed. This study contributes to a more informed understanding of the development of IRs and identifies a model framework for future IR developers.

# Table of Contents

Abstract.....	v
Table of Contents.....	vi
List of Tables.....	xiv
List of Figures.....	xv
List of Charts.....	xvi
List of Charts.....	xvi
List of Abbreviations.....	xvii
Chapter 1. Introduction.....	1
1.1 Institutional Repositories.....	2
1.2 Institutional Repository Research.....	8
1.3 Research Questions.....	13
1.4 Significance.....	14
1.5 Research Design.....	15
Chapter 2. Literature Review.....	17
2.1 Scholarly Communication.....	17
2.2 Open Access.....	18
2.3 Institutional Repositories.....	18

2.3.1 Models.....	19
2.3.2 Software .....	20
2.3.3 Permanence .....	22
2.3.4 Content.....	23
2.3.5 Community-Driven .....	25
2.3.6 Growth .....	28
2.3.7 Development .....	28
Chapter 3. Methodology .....	29
3.1 Case Study Literature.....	29
3.2 Case Study Design .....	30
3.3 Research Questions.....	33
3.4 Definitions.....	34
3.5 Preliminary Interviews.....	39
3.5.1 Preliminary Interview Discussion.....	41
3.6 Sample Selection.....	43
3.7 Selecting Sites for Study.....	43
3.8 Case Study Sites.....	45
3.9 Data Collection Methods .....	48
3.9.1 Semi-Structured Interview .....	48
3.9.2 Data analysis .....	49
3.9.3 Outline of the case study report .....	50

3.10 Limitations .....	51
3.11 Anticipated results .....	51
Chapter 4. Case Studies .....	52
4.1 Indiana University.....	52
4.1.1 Background.....	52
4.1.2 Repository Development .....	52
4.1.3 Content.....	54
4.1.4 Assessment.....	59
4.1.5 Benefits .....	60
4.1.6 Challenges.....	60
4.1.7 Project Outcomes .....	61
4.2 Purdue University .....	62
4.2.1 Background.....	62
4.2.2 Repository Development .....	62
4.2.3 Content.....	66
4.2.4 Assessment.....	69
4.2.5 Benefits .....	70
4.2.6 Challenges.....	70
4.2.7 Project Outcomes.....	71
4.3 Indiana University Purdue University at Indianapolis.....	72
4.3.1 Background .....	72
4.3.2 Repository Development .....	73

4.3.3 Content.....	74
4.3.4 Assessment.....	78
4.3.5 Benefits .....	79
4.3.6 Challenges.....	79
4.3.7 Project Outcomes .....	80
4.4 University of Washington .....	81
4.4.1 Background .....	81
4.4.2 Repository Development .....	81
4.4.3 Content.....	83
4.4.4 Assessment.....	86
4.4.5 Benefits .....	86
4.4.6 Challenges.....	87
4.4.7 Project Outcomes .....	88
4.5 University of British Columbia.....	89
4.5.1 Background .....	89
4.5.2 Repository Development .....	89
4.5.3 Content.....	91
4.5.4 Assessment.....	94
4.5.5 Benefits .....	94
4.5.6 Challenges.....	95
4.5.7 Project Outcomes .....	95
4.6 Simon Fraser University .....	96
4.6.1 Background.....	96

4.6.2 Repository Development .....	97
4.6.3 Content.....	99
4.6.4 Assessment.....	101
4.6.5 Benefits .....	101
4.6.6 Challenges.....	102
4.6.7 Project Outcomes .....	104
4.7 Subject Repositories.....	106
4.7.1 IUBio Archive.....	106
4.7.2 DSpace at the University of Washington Health Sciences Libraries (HSLIC) .....	106
4.7.3 Centre for the Study of Historical Consciousness (CSHC) .....	107
4.7.4 Comparisons with Institutional Repositories .....	107
Chapter 5. Analysis of Cases .....	109
5.1 Examining Core Functions .....	109
5.1.1 Materials Submission.....	109
5.1.2 Metadata Application.....	110
5.1.3 Access Control .....	111
5.1.4 Discovery Support .....	112
5.1.5 Distribution .....	113
5.1.6 Preservation.....	113
5.1.7 Resources .....	114
5.1.8 Responsibility .....	115

5.2 Examining Development Factors.....	117
5.2.1 Adoption .....	117
5.2.2 Motivating Factors .....	122
5.2.3 Planning .....	128
5.2.4 Stakeholders .....	129
5.2.5 Assessment.....	131
5.2.6 Defining Success.....	133
5.2.7 Content.....	135
5.2.8 Policy Development.....	139
5.2.9 Marketing.....	140
5.2.10 Sustainability.....	140
5.2.11 Perceived Benefits .....	141
 Chapter 6. Conclusions and Future Research Recommendations .....	 144
6.1 A Narrative Approach to IR Development .....	145
6.1.1 Narrative Development.....	146
6.1.2 Storytelling and organizational change.....	148
6.2 Project Management .....	151
6.2.1 Working Groups.....	151
6.2.2 Project Planning.....	153
6.2.3 Resource Allocation.....	154
6.2.4 Marketing.....	155
6.2.5 Assessment Measures .....	157

6.3 Campus community .....	162
6.3.1 Communities of Practice.....	162
6.3.2 Diffusion of Innovation.....	165
6.4 Future Repository Development Plans .....	167
6.5 Resolving Anticipated Results .....	168
6.6 Implications and Contributions.....	171
6.6.1 Primary Contributions.....	172
6.6.2 Secondary Contributions.....	175
6.7 Limitations .....	176
6.8 Directions for Further Research.....	176
APPENDICES .....	178
Appendix A: Interview Guide.....	178
Appendix B: Preliminary Interviews .....	184
Baccalaureate - Liberal Arts (Midwest).....	184
Baccalaureate - Liberal Arts (Northeast) .....	185
Masters' Universities & Colleges I (South) .....	187
Masters' Universities & Colleges I (Midwest).....	189
Doctoral/Research Intensive (East).....	191
Doctoral/Research Extensive (Midwest) .....	192
Doctoral/Research Extensive (South).....	194
Appendix C: Informed Consent Form .....	196

Appendix D: Participant List .....	199
Indiana University.....	199
Indiana University Purdue University Indianapolis.....	199
Purdue University .....	199
University of Washington.....	199
University of British Columbia.....	199
Simon Fraser University .....	200
 Appendix E: Themes, Categories, and Sub-categories Developed for Data Analysis	201
 References.....	203

## List of Tables

1. Characterization of IR Involvement.....	11
2. IR Development in ARL institutions.....	12
3. Repository Software Use .....	21
4. Preliminary Interview Summary.....	40
5. Case Study Sites (December 2007).....	46
6. Comparison of Access Control .....	112
7. Resource Comparison .....	114
8. Responsibility Issues.....	116
9. Adoption .....	117
10. Types of Repositories* .....	119
11. What Motivated Your Institution to Establish an IR? (Bailey, 2006, p. 25) .....	122
12. Motivators for Developing a Repository .....	124
13. Library Goals and Values Associated with Repository Development.....	126
14. Content.....	136
15. Intrinsic Motivation, Factors and Benefits of IR Development.....	143
16. Examples of Stories Associated with IR Development.....	150

## List of Figures

1. Gartner Hype Cycle for Education Technology with E-Learning Repositories circled.	9
2. Product Life Cycle with Repository Development Identified in Growth. ....	10
3. IR Adoption Based on Statistics from the Berlin 3 Meeting. ....	11
4. DSpace's Project Planning Timeline.....	22
5. Academic Publishing in Institutional Repositories.....	24
6. Researcher Page in IUScholarWorks.....	58
7. Purdue Libraries e-Scholar Portal.....	64
8. Purdue e-Pubs Growth from Inception to Case Study Visit. ....	67
9. Screenshot with links to Umbrella in CONTENTdm and IDeA .....	75
10. Project Deliverables from the UBC Institutional Repository (IR) Pilot Project.....	93
11. Critical Success Factors. ....	94
12. SFU Institutional Repository Home Page (Screen Capture March 2, 2008). ....	105
13. Comparison of subject repository development .....	108
14. Repository software instances* .....	118
15. Innovation-Decision Process for Repository Implementers, Librarians and Faculty	120
16. Purdue e-Pubs screen capture July 24, 2007.....	131
17. Repository Growth (data from ROAR).....	138
18. Diffusion of Innovation/Adoption (Rogers, 2003) .....	166
19. Basic IR Development Framework.....	172
20. Best Practices Framework Based on Case Study Analysis.....	173

**List of Charts**

1. Assessment Measures ..... 132

2. Definitions of Success..... 134

3. Perceived Benefits ..... 142

## List of Abbreviations

<b>Abbreviations &amp; Acronyms</b>	<b>Definition</b>
AAUC	Association of Universities & Colleges of Canada
ACRL	Association of College and Research Libraries
ALA	American Library Association
ARL	Association of Research Libraries
CIC	Committee on Institutional Cooperation
CNI	Coalition for Networked Information
CoP	Community of Practice
ETD	Electronic Theses and Dissertations
FOLIO	Foundation Literature Online
FTE	Full-time enrollment; Full-time equivalency
IDeA	IUPUI Digital Archive
IMLS	Institute of Museum and Library Services
IR	Institutional Repository
IU	Indiana University
IUPUI	Indiana University Purdue University at Indianapolis
MIRACLE	Making Institutional Repositories A Collaborative Learning Environment
MIT	Massachusetts Institute of Technology
NCPH	National Council on Public History
OA	Open Access
OAI	Open Archives Initiative
OAI-PMH	Open Archives Initiative - Protocol for Metadata Harvesting
OCR	Optical character recognition
ODLIS	Online Dictionary for Library and Information Science
PI	Principal Investigator
PUL	Purdue University Libraries
ROAR	Registry of Open Access Repositories
RoMEO	Rights METadata for Open archiving
SFU	Simon Fraser University
SPARC	Scholarly Publishing and Academic Resources Coalition
UBC	University of British Columbia
USG	University System of Georgia
UW	University of Washington

## Chapter 1. Introduction

Scholarly communication has changed dramatically since the development of the Internet. Schauder (1994) identified electronic publishing as providing “greater diversity and choice” (p. 94) in scholarly communication. The ability to publish, access and store scholarly works digitally has transformed not only how libraries store information, but academic publishing. Libraries, traditionally storehouses of print publications, are expanding services by collecting digital content and becoming content providers by digitizing archival and special collections. No longer are libraries mere storehouses of print materials; by providing access to electronic documents and faster retrieval of information, libraries are adding to the choice and diversity in scholarly publishing.

Traditional publishing models are shifting as electronic publishing changes the ways that libraries purchase and access scholarly materials (ARL, 2006). E-book models are still developing, as more content providers like ebrary and Ingram join NetLibrary as online book distributors. Journal publishing has developed more robustly with electronic copies of articles available in online databases and publisher sites. Scholarly communication on campus is changing as well. Documents, reports and other publications are born digital and may not be available in print on a campus.

Different models are being developed to provide access, manage costs, and manage an organization’s scholarly output, especially at colleges and universities. Developing repositories of scholarly information is one solution. Repositories can be developed at an organizational or institutional level, at a discipline specific or subject level in a collection developed across institutions, or by a particular department or school

within a larger organization. While other subject-based collections exist, they are often developed by individuals rather than an institution. Individual and subject-based collections are not necessarily permanent collections on the Internet and may disappear or close if research interests or faculty<sup>1</sup> appointments change (Gibbons, 2004). To manage and preserve digital content, universities are creating institutional repositories to manage the scholarly materials created by the campus community. These repositories can include traditional forms of scholarly communication such as journal articles and technical reports. They can also be used to capture undergraduate and graduate student work, curricula, and learning objects.

### ***1.1 Institutional Repositories***

The development of institutional repositories at academic institutions worldwide has greatly increased with the growth of open source initiatives in scholarly communication and software development (Swanepoel, 2005). According to Lynch (2003):

a university-based 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. (p. 2)

---

<sup>1</sup> Faculty in this study refers to teaching faculty and researchers with faculty status. Librarians with faculty status are referred to as librarians.

An IR is a digital library, in that it has an organized collection of digital objects pertaining to a particular research or educational organization (Jones, Andrew & MacColl, 2006). Johnson (2002) identifies an IR as: scholarly; cumulative and perpetual; and open and interoperable. Although new technology and publishing have reinvigorated the discussion about institutional repositories, the concept of developing an organized collection of a university's scholarly output has been discussed before.

During the 1980s, college and university archives began institutional repository collections and special-subject collections. Dealing primarily with print materials and papers, university archives have attempted to acquire, retain and preserve faculty papers to document scholarly activity on campus (Honhart, 1983; Shervington, 1983; Wolff, 1981). In addition, the role of the library as a repository and disseminator has also been discussed in terms of an institutional repository associated with a college or university archive (Wolff, 1981; Jones, 1984). Not only were these special collections seen as a growth area for archives, but the purpose of special collections was seen as a way to “project the image of the individual university as an institution of higher learning” (Gakobo, 1985, p. 405).

In the 1990s, librarians began digitizing content in special collections, making it accessible worldwide. Johnson (2002) identifies online IRs as either competing with or complementing these archival repositories. Johnson goes on to suggest that both should be supported since university archives have an additional role in preserving administrative records. Johnson sees a distinction between the intellectual output of faculty, students and staff in an online repository and the “materials pertaining to the institution's history and to the activities and achievements of its officers, faculty, staff, students, and alumni” which an archival institutional repository retains (p. 4).

Within the library various groups may be involved in repository development: technical services, public services and archivists. Technical services, according to the *Online Dictionary for Library and Information Science (ODLIS)*, encompasses the “library operations concerned with the acquisition, organization (bibliographic control), physical processing, and maintenance of library collections, as opposed to the delivery of public services” (Reitz, 2004-06). Public services, on the other hand, includes “[a]ctivities and operations of a library that bring the staff into regular direct contact with its users, including circulation, reference, online services, bibliographic instruction, serials assistance, government documents, and interlibrary loan/document delivery” (Reitz, 2004-06). According to the *Oxford Dictionary of English*, an archivist is an individual who collects and manages a collection of historical documents or records providing information about a place, institution, or group of people (Soanes & Stevenson, 2005). Each group in the library may be involved in repository development at a particular institution, and each brings different skills and outlooks on a repository’s role within the greater organization.

In part, the demand for IRs has developed out of changes in scholarly communication models, as well as the need or desire to develop a clearinghouse for the scholarly output of a particular college or university (Crow, 2002). The preponderance of electronic communication allows authors to easily distribute and share information and articles. Esposito (2004) identifies Open Access (OA) publishing as Internet-based information available to anyone at any time. OA decreases the costs for journals, provides new services and provides access to materials (Esposito, 2004). OA does not negate the costs of scholarly materials nor does it provide an alternative to the peer review process. Rather than placing scholarly information on blogs, wikis or personal web pages, adding content to a university IR provides a controlled venue for OA documents. While not peer reviewed, the institution does lend some authority to the documents available in its repository. In order to control the documents, several universities have developed software that allows for the collection, dissemination and archiving of scholarly materials.

The development of software to manage electronic documents, like DSpace by Massachusetts Institution of Technology (MIT) and EPrints at the University of Southampton, created open source alternatives for IR development. Like EPrints, DSpace is open source software implementing the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) (Smith et al., 2003).

According to the Association of College and Research Libraries' (ACRL) 2003 Environmental Scan, no formal definition has been developed for open access. Most often, the open access contains "free, non-exclusive, access to content" (ACRL, 2003, p. 23). The Open Archives Initiative (OAI)<sup>2</sup> was developed to enhance access to e-print archives as a means of increasing availability as well as developing and promoting "interoperability standards that aim to facilitate the efficient dissemination of content" (OAI, n.d.).

IRs can provide permanent access to digital objects and allow for increased accessibility through open access initiatives, such as OAIster,<sup>3</sup> an OAI-compliant metadata harvester that can search any digital repository, or Google Scholar<sup>4</sup>, which includes IR content in its searches.

The Scholarly Publishing and Academic Resources Coalition (SPARC)'s position paper identified several factors leading to the growth and development of IRs: technology, the increasing volume of research, "dissatisfaction" with current publishing models, and preservation concerns (Crow, 2002). These factors are changing the library's role from "manag[ing] and preserving" information to the traditional publisher roles of "aggregators and distributors" of scholarly information (Crow, 2002, p. 4). In addition, libraries are also working with faculty to retain rights to self-archive and redistribute their scholarly work. To this end, SPARC has identified rationale for IR development: the development of a "new scholarly publishing paradigm" and "institutional visibility and prestige" (Crow, 2002, p. 6).

---

<sup>2</sup> <http://www.openarchives.org/index.html>

<sup>3</sup> <http://oaister.umdl.umich.edu/o/oaister/>

<sup>4</sup> <http://scholar.google.com/>

With the development of electronic communication, academic research has introduced new technologies and tools for scholarly communication. As the landscape surrounding scholarly communication changes, alternatives and uncertainty grow. The Internet's decentralization of publishing creates obstacles to finding a university's intellectual output. According to Rogers (2003, p. 6), uncertainty "implies a lack of predictability, of structure, of information." Time constraints, fear or distrust of online documents and a "loss of control" seem to be limiting faculty participation in open access repositories. Occurring at the university level, IR developers can allay faculty concerns by creating policies and procedures to aid faculty participation.

Many universities have faculty statements that support the use of articles placed in an open access repository for tenure and promotion. Most IR policies allow for the deposit of scholarly output by anyone connected to the host institution. This output includes preprints, post prints, data sets, images, music, and multi-media content (Johnson, 2002). However, faculty members often do not understand copyright issues involved with archiving at either the institutional or personal level.

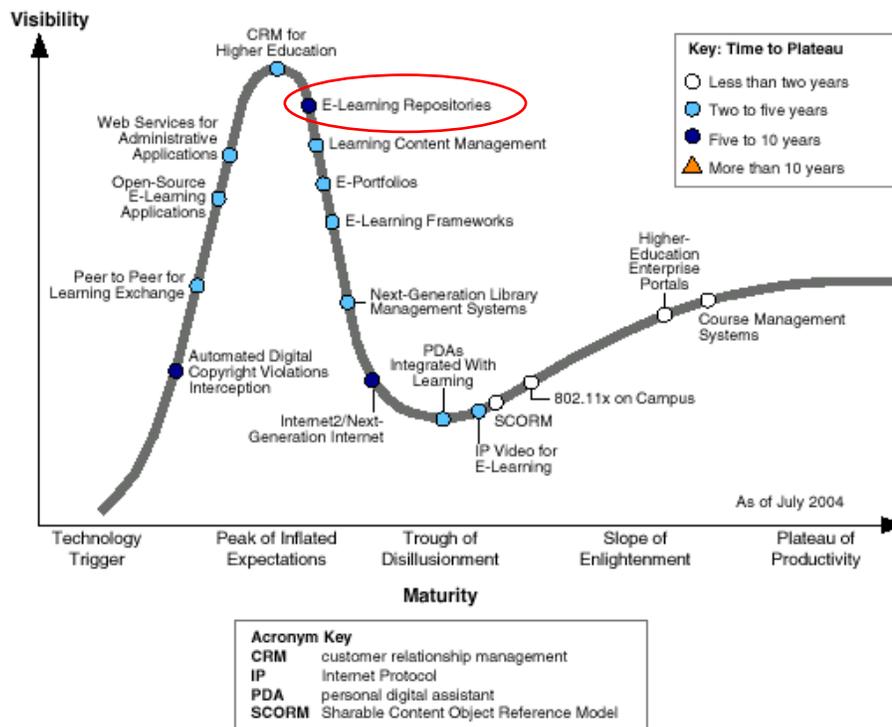
IRs provide faculty with a venue to archive papers within copyright guidelines. They create a record of scholarly publishing for an institution. In addition, while searching for articles in library databases and on the Web, it is not always possible to search by institution or the institutions of second or third authors. With the appropriate metadata, it is possible to search across multiple repositories on the Web, finding articles that are not housed at an author's local repository. Identifying the scholarly output of an institution is a benefit; IRs provide a clearinghouse for the institution and a mechanism for identifying researchers within a university with similar interests.

## ***1.2 Institutional Repository Research***

Clifford Lynch is the Director of the Coalition for Networked Information (CNI), which is jointly sponsored by the Association of Research Libraries (ARL) and EDUCAUSE. Lynch and Joan Lippincott, the Associate Executive Director of the CNI, found that “there has been relatively little systematic examination of the actual state of deployment of institutional repositories in higher education (or even among research universities) across the United States” (2005).

Marius Swanepoel (2005) employed the Gartner Hype Cycle, product life cycle (Sigmoid curve) and diffusion of innovation theory to IR development (Figure 1). A Hype Cycle is a graphic representation of the maturity, adoption and business application of specific technologies. According to his research, IR development is exiting the early adapters phase of the Gartner Hype Cycle (p. 22). Using the Hype Cycle for Education Technology, which identifies IRs as E-Learning Repositories, Swanepoel views IR technology as “on its way to the trough of disillusionment...pass[ing] the slope of enlightenment reasonably soon to reach the plateau of productivity within 5-10 years” (p. 23).

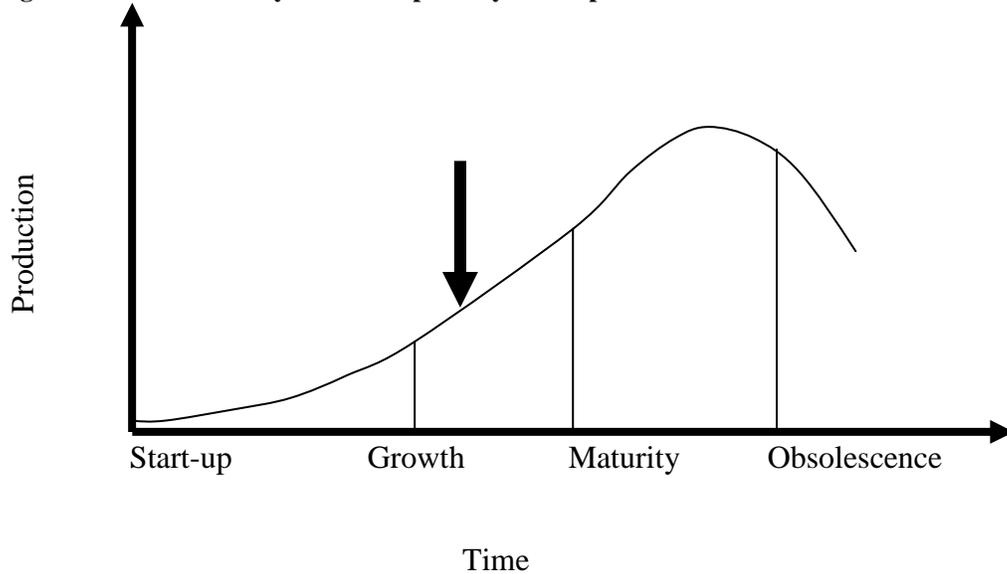
Figure 1. Gartner Hype Cycle for Education Technology with E-Learning Repositories circled.



From Swanepoel, M. (2005). Digital repositories: All hype and no substance? *New Review of Information Networking*, 11(1), p. 20.

Placing IR in the growth phase of the Product Life Cycle, Swanepoel identifies the phase as a period of growth when a new product takes off (p. 21, Figure 2).

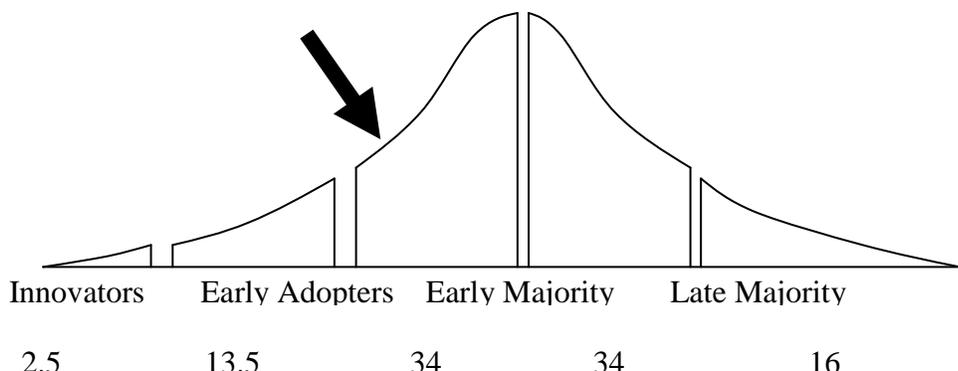
**Figure 2. Product Life Cycle with Repository Development Identified in Growth.**



From Swanepoel, M. (2005). Digital repositories: All hype and no substance? *New Review of Information Networking*, 11(1), p. 17.

Since software development has moved beyond start up, but is still not mature, Swanepoel positions IR production at the beginning of the product life cycle. Finally, using diffusion of innovation theory, Swanepoel places adoption in the Early Majority phase, based on statistics from the Berlin 3 meeting on Open Access in February 2005 (Figure 3). During this meeting, Swanepoel identified 25% of faculty as providing open access to their scholarly output, with a further 79% indicating a “willingness to self-archive” (p. 22). With only 201 registered repositories and over 4,000 institutions of higher learning in the United States, examining the development and planning process of Early Adopters will identify risks for those planning to implement an IR in the future.

**Figure 3. IR Adoption Based on Statistics from the Berlin 3 Meeting on Open Access (February 2005).**



From Swanepoel, M. (2005). Digital repositories: All hype and no substance? *New Review of Information Networking*, 11(1), p. 22.

Current studies are underway at the University of Michigan on IR user behavior.

The MIRACLE (Making Institutional Repositories A Collaborative Learning Environment) Project<sup>5</sup> “addresses the effectiveness of institutional repositories, it accounts for the perspectives of both users and administrative staff” (MIRACLE Proposal, 2005). Preliminary results from 273 respondents to an initial survey of 2,117 college and university libraries found that only 10% of libraries had IRs that were considered fully implemented, 15% were in a pilot and testing stage, 24% were planning and 51% had no plans to date (Markey et al, 2006; Table 1).

**Table 1. Characterization of IR Involvement.**

Number of respondents, n=273

Status	Number of Respondents	Percentage
Implemented	28	10%
Planning & pilot testing	42	15%
Planning only	65	24%
No planning	138	51%

Adapted from Markey, K., St. Jean, B., Rieh, S.Y., Yakel, E., Kim, J., Kim, Y.M. (2006, June). National Census of Institutional Repositories: Preliminary Findings. Paper at the Joint Conference on Digital Libraries 2006, Chapel Hill, NC. Retrieved November 15, 2006 from <http://www.ils.unc.edu/tibbo/JCDL2006/Markey-JCDLWorkshop2006.pdf>.

<sup>5</sup> <http://miracle.si.umich.edu/index.html>

According to Charles Bailey, Assistant Dean for Digital Library Planning and Development at the University of Houston, 30% of all Association of Research Libraries (ARL) institutions “had an operational IR at the beginning of 2006” (2006, p. 13). In January 2006, 123 ARL member libraries were surveyed about IR development; eighty-seven libraries responded to the survey (Bailey, 2006; Table 2).

**Table 2. IR Development in ARL institutions.**

Number of respondents, n=87

<b>Status</b>	<b>Number of Respondents</b>	<b>Percentage</b>
Operational	37	43%
Planning only	31	35%
No plans	19	22%

Adapted from Bailey, C.W. (2006). Institutional Repositories: SPEC Kit 292. Washington, DC : Association of Research Libraries.

From the survey, Bailey found that the top three reasons for implementing an IR were: 1) to increase global visibility, 2) preservation, and 3) the free provision of the institution’s scholarship (2006, p. 14). Finding that most IRs were started within the last two years, the survey identified several benefits and challenges to IR development. The benefits identified by survey participants closely match those listed for implementing the IR: enhanced visibility and increased dissemination; preservation, access and stewardship; and collocation and organization of “assets” (Bailey, 2006, p. 20).

The building and development of an IR has yet to be addressed empirically in the literature. This is especially true of liberal arts institutions that are often working with fewer resources than doctoral institutions. Instead, much of the current literature on institutional repositories is related to the technical and structural problems inherent in building the repositories. In addition, limited research has been done on the political and cultural changes required to develop a successful repository. By using the comparative case study approach, the political and cultural changes can be examined through the repository development process.

### ***1.3 Research Questions***

This study seeks to investigate factors influencing the development of institutional repositories at academic institutions. A comparative case study approach will be used to provide a detailed account and cross case analysis of six repositories at doctoral institutions. Smaller comparisons of development at three discipline repositories are also included in the analysis. The research questions that guide the study are:

1. How are IRs developing?
2. What are the factors influencing IR development?
3. How should IRs develop?
4. What are the best practices for IR development?
5. What is the best model(s) for IR development?

## ***1.4 Significance***

Current literature in library science identifies IRs as a growing phenomenon, with an increasing number of libraries planning to implement collections. In the United States, the growth of IRs has been slower than in Europe. A survey by Lynch and Lippincott (2005) found that more than 40% of the ninety-seven doctoral universities surveyed had developed an institutional repository. While only 178 colleges and universities participated in the survey, many of the respondents were developing or interested in developing an institutional repository. With more than 4,000 degree granting institutions in United States, technology and process improvements to IR software should result in an increase in the development of repositories (Pocket Guide, 2005). The ACRL Environmental Scan (2003) recognized IRs as an “emerging issue that may affect the future of ...academic libraries” (ACRL, 2003, p. 3).

The development of an IR is one of the more complex projects that librarians may undertake. While many librarians have experience with large information system projects, like implementing a new online catalog, IR projects require a large stakeholder group, technical services, public services and administration to succeed. Gibbons (2004, p. 3) identifies “building advocacy, garnering collections, and determining policies” as areas requiring time and effort during IR development; however, it is unclear why libraries choose to undertake IR development without first procuring additional funding from the institution.

In addition, reported start-up costs for IR implementation range from \$8,000 to \$1,800,000 with ongoing operations budgets running between \$8,600 and \$113, 543 for

ARL libraries (Bailey, 2006). With decreasing library budgets competing with the ever increasing costs of resources, IRs can consume a large portion of the library budget.

By identifying crucial factors in IR development and the challenges that the factors pose, a developmental framework can be identified for libraries interested in IRs. Improving IR development may impact other types of libraries or collections wishing to share a digital collection among its members. A framework based on the case studies may provide a model that will aid other university libraries in developing IRs in the future. The factors in the study focuses on the steps taken during IR development, including: studies, policy decisions, timelines, resources and costs. In addition, the study seeks to identify the key stakeholders necessary for project success. Preliminary interviews (Chapter 3; Appendix B), suggest that librarians in public services and university administration played key roles in the successful development of the repository; however, these individuals were not necessarily included in the IR planning process at an early stage. Identifying stakeholders and other key factors will provide an enhanced understanding of the role of an IR at an institution of higher education, as well as a model for the development of IRs.

### ***1.5 Research Design***

This study used comparative case study analysis to investigate the development of nine IRs. Comparative case study analysis highlights the similarities and differences between cases, identifying areas that have direct implications for IR development. Data was collected from multiple sources, including semi-structured interviews and analysis of operational and archival documents. For each case study, data was analyzed during the

collection process and interviews were coded and categorized. Then cross case analysis was used to compare and contrast individual case studies, identify relevant themes and draw conclusions.

## **Chapter 2. Literature Review**

### ***2.1 Scholarly Communication***

Electronic publishing and electronic access have wrought tremendous changes in scholarly communication, especially for periodicals. Scholarly articles are often given to publishers who then sell the content back to the academic community, primarily libraries. Universities, libraries, and scholars are re-examining publishing models, especially with ever-increasing journal prices and constrained library budgets (Gibbons, 2004).

SPARC, OAI and the Budapest Open Access Initiative have been established to bring about change in scholarly publishing. SPARC, a coalition of academic and research libraries, was developed to be “a constructive response to market dysfunctions in the scholarly communication system” (SPARC, 2006). As such, SPARC is focusing on access to peer-reviewed scholarship, promoting open access and the retention of copyright by authors. OAI, on the other hand, seeks to “facilitate the efficient dissemination of content,” focusing on the technology and standards necessary to promote access to scholarly information (OAI, n.d.). The Budapest Open Access Initiative was formed in 2001 to “accelerate progress in the international effort to make research articles in all academic fields freely available on the Internet” (Budapest, n.d.). Together these organizations are working to provide free standardized access to scholarly communication through the development of repositories and repository software.

## ***2.2 Open Access***

According to Peter Suber (2005), self-archiving allows authors to expand access to their works, by providing additional access points. High-impact journals are more likely to permit authors to self-archive, though the article is usually a pre-print or post-print document rather than an electronic copy in Portable Document Format (PDF) of the journal article (Suber, 2005). Harnad and Brody (2004) found that more than 50% of publishers identified by the RoMEO Project (Rights METadata for Open archiving) allowed authors to self-archive.<sup>6</sup> Wren (2005) found correlation between a journal's impact factor and the availability of its articles on the Web. Harnad and Brody (2004) also found that OA articles were more frequently cited. Publications are increasingly available online, posted through repositories, personal websites or publisher sites. With enhanced access, it becomes even more important to place items in a digital collection. Metadata and preservation value-add to deposited articles which are placed in context among a faculty member's publications. This level of indexing and preservation is usually not delivered from a personal web site.

## ***2.3 Institutional Repositories***

The development of IRs has been discussed by Crow (2002), Pinfield, Gardner and MacColl (2002), Shearer (2003), and Gibbons (2004), among others. Pinfield et al. (2002) outline the necessary stages to implement an e-print repository. Shearer (2003) identifies steps to develop a repository using E-Print. Gibbons (2004) discusses development issues in repositories, including content acquisition and faculty

---

<sup>6</sup> <http://www.lboro.ac.uk/departments/lis/disresearch/romeo/>

participation. However, none of the literature includes comparative studies of IR development.

Gibbons (2004, p. 6) identifies five core features of IRs: “digital content; community driven and focused; institutionally supported; durable and permanent; and accessible content.” Since IRs are institutionally defined, Lynch’s 2004 description as a “set of services” provides a general overview of what an IR is, while Gibbons describes features included in the services.

### **2.3.1 Models**

Crow (2002, p. 7) identifies four components of scholarly publishing: registration, certification, awareness and archiving. Current IR models are able to provide some, but not all, of the components. Simply entering the research into the IR will register the concept by date stamping its inclusion in the database. The interoperability and open access will allow search engines to access the content when it is added to the IR. Inclusion in an IR provides researchers with a permanent archive. It is assumed the institution will maintain the accessibility of the research by updating formats and/or providing a stable storage location and medium. Problems arise with certification. Most academic journals have a tradition of peer review built into the acceptance process, however; IR software has yet to develop such a function.

Pinfield et al. (2002), on the other hand, discuss the Eprint pilot projects at the Universities of Edinburgh and Nottingham. The discussion focuses on document types, preservation, submissions and meta-data. The authors also discuss the costs involved with IR development. EPrints, like DSpace, are open source; the program is freely available so

that programmers can read, redistribute, and modify the source code for a piece of software, and the software evolves. Users are expected to make additional code freely available to other users. According to Pinfield et al. (2002), the cost of initial set-up is staff time. The IR requires one to two days for software installation and another three days for Web interface customization. Though no specific monetary costs are discussed, the authors identify managing the IR and the costs associated with encouraging faculty participation as the largest ongoing expenses in an IR budget.

### **2.3.2 Software**

Both open source and proprietary software are available for IR development. Software choices include: EPrints, DSpace, Fedora and Proquest's Digital Commons, formerly Bepress. According to recent Registry of Open Access Repositories (ROAR) data, over 30% of the registered repositories worldwide are located in the United States (26.3%) and Canada (4.3%). Of the repositories listed in the United States, in ROAR, DSpace and Digital Commons software are used more extensively for institutional repositories, while EPrints is used for subject-specific collections and other publications (Table 3).

**Table 3. Repository Software Use**

Repository Software	Worldwide	%	United States	US%	Canada	CDN%
GNU EPrints	213	27.9	52	25.9	13	50.0
DSpace	179	23.4	58	28.9	11	42.4
Bepress (Digital Commons)	51	6.7	41	20.4	1	3.8
ETD-db	23	3.0	12	6.0	1	3.8
OPUS	21	2.7	0	0.0	0	0.0
DiVA	15	2.0	0	0.0	0	0.0
CDSWare	9	1.2	0	0.0	0	0.0
ARNO	6	0.8	0	0.0	0	0.0
HAL	6	0.8	0	0.0	0	0.0
Fedora	5	0.7	1	0.5	0	0.0
DoKS	3	0.4	0	0.0	0	0.0
EDOC	2	0.3	0	0.0	0	0.0
Open Repository	2	0.3	0	0.0	0	0.0
MyCoRe	1	0.1	0	0.0	0	0.0
Other Systems	228	29.8	37	18.4	7	
Total Registered Repositories	764	100	201	100	33	100

\* Data from Registry of Open Archive Repositories. (2007). Retrieved November 10, 2006, from <http://roar.eprints.org/>.

Of the documentation currently available, most thoroughly describe the work flow processes within the repository software but spend little time defining the entire project process. Although colleges and universities do not fall into a single operating platform, there should be enough similarities between institutions to develop a basic project plan. DSpace is the only site with any project planning steps; however, the steps they identified may be too general to provide assistance to other institutions (Figure 4).

**Figure 4. DSpace's Project Planning Timeline**

Project Planning Timeline					
	Early weeks	Beginning phase	Before launch date	Launch	Long-term events
Define service definition					
Assemble a team					
Choose technology					
Early Adopter programme					
Marketing the service					
Launch events					
Run the service					

MIT Libraries. (2004). *Work Sheet: Project Planning Timeline*. Retrieved June 15, 2006 from <http://dspace.org/implement/project-plan.pdf>.

For example, the timeline developed by MIT does not include information about stakeholder groups other than the needs assessment during the Service definition phase. Marketing and training are incorporated into the timeline, but assessment is not discussed.

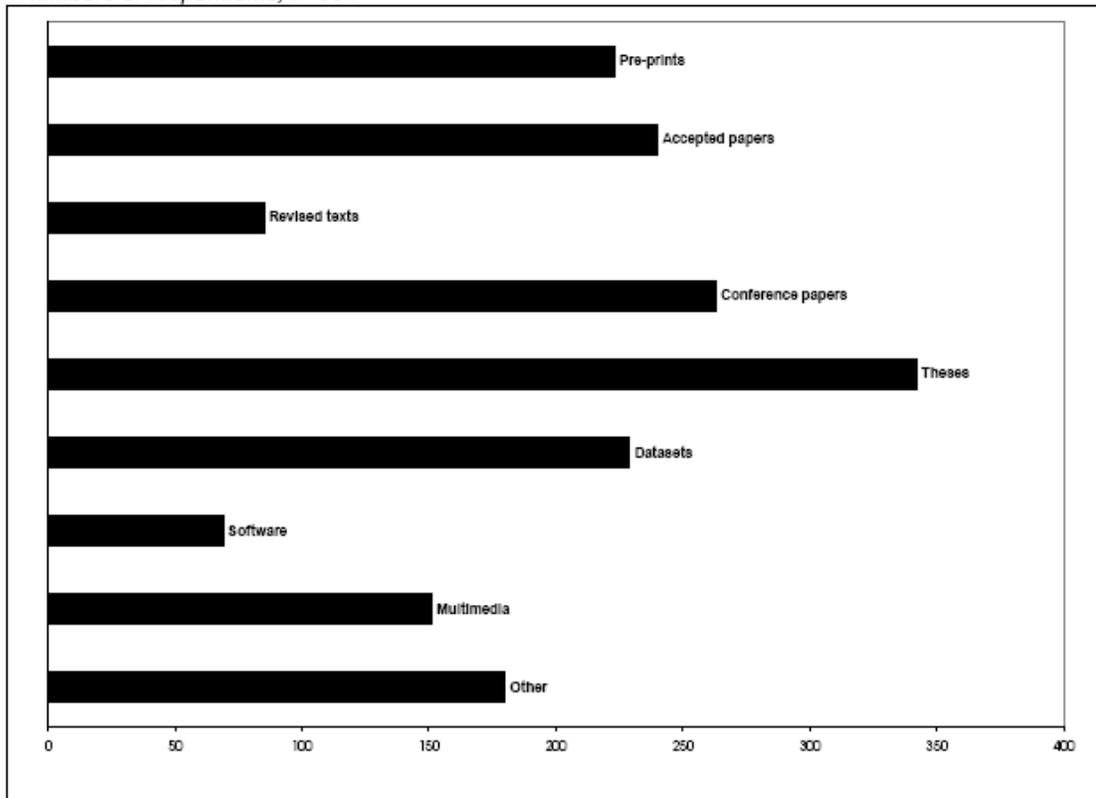
### **2.3.3 Permanence**

Johnson (2002) identifies IRs as a collection that is both cumulative and operates in perpetuity. As such, colleges and universities must have specific policies in place to ensure access to documents. Johnson (2002) notes that removing a document from an IR is “the functional equivalent of revoking the registration initially granted to the contribution on accession into the repository,” (p. 5) something only done in rare cases. However, current practice allows for the removal of items and, in some cases, the limits on access for certain time periods.

### **2.3.4 Content**

A 2004 study commissioned by the Publishers' Association surveyed 4,000 researchers from ninety-seven countries about scholarly communication (Rowlands, Nicholas, & Huntingdon). The survey found that authors do not know much about open access and are unsure about quality and permanence of open access. IRs are able to fill this void by creating trust and maintaining access to deposited items. The repository is, in fact, a reflection of the university's image and status. Out of 3,787 initial respondents, the authors found that 32% of respondents had published on the Web, while more than 53% were considering Web publishing (p. 19). When questioned about institutional repositories, 26% of respondents had deposited material in an institutional repository, while 55% planned to deposit material in the future (p. 20). Content for repositories is often difficult to acquire. Genoni (2004) suggests that collection development plans should be developed to prioritize content added to a collection.

**Figure 5. Academic Publishing in Institutional Repositories**  
Numbers of respondents, n=804



From Rowlands, I., Nicholas, D. and Huntingdon, P. (2004). *Scholarly Communication in the Digital Environment: What do authors want?* Retrieved September 14, 2005, from <http://www.homepages.ucl.ac.uk/~uczciro/ciber-pa-report.pdf>, p. 21.

While preservation and access are two commonly cited benefits, many sites have not developed a preservation plan. The Deep Blue repository at the University of Michigan is an exception, with a highly detailed preservation plan. Three levels of support are identified for scholarly materials, and best preservation practices are listed for various file formats (Deep Blue, n.d.). IRs provide a visible record of a university's scholarly output, wider distribution of scholarly materials and a site for alternative publishing (Gibbons, 2004). As better search engines like Google Scholar are developed, the ability to access these materials increases.

### **2.3.5 Community-Driven**

IRs are driven by the campus community and can become a forum for a campus community of practice (CoP). According to Lave and Wenger (1991), a community of practice is “a set of relations among persons, activity, and world, over time and in relation with other tangential and overlapping communities of practice” (p.98). Occurring within a particular domain, a knowledge sharing community evolves through trust and respect (Wenger et al, 2002). By sharing information, stories, language and resources, the community negotiates normative behaviors such as reciprocity (Wenger et al, 2002; Wasko & Faraj, 2005). Communities of practice provide a way to assimilate new members into a particular profession, much like an apprenticeship (Lave & Wenger, 1991). A community of practice is a “flexible group of professionals, informally bound by common interests, who interact through interdependent tasks guided by a common purpose thereby embodying a store of common knowledge” (Jubert, 166). A CoP serves as a “living repository” for the knowledge of a group, which develops to make tacit information held by individual members explicit (Wegner et al, 2002). In a similar vein, institutional repositories make explicit the scholarly output of an institution that is tacit in specialized journals, information that may not be available or known to the greater campus community.

According to Hodkinson (2004), new researchers learn how to “judge research through engaging in the authentic practices of whichever research community they join” (p. 13). The ability to write for academic journals and create research proposals is a skill that is not formally taught in most academic programs, but something that is gained through experience in the community (Hodkinson, 2004). IRs create a campus-wide

community, beyond the school or department level, to enhance scholarly communication. Information resources, emotional resources and identity support are all benefits for members of a community of practice (Finholt & Sproull, 1990). When researchers begin working with advisors on specific projects, they are entering a scholarly community of practice. Once researchers obtain professional positions, virtual communities of practice are often their only links to the greater community of scholars in a particular field. In addition, the information supplied by discussion lists and bulletin boards enhances other sources of scholarly information like e-conferences (Kovacs, Robinson & Dixon, 1995). Thus the development of an IR benefits the entire university, not just the researchers who participate.

Three views of these communities (information-based, technology-based, and culture-based) apply to repositories (Alavi & Leidner, 1997; Hildreth, 2004). IRs are information based, providing access to and preserving scholarly output on campus. By creating an electronic location for scholarly output, an IR is technology-based. The repository is culture based because it identifies faculty and student research and showcases research projects on campus.

Three main differences exist between electronic and face to face communities: location is not important, participants may not be identified, and the social and logistical cost of participating is low (Sproull & Faraj, 1997). Virtual communities like an IR have to overcome many problems associated with these differences. Problems develop because of distance (Wenger et al, 2002). In virtual communities, the members may be too remote or have difficulty creating an online presence. Size may create issues as well. Without face-to-face meetings, it is difficult to know exactly how many people are affiliated with

the group and members may have different goals or boundaries that will affect community development. Virtual communities allow specialized groups that are not collocated to communicate at low cost. This is especially true of academic virtual communities that allow members of a specialized field to share information and resources worldwide. Wasko and Faraj (2005) define “an electronic network of practice as a self-organizing, open activity system focused on a shared practice that exists primarily through computer-mediated communication” (p.37). While academic communities can be referred to as networks of practice, they more closely resemble communities of practice by combining various levels of academics and allowing members to share information.

While many communities of practice revolve around commercial product design, common problem or a particular profession, an academic CoP provides members with an intercollegiate network of peers and increased research and publication opportunities (Lowrie & McKnight, 2004). The CoP provides a forum to increase access to position announcements, calls for papers, conference opportunities and potential research partners for scholarly publication (Granovetter, 1973; Lowrie & McKnight, 2004). Academic communities of practice typically fall into the information-based category, where information between academics is distributed.

Lawrence (2001) investigated the impact of online articles by analyzing citation rates. Although the articles were not in a specific collection or tool, the study found that online availability facilitates access and use. Repositories offer researchers another venue to publicize their work. The added interoperability and federated searching functions also allow access through a variety of search engines and databases, increasing the potential audience for an article or other scholarly output.

### **2.3.6 Growth**

According to Swanepoel (2005), IRs are still in a growth phase. Current studies are underway, such as a research project conducted at the University of Michigan on IR user behavior. The building and development of an IR has yet to be addressed empirically in the literature.

### **2.3.7 Development**

By identifying how institutional repositories are developing and the challenges that they face, a developmental framework can be identified for libraries interested in IRs. Examining the process of developing an IR at several institutions can provide a generalized view of how an IR is developed. According to *A Guide to the Project Manager's Book of Knowledge*, "the better you know your project, the better you are able to manage it" (Project Management Institute, 2000, p. 6). This can be done by identifying key areas, activities and stakeholders and applying a project management context, thus developing a project process.

While Cervone (2007) discusses the use of a project charter for repository development, the focus is on the technical infrastructure needed for a repository, not the services needed to develop repository content.

In order to provide a project framework for institutions developing IRs, this study will survey those currently working with IRs to identify key project management processes. In addition, the survey will identify various steps, issues and concerns that librarians developing IRs should be aware of prior to the project's inception.

## **Chapter 3. Methodology**

### ***3.1 Case Study Literature***

Associated primarily with the fields of anthropology and sociology, case studies seek to provide an in-depth description of the features or attributes of a particular phenomenon (Hamel, Dufour & Fortin, 1993). The early use of this form of research can be traced to Europe, predominantly to France. In the United States, this methodology linked closely with the University of Chicago Department of Sociology in the early 1900s.

Case studies can be single or multiple-case designs, where a multiple design must follow a replication rather than sampling logic. When no other cases are available for replication, the researcher is limited to single-case designs. Yin (2003) pointed out that generalization of results, from either single or multiple designs, is made to theory and not to populations. Multiple cases strengthen the results by replicating the pattern-matching, thus increasing confidence in the robustness of the theory. Applications of case study methodology have been carried out in libraries by several researchers

Over the years, case study research has fallen in and out of favor; primarily because of the lack of quantitative measurements in the research design and analysis (Hamel et al, 1993). A single case or observation point, unless carefully selected, may fail to represent the object of study, while dependence on a single case renders a study incapable of providing a generalizable conclusion. Giddens considered case methodology "microscopic" because so few cases were studied (Yin, 2003). Hamel et al. (1993) and

Yin (2003) both argue that the size of the sample does not transform multiple cases into a macroscopic study. Instead, a case study should establish the parameters which allow a single case to be considered acceptable (Tellis, 1997).

As the use of quantitative methods advanced, case study research declined. With concern about the limitations of quantitative methods, interest in case studies was renewed in the 1960s, especially after the development of grounded theory (Strauss and Glaser, 1967). According to Borgatti (2005, Goals and Perspective section, ¶ 1), grounded theory “refers to theory that is developed inductively from a corpus of data.” Using grounded theory, a researcher “attempts to derive a general, abstract theory of a process, action or interaction grounded in the views of participants” (Creswell, 2003, p. 14).

### ***3.2 Case Study Design***

Yin (2003) identified five components important in case study design: a study's questions; its propositions, if any; its unit(s) of analysis; the logic linking the data to the propositions; and the criteria for interpreting the findings (p. 21). The development of "how" and "why" questions and their definition is the first task of the researcher. It captures what one is interested in studying. An exploratory study, rather than having propositions, has a stated purpose or criteria on which success will be judged. The unit of analysis can be an individual, an event or process.

Construct validity, the extent to which a measurement corresponds to the concepts under study, is problematic in case study research. Yin (2003) proposed using multiple sources of evidence, establishing a chain of evidence, and having a draft case study report

reviewed by key informants to mitigate the problem. Internal validity, establishing causal relationships, is a concern only in causal cases, not descriptive or exploratory ones.

Reliability is achieved in many ways in a case study. One of the most important methods is the development of the case study protocol and the consistent use of uniform procedures.

Case studies can be either single or multiple-case designs. Single cases are used to confirm or challenge a theory or to represent a unique or extreme case (Yin, 2003).

Single-case studies are also ideal for revelatory cases where an observer may have access to a phenomenon that was previously inaccessible. Single-case designs require careful investigation to avoid misrepresentation and to maximize access to key documents and people. Multiple-case studies follow replication logic. Each individual case study consists of a "whole" study, in which facts are gathered from various sources and conclusions drawn from those facts.

A case study protocol, essential in a multiple-case study, contains more than the survey instrument; it should also contain procedures and general rules that are to be followed during the study. The protocol, according to Yin (2003), is a major component in demonstrating the reliability of the case study research. A typical protocol should have the following sections:

- An overview of the case study project (objectives, issues, relevant readings);
- Field procedures (credentials and access to sites, sources of information);
- Case study questions (specific questions that the investigator must keep in mind during data collection) ; and
- A guide for case study report (outline, format for the narrative)

(Yin, 2003, p. 69).

The overview provides an introduction to the project and the issues being investigated. The field procedures involve data collection issues, as well as ensuring that people and resources are available for the study. Open-ended interviews should be planned based on the participants' schedules (Stake, 1995). Prior to the interview, the researcher must gain access to the subject organization, have sufficient resources while in the field, clearly schedule data collection activities, and provide for unanticipated events. Case study questions serve to guide the interviewer and keep the investigation on track.

Stake (1995) and Yin (2003) identify sources of evidence in case studies as documents, archival records, interviews, direct observation, participant observation and physical artifacts.

This study of institutional repositories relies heavily on qualitative methods of inquiry. Creswell (1998) defines qualitative research as:

Qualitative research is an inquiry process of understanding based on distinct methodological traditions of inquiry that explore a social or human problem. The researcher builds a complex, holistic picture, analyzes words, reports detailed views of informants, and conducts the study in a natural setting. (p. 15)

The case study methodology was chosen for its focus on context and “lessons learned” (Lincoln & Guba, 1985). By interviewing those involved with IR development at several institutions, noticeable patterns or regularities of IR development became apparent. Using multiple cases, developed in different contexts, expands the generalizability of the findings (Yin, 2003). In addition, multiple cases expand the usefulness of the research that are planning and developing IRs, especially since much of the information about development is local and not published.

The unit of analysis for the study is the IR. A comparative case study analysis approach was employed to gather and analyze data, and provides a detailed account and analysis of six academic institutional repositories.

### ***3.3 Research Questions***

The research questions that guide this study are:

1. What are the factors influencing IR development?
2. How are IRs developing?
3. How should IRs develop?
4. What are the best practices for IR development?
5. What is the best model(s) for IR development?

### ***3.4 Definitions***

*Institutional Repository* - In this study, an institutional repository will be defined as: an organized collection of digital objects, pertaining to a particular research or educational organization (Jones et al., 2006). A SPARC position paper further defines an IR as a “digital archive of the intellectual product created by the faculty, research staff, and students of an institution and accessible to end users both within and outside of the institution, with few if any barriers to access” (Crow, 2002, p. 16).

*Development* - The act, process, or result of developing. Gibbons (2004) identifies six core functions of an IR as: materials submission, metadata application, access control, discovery support, distribution and preservation. Questions about each of these functions will illuminate development for each IR.

- *Materials submission* identifies how authors are submitting content and whether the IR has an editor or any submission rules.
- *Metadata application* addresses who adds the metadata and what types of metadata are incorporated in the IR.
- *Access control* includes digital rights management, authentication, and identity management. Inherent in access control are the rights to add, delete, approve and edit content.
- One of the key benefits of IRs is the increased visibility of content through *discovery support* (Gibbons, 2004; Bailey, 2006). Discovery support identifies how the IR content is searched.

- *Distribution* of content identifies any limits on viewing content, as well as available file types and the plug-ins needed to view it.
- *Preservation* identifies both long- and short-term preservation policies, ranging from format conversions to content back-ups. Overall, these core functions are “essential” to an IR (Gibbons, 2004, p. 7).

By identifying how libraries are defining these functions, insight is gained about how they develop IR collections.

Development also includes resources, responsibility, and management.

- Resources are defined as the staffing, software, equipment, and maintenance needed for IR development. Resources closely parallel the costs identified by Gibbons (2004): staffing, software, maintenance, content recruitment, and equipment.
- *Responsibility* is the state or fact of having a duty to something or of having control over someone. While the sense of duty and control is important, in this case the Hawaiian word *kuleana* may better add to the definition. The definition of *kuleana* in the *Hawaiian Dictionary* includes: “right, privilege, concern, responsibility, portion, jurisdiction, authority, liability, interest, claim, ownership, province; reason, cause, function, and justification” (Pukui & Elbert, 1986). The research questions in this case would involve *kuleana* as well.

The management of the repository includes individuals involved in the day-to-day management, the department responsible for managing the project, and the task roles of those involved.

- *Library departments* in this study will refer to technical services, public services, and archives.
- The *Dictionary of Business and Management* by Oxford University Press uses the term *task roles* for “a set of coherent roles that are often adopted by the different members of a group in order to solve problems, make decisions, and meet targets.”

*Factors* - Throughout the literature and in the preliminary interviews, several factors were identified as influencing IR development. These factors include: stakeholders involved in the IR process, planning, assessment, content, policy development, marketing, costs, sustainability, motivation and perceived IR benefits.

- *Stakeholders* - Stakeholders include all of those with an interest in an organization. In this instance, it includes users and content owners as well as library and university administration.
- *Planning* - Planning is defined by the *Merriam Webster Dictionary* as the establishment of goals, policies, and procedures for a social or economic unit. The *planning process* “involves doing something that has never been done before” (PMBOK, 2000, p. 32). Although numerous academic institutions have implemented IRs, an empirically developed, generalized project model has not been identified. The core processes or activities that occur in this stage include: scope definition, activity definition and sequencing, resource planning, risk planning, cost estimating and schedule development (PMBOK, 2000).

- *Content* - What the repository collects will often inform the policies it develops for content. There are often multiple digital collections at an institution, ranging from theses and dissertations to special collections to online texts.
- *Policy development* - The policies an institution develops define who can submit content, what that content shall be and who can access it. In addition, policies cover versioning, approval process, and other key functions of a repository (Gibbons, 2004).
- *Assessment* - According to the *Oxford English Dictionary*, assessment is: To evaluate (a person or thing); to estimate (the quality, value, or extent of), to gauge or judge. In this context, assessment is the process of observing, describing, collecting, recording, scoring, and interpreting information about a project in order to judge its value.
- *Marketing* – In this case, marketing refers to advertising and promoting IRs. Preliminary interviews found that the institutions surveyed did little to market the repository to faculty for content recruitment. These same institutions were having difficulty getting content. Content recruitment was also identified as one of the biggest challenges in IR development by Bailey (2006).
- *Costs* - Bailey (2006) places the cost of IR development at \$100,000 with an additional \$100,000 needed each year for maintenance and preservation. Identifying how libraries cover these additional expenses becomes an issue. As well as identifying the possible costs listed above, Gibbons (2004) identified three possible models: 1) the library assumes all the costs, 2) the library assumes core

costs, and “ancillary” costs are charged to the content providers’ department, and 3) individual content providers are charged for each deposit (p. 56).

- *Sustainability* - Linked with preservation, IRs are intended to provide long-term access to an institution’s scholarly output. Several of the preliminary interviewees mentioned it was unclear if the project would continue. Campus support for IR funding, as well as faculty participation, may hinge on whether an IR will continue into the future.
- *Motivating Factors* –Motivating factors are the internal and external forces and influences that drive an individual or an organization to develop an institutional repository. Bailey (2006) identified several motivating factors in the ARL survey, but 22% of responses were listed as *Other* (Table 11).
- *Perceived IR benefits* - According to Bailey (2006), benefits and motivation are related. Gibbons (2004) identifies IR benefits as: stewardship, efficiencies, a scholarly showcase, wider distribution and a response to the crises in scholarly communication. Bailey (2006) identified three main benefits in his 2006 survey: visibility and increased dissemination of the institution’s scholarship; free, open, timely access to scholarship; and preservation and stewardship of digital content (p.20). In addition, the IR benefits the students and education at a university. These benefits relate to core functionality and to an IR’s management and funding within an institution.

*Best practices* - The ODLIS defines best practices as “the application of theory to real-life situations, procedures that, when properly applied, consistently yield superior results and are therefore used as reference points in evaluating the effectiveness of

alternative methods of accomplishing the same task. Best practices are identified by examining empirical evidence of success.”

### ***3.5 Preliminary Interviews***

I conducted preliminary interviews with IR developers at six colleges and universities throughout the United States to test and revise interview questions. I identified a sample of preliminary sites from different Carnegie Classifications was chosen to see if the size or type of institution affected IR development. According to the Carnegie Foundation, its classification system “has been the leading framework for describing institutional diversity in U.S. higher education” (Carnegie Foundation). Using ROAR data, I identified eight sites for a pilot inquiry. The repositories at the selected sites serve the entire campus, rather than a department or school, and were located in the United States. Two colleges, two masters institutions, two doctoral intensive and two doctoral extensive academic institutions were chosen for preliminary interviews. Seven sites consented to my request for phone interviews about IR development.

The purpose of the preliminary interviews was to (1) identify factors that influence IR development; (2) capture the language being used by IR developers; and (3) review and revise the draft interview questions. Synopses of the preliminary interviews are available in Appendix B.

**Table 4. Preliminary Interview Summary**

<b>Carnegie Classification</b>	<b>Baccalaureate - liberal arts (Midwest)</b>	<b>Baccalaureate - liberal arts (Northeast)</b>	<b>Masters' Universities &amp; Colleges I (South)</b>	<b>Masters' Universities &amp; Colleges I (Midwest)</b>	<b>Doctoral/ Research Intensive (East)</b>	<b>Doctoral/ Research Extensive (Midwest)</b>	<b>Doctoral/ Research Extensive (South)</b>
<b>Public/Private</b>	Private	Private	Private	Private	Public	Public	Private
<b>Depository Software</b>	Digital Commons	Digital Commons	Digital Commons	DSpace	DSpace	Digital Commons	DSpace
<b>Started</b>	May 2004	May 2005	May 2004	October 2004	April 2005	June 2005	2003
<b>Status</b>	pilot	pilot	launched	launched	pilot	launched	pilot
<b>Records</b>	44	6,000+ <sup>7</sup>	104	130	18	1,400+; 9,500+ dissertations from UMI content	80
<b>Current Content</b>	Student thesis (maybe only Honors)	Theses; philosophy journal/ lectures	Masters' theses; some University documents	Anything associated with the university	Theses, dissertations, research	Anything associated with the university	Anything associated with the university
<b>Content Policy</b>	no	no	developing	no	no	yes	developing
<b>Evaluation</b>	no	no	no	no	no	no	no
<b>Marketing</b>	no	no	no	yes	no	no	yes
<b>Training</b>	no	no	no	yes	no	no	no

<sup>7</sup> No full text theses was available at the time of the interview. Increased visibility of the indexed theses has led to an increase in thesis requests at the site.

### **3.5.1 Preliminary Interview Discussion**

While conducting the preliminary interviews, it became clear to me that IR developers were not aware of many of the factors that would influence the development of the IR at their institutions. For example, several institutions neglected to include someone from Academic Affairs on the IR committee. At some institutions, permission from Academic Affairs is needed to incorporate student research, such as theses and dissertations, into a repository. Although a large amount of content exists on campus in a digital form, many sites had difficulty identifying who owned the copyright. Electronic access to journals and conferences may be archived at an institution, but the rights signed by authors may not include distribution through a repository. If copyright ownership is unclear, it becomes necessary to contact individual authors for permission, with clerical time and postage increasing IR costs.

The individuals developing content were having more difficulties developing the IR than those charged with platform development. Open-ended questions prompted in depth answers, participants were interested in discussing their experiences, often seeking advice on how to proceed with the IR projects. In some cases, the discussion of library-related problems outside of IR development prompted the participants to discuss similar problems with IR development.

IRs developed by one person, rather than a project team, seemed to have more difficulties identifying resources available on campus and recruiting faculty participation. Although other librarians volunteered their services at these sites, formal development was left to one person. IRs developed by a project team, on the other hand, had created a structured environment for development that mitigated some of the problems.

Lessons learned from the preliminary interviews include:

- Define what an IR is and make sure those involved in the project have the same definition.
- Develop a project plan.
- Define content early in order to identify necessary permissions and stakeholders.
- Review the IR software and know its functionality and limits prior to discussing options with faculty.
- Include campus administration on the project team.

### ***3.6 Sample Selection***

I investigated a variety of information sources in order to identify potential sites for study. Institutional repository sites were of primary interest.

*Institutional repositories sites.* A number of Web sites associated with open access and institutional repository development maintain links to sites with institutional repositories.

The links at these sites were explored in an effort to discover potential sites for study.

Links were identified through:

- The Registry of Open Access Repositories (ROAR)<sup>8</sup>
- ProQuest Digital Commons<sup>9</sup>
- DSpace Federation<sup>10</sup>
- EPrints<sup>11</sup>

### ***3.7 Selecting Sites for Study***

A list of repositories at the university level was compiled from the various sources. The study initially identified IRs that contained 2,000 – 9,000 items in the United States. According to ROAR, only six university repositories had more than 9,000 items in December 2006. These sites were removed as study sites since they were either 1) principal developers or removed from the development process because they reached a critical mass of articles, 2) may have acquired content from providers, like ProQuest or Elsevier or 3) participated in the preliminary interviews. Ten university repositories from

---

<sup>8</sup> <http://archives.eprints.org/>

<sup>9</sup> [http://www.umi.com/products\\_umi/digitalcommons/](http://www.umi.com/products_umi/digitalcommons/)

<sup>10</sup> <http://www.dspace.org/>

<sup>11</sup> <http://www.eprints.org/>

the 2,000 to 9,000 item group were emailed and asked to participate in this IR development study. Fearing that the sample was too small, I emailed the sites with identifiable contact information and fifteen additional repositories at various stages of development as well.

From the 2,000 – 9,000 record groups, repositories at Purdue University and Oregon State, Georgia Tech, the University of Connecticut and the University of Maryland responded and were interested in participating in my study on institutional repository development. Columbia University, IUPUI and Texas State University from the fifteen repositories listed on ROAR with contact information also expressed interest.

After comparing the list of respondents with the location of other IRs in development, I chose to focus research in Indiana, Washington and British Columbia because this would provide six institutional repositories and three subject repositories located at institutional repository sites. Indiana University, originally non-responsive, was contacted a second time and agreed to participate as a case study site. Contact information for the University of Washington was also identified and librarians at the site agreed to participate as well. The University of British Columbia (UBC) and Simon Fraser University (SFU) were not initially included in the study, but were added because 1) the location was convenient to the researcher in Vancouver, BC; 2) the UBC is initiating an IR and it would be possible to document early decision making process to some extent; 3) Adding SFU and UBC widens the geographic area to include Canadian schools; and 4) Comparing two Canadian institutions may identify any site anomalies that may be identified as Canadian in nature.

Indiana University, the University of British Columbia and the University of Washington are also home to department level or subject repositories registered in ROAR. Including IRs developed or developing at the department or subject level provides insight into faculty participation. If sponsored by a particular department or in a research field, the development factors may be quite different. For example, faculty participation is expected to be higher if one's department or school is running the repository.

Repository administrators at the IUBioArchive (IU) and the Centre for the Study of Historical Consciousness (UBC) answered questions about repository development. The administrator at the Health Sciences Libraries (UW) had recently left the institution, but a library employee was able to comment on the project.

The UBC repository is a pilot project; the other eight sites were live at the time of the interviews. Eprint, Digital Commons, PKP and DSpace sites are represented in the study. Although all of the IRs are at institutions that offer PhD programs, preliminary research showed that regardless of size, librarians were experiencing similar issues during repository development.

### ***3.8 Case Study Sites***

The sites for this study are: Purdue University, Indiana University, Indiana University-Purdue University Indianapolis (IUPUI), Simon Fraser University, University of British Columbia, and the University of Washington. The case studies included site visits for interviews, documents and archival records where collected during the visits and from each university website.

**Table 5. Case Study Sites (December 2007)**

University	Students <sup>1</sup>	Repository	Software	Number of Items <sup>2</sup>
Purdue University	38,712	Purdue e-Pubs (part of Purdue e-Scholar)	Digital Commons	6,674
IUPUI	29,933	IDeA	DSpace	384
Indiana University	37,958	IUScholarWorks	DSpace	260
		IUBio Archive	E-Prints	494
Simon Fraser University (SFU)	26,670 <sup>12</sup>	Simon Fraser University Institutional Repository	DSpace	1,127
University of British Columbia	43,000	cIRcle	DSpace	0
		Centre for the Study of Historical Consciousness	PKP	23
University of Washington	40,216	DSpace at The University of Washington	DSpace	2,229
		DSpace at the University of Washington Health Sciences Libraries (HSLIC)	DSpace	242

<sup>1</sup> Based on most recent enrollment data available.

SFU (2004/2005): [http://www.sfu.ca/mediapr/for\\_media/sfu\\_facts.html](http://www.sfu.ca/mediapr/for_media/sfu_facts.html)

UBC Vancouver (2004/2005): <http://www.publicaffairs.ubc.ca/ubcfacts/index.html#students>

Purdue West Lafayette (Fall 2005): <http://www.purdue.edu/Purdue/about/index.html>

IUPUI (Fall 2005): [http://factbook.indiana.edu/~urr/factbook/fbook05/fast\\_facts/fastfacts3.shtml](http://factbook.indiana.edu/~urr/factbook/fbook05/fast_facts/fastfacts3.shtml)

Indiana University (Fall 2005):

[http://factbook.indiana.edu/~urr/factbook/fbook05/fast\\_facts/fastfacts2.shtml](http://factbook.indiana.edu/~urr/factbook/fbook05/fast_facts/fastfacts2.shtml)

University of Washington (2006-2007): <http://www.washington.edu/admin/factbook/>

<sup>2</sup>As of November 3, 2006.

Purdue University, Indiana University, the University of British Columbia and the University of Washington are institutions similar in size and scholarship. In addition to the convenience of the sites, the sites are at various stages in development. This allowed the researcher access to different stages of development and the ability to verify through

<sup>12</sup> Enrollment may include students from Vancouver (2,200 students) and Surrey (1,200 students) campuses.

interviews whether the development experiences between institutions are similar. In addition, the researcher was able to compare IRs developed using DSpace to those using Digital Commons software.

IUPUI and Simon Fraser are institutionally similar as well. Both sites offer fewer than 50 doctoral degrees, have similar-sized library collections, and were founded in the 1960s (SFU Facts, 2006; SFU Calendar, 2006; IUPUI, 2006; Indiana University, 2006).<sup>13</sup> Both sites are expanding and have smaller associated campuses. Although the sites are at different stages of IR development, the sites are comparable.

Finally, the inclusion of non-institutional repositories provides data about smaller collections. These subject and departmental repositories include the IUBioArchive, DSpace at the University of Washington Health Sciences Libraries, and the Centre for the Study of Historical Consciousness at the University of British Columbia. The case studies answered the following questions:

- Are subject repositories comparable?
- Are the repositories comparable to institutional repositories?
- Are there any lessons learned that are applicable to IR development?
- How do repository managers get faculty members to participate?

---

<sup>13</sup> According to the sources cited in the text, the IUPUI library collection contains 1,659,129 items, while the SFU collection contains approximately 1.4 million.

### ***3.9 Data Collection Methods***

This section describes the methods used to gather data for this study. These methods included: site visits; semi-structured interviews and informal discussions; and collection and analysis of the organizational, operational, and archival documents.

#### **3.9.1 Semi-Structured Interview**

A semi-structured interview approach was selected to address the ambiguous language of institutional repositories as well as local terms and meanings that may be unique to the individual cases.

Based on the preliminary interviews, questions were developed to guide interviews and identify the process and procedures in place during IR planning, as well as to collect information about the perceived usefulness of various project management techniques. The questionnaire consisted of both quantitative data collection and open-ended questions that allowed for participant response (Creswell, 2003; Appendix A). Interviews were conducted on site with participants. Site visits allowed the researcher to probe for more details. Although face-to-face interviews can result in interviewer bias, they provide visual cues or aids to the discussion. Telephone interviews and email were used to clarify information and to contact additional people involved in IR development. In order to prevent misunderstanding or misinterpretation, transcripts of the interview were sent to interviewees for accuracy. Not all questions in the guide (Appendix A) were asked to every participant. Every question in the guide, however, was asked at each institution. In group interviews, all questions were asked; however, only questions relevant to a participants role or experience were asked during individual interviews.

Interviews with repository developers were taped and transcribed. Transcripts of the interviews were sent to participants to ensure the content and contexts were correctly recorded by the researcher. An interview with a faculty member at Indiana University was not taped, but the interview was summarized and sent to the faculty member for review. As a professional courtesy, the final case report was sent to repository developers at each site to verify the facts of the case, not the conclusions (Yin, 2003). Not only did this allow for corrections and additions, but it also increased construct validity.

### **3.9.2 Data analysis**

After I transcribed the tapes, all of the interviews were placed into a single MS Word document ensuring the combination of page number and line number is unique and identifiable for re-interpretation. Comments, that clarified content or identified an area of interest, were also added to the Word document. These included comments and notes that I made during the site visits.

A MS Access database was populated with fields including:

1. *unique ID* (auto-generated by MS Access);
2. *name* of the participant;
3. *role* of the participant;
4. *organization* of the participant;
5. *theme* that this data record fit in;
6. *coding* based on research questions;
7. *page* and *line number* from the MS Word document; and
8. *observed data* from the textual content from the transcriptions.

I created over 400 records using the themes and codes developed from the questionnaire (Appendix E). The observed data was kept in context by including surrounding text that clarified the content. Keyword searches of the documents identified similar content within the Word document. I then sorted the content in MS Access in order to identify themes. Categories and subcategories were identified within the themes to further sort the data.

Once populated, I sorted the records by institution, theme and category and used the resulting table to develop case studies. I also returned to the transcript documents during this period to identify additional content for the database and verify the context of the quotes. For the comparative case studies, I sorted the records by theme and category to compare them across institution.

### **3.9.3 Outline of the case study report**

A multiple-case report was developed for the study. Each institutional repository has its own narrative with a separate section covering cross-case analysis. An additional section contains a brief description and comparison of the subject repositories. The case reports include:

- A description of the IR, providing historical information about its development.
- A discussion of the collection policy and current contents.
- Discussion of the planning process of IR development, including resources and stakeholder groups.
- A discussion of evaluation measures being used or considered.
- University context and history pertaining to the library and/or IR development.

Cases were compared, in order to find patterns across the cases. The patterns were used to develop framework for IR development.

### ***3.10 Limitations***

The sample size of eight repositories is a limit, as is construct validity and developer bias during the participant interview. According to Yin (2003), a single case is considered acceptable, provided it meets the established objective. The variability, descriptions and multiple interviews per site increases validity. Construct validity was increased by providing case study reports and interview transcripts to the participants for review prior to submission. Although face-to-face interviews can result in interviewer bias, they provide visual cues or aids to the discussion. Telephone interviews and email were used to clarify information and contact additional people involved in IR development. In order to prevent misunderstanding or misinterpretation, a transcript of the interview was sent to interviewees for accuracy.

### ***3.11 Anticipated results***

From the literature and the preliminary interviews, it is expected that:

- There are multiple models for development depending on those involved in the project.
- Marketing the IR leads to increased faculty participation.
- Project teams have fewer challenges than individuals responsible for IR development.
- There is a relationship between motivation and who is leading IR development.
- Development models depend on where the IR falls organizationally (e.g., technical services).
- The IR may be competing for resources with other digitization projects.

## **Chapter 4. Case Studies**

### ***4.1 Indiana University***

Well, our institutional repository is a set of services to make the web of IU scholars' work truly available while ensuring that they are preserved and organized for the future. (ID 5)

#### **4.1.1 Background**

Indiana University (IU) founded in 1820, serves 38,990 on its Bloomington campus (Indiana University, 2008). In addition to the Bloomington campus, there are six regional campuses throughout Indiana and a unique relationship with Purdue University running IUPUI. The Indiana University Bloomington (IUB) Libraries have an extensive Digital Libraries Program.

#### **4.1.2 Repository Development**

The IUScholarWorks Repository<sup>14</sup>, Indiana University's DSpace repository, began with an official pilot project in the spring of 2006 (ID 39). By the summer of 2006, the repository was considered "a production service" and was registered with various directories (ID 39). At this point, repository developers began actively seeking content (ID 39). The repository was developed in response to the open repository movement; a previous dean was interested and thought the library should be involved (ID 56).

IUScholarWorks is a set of services developed by the IU Libraries and Digital Library Program (IUScholarWorks, 2007). As part of a mature digital library program,

---

<sup>14</sup> <http://scholarworks.iu.edu/>

the repository director is trying to understand “the particular niche or milieu that the institutional repository should fill – [a niche] that is not filled by Mass Store, the Digital Libraries Program, or OnCourse, the course management system” (ID 40). With so many digital projects and storage options available on campus, the DSpace repository is just one of many selections available to faculty, departments and centers on campus.

Initially, the university archives was involved with the pilot project. It was hoped that DSpace could be used to house the articles, lectures and manuscripts of highly visible retired faculty members to create a “portrait” of the faculty member’s scholarly output. Unfortunately, DSpace was “very bad at that” and “...they just stopped completely and decided not to continue; it was just an incredible amount of work” (ID 74).

The Assistant Director of Collection Development and Scholarly Communication spends 50% of allocated time managing IUScholarWorks. The Assistant Director has been with the library for twenty-five years, and has held positions in “public services or automation and technology” (ID 64). In addition to library experience, she has “been active in the Bloomington Faculty Council” and has “lots of faculty contacts and recognition” (ID 64). Although the position reports to the Director of Collection Development, interacting with faculty is a large part of the position (ID 68).

Initially, the repository director worked with a reference librarian, “one of these people who are always finding new and incredibly useful uses of technology,” who was assigned 50% to the project (ID 69). She is currently assisted by a School of Library and Information Science (SLIS) Graduate Assistant for fifteen hours a week and a programmer for ten hours a month (ID 69, 78). Like many repositories, other library staff

members provide assistance to the director as needed, but do not commit a percentage of their time to the project. This includes a “metadata librarian... [who] keeps us on track with using Dublin Core and the appropriate use of metadata” and a cataloger, whose role hasn’t “become clear” yet (ID 68). Further assistance about licensing and rights was provided by the University Counsel (ID 69).

### **4.1.3 Content**

One of the things that makes it so difficult for me to develop a sort of *raison d’être* for this, one of them, as I have said before, is the existence of a mature digital library program. (ID 66)

The repository director sees the initial decision not to limit the collection to peer-reviewed publications as the “right choice” and a “big step” (ID 81). Within DSpace, groups are set up for units on campus in addition to locations for individual faculty members (ID 41, 44). The repository director acts as a group editor for many of the collections, adding the documents and associated metadata, and in some cases scanning documents and turning them into PDFs (ID 44, 47). Some of the collections are associated with “quasi-library groups,” like Archives of Traditional Music (ID 48). Many of these are staffed by librarians who “are very careful and very much want to control their metadata and submissions and are doing it all themselves” (ID 48). The repository is focusing on PDFs at the moment, discouraging the submission of video and audio files. These larger files are more difficult to deal with in DSpace and may be better suited in a different collection within the Digital Libraries Program (ID 33).

#### **4.1.3.1 Planning**

Initial planning was conducted by a small team that included the repository director, some people from the Digital Libraries Program and another librarian (ID 70). One of the biggest decisions the team had to make was whether to use an open source or vended product (ID 59). The team compared functionality and estimated costs of different software options with “pros and cons of each of them, and a recommendation” was submitted the library administration (ID 70). The team saw the library as becoming part of the “actual publication processes,” a functionality which was not available with Digital Commons (ID 59). The director is still not sure if that was the right decision, but an open source product “just made a whole lot more sense” with the existing Digital Libraries Program and its technical infrastructure (ID 59).

The Institutional Repositories Development Team did write a business plan for the repository. The plan what resources were devoted to the project, how decisions and policies would be made, who was responsible for reports, and a general timeline for implementation (IUB Libraries Institutional Repository Pilot Team, 2005).

The plan was based on a document developed by Bishoff and Allen (2004) which outlined how to write a business plan for non-profits (ID 71). Bishoff and Allen developed the document for small institutions embarking on digital projects “that have no chance of being sustainable over the long term” (ID 71). The document “was an effort to force [small institutions] to consider those issues before they started the project,” but the document template aided the development of an action plan (ID 71). While writing the business plan was “torture,” it proved an effective way “to move ahead, and [provided] the directors something to approve” (ID 72).

#### **4.1.3.2 Policies**

Our policy is that it's up to the community to decide. And that if they choose to delete an item, we will do it for them... (ID 62)

Rather than set up barriers for groups and departments on campus, policy development has been left to communities (ID 62). As it stands, faculty can deposit anything in the faculty community and graduate students can deposit dissertations. Any other content has to come through a community; this includes content from graduate and undergraduate students (ID 52).

Authors are required to “sign a license that gives Indiana University the nonexclusive right to distribute it in perpetuity, but the rights holder gives up none of the rights that they hold” (ID 38). Individual authors are still able to distribute the work as they see fit.

#### **4.1.3.3 Recruitment**

Research centers on campus, usually considered to be “low hanging fruit,” were contacted as well as departments that had worked with the library previously (ID 41). As of January 2007, the repository director had spoken to over fifty departments on campus, twelve of which are actively participating in IUScholarWorks (ID 41). In addition to these areas, the director has been working on two other areas of content: dissertations and electronic faculty annual reports (ID 37). With change in administration at the Graduate School, the library has “entered into discussions” with the graduate school about placing e-dissertations in IUScholarWorks (ID 421). Currently, students can choose to place

dissertations in IUScholarWorks, but the students must fill out additional forms and register copyright in order to submit (A Guide to the Preparation of Theses and Dissertations, n.d.). The repository director is trying to streamline the process by creating an “interface” for Proquest’s Electronic Dissertation Submission workflow that will allow students to submit a PDF to the repository while submitting to UMI (ID 422, 423). In a similar vein, Indiana University is moving toward electronic annual reports from faculty members (Project: Faculty annual reporting, 2007). The repository director has been “lobbying” to “piggyback” article submission onto the annual report process, so that when a faculty member adds a published article to the report, he or she has the option to upload a copy to IUScholarWorks (ID 37).

#### **4.1.3.4 Content Examples**

The Department of Folklore and Ethnomusicology has posted several collections to IUScholarWorks. In addition to peer reviewed papers, the department is also archiving its syllabi. The IUScholarWorks collection allows prospective students a chance to view current research within the department and examine past syllabi (J. Jackson, personal communication, July 3, 2007). In addition to document and syllabi collections, the faculty in the Folklore and Ethnomusicology Department include locally produced journal content from publications like Folklore Forum in the repository. One of the faculty members, Jason Baird Jackson, was an early adopter and the first to include a researcher page within the DSpace environment.

**Figure 6. Researcher Page in IUScholarWorks**  
**Researcher Page: Jason Jackson**

▼  Research	
<a href="#">Native American Voices on Identity, Art and Culture: Objects of Everlasting Esteem</a>	
<a href="#">Spirit Medicine: Native American Uses of Common Everlasting (<i>Pseudognaphalium obtusifolium</i>) in Eastern North America</a>	
<a href="#">William C. Sturtevant (1926-2007)</a>	
▼  Teaching Related	
<a href="#">FOLK F722 Putting Cultural Theory to Use</a>	<b>About The Researcher: Jason Jackson</b> Associate Professor Department of Folklore and Ethnomusicology Folklore, Anthropology, Museum Studies

IUScholarWorks also houses annual reports created by visiting scholars at the IU Cyclotron. Although it has been difficult to include scientific notation in the metadata, the two to three page reports are digitized by the library. Currently the collection contains reports from 1983-1996, many of which were not previously available online (ID 47).

Also in the collection is the first volume of an undergraduate journal developed in an education class. The professor put together an editorial board of students that selected articles written in the class for publication. Although available in print, the professor was excited to make the articles available online as well. Two of the articles have become the most heavily downloaded in the repository (ID 63).

#### **4.1.3.5 Marketing**

The repository has been marketed to librarians on campus, and a brochure was created for librarians to take to the departments. The collection managers are also “trying to get into faculty meetings” to discuss the repository, but nothing has been done campus-wide (ID 77).

#### **4.1.3.6 Preservation**

At this time, there is not a preservation plan specifically for IUScholarWorks; however, the Digital Libraries Program does have one. The DSpace environment is one of many in the program, all running on separate infrastructures. The goal is to have one repository infrastructure for the entire program; it will be systematically preserved. DSpace is viewed as a temporary application, “getting us to a more integrated point” (ID 53).

#### **4.1.4 Assessment**

Although statistics are being kept and the repository director is trying to identify trends in use and collection growth, these measures may not be “a good use of [the director’s] time,” because the repository is changing (ID 73). She has found that the assessment measures identified by Stevan Harnad, based on the number of peer-reviewed journal articles, do not provide an appropriate measure (ID 6). One measure of success identified is “learning what content is out there, on our campus, that we need to help preserve” (ID 15).

#### **4.1.5 Benefits**

It is possible that the single biggest benefit from this whole project will be discussions with faculty on this topic to inform us and to inform them. (ID 14)

According to the repository director, “the content is accessible to researchers, so it improves the quality of research, and, of course, knowledge” (ID 79). The repository is opening avenues of discussion between librarians, faculty and administration that were not available before.

#### **4.1.6 Challenges**

As previously discussed, some of the biggest challenges involve incorporating the repository into workflows for dissertations and faculty annual reports (ID 21, 37, 421, 422, 423). However, there are also challenges within the library. According to the director, the subject librarians are trusted by faculty and departments, an advantage not shared by other support units on campus (ID 58). Since it is unclear if DSpace will work for faculty, many librarians are not willing to jeopardize the trust librarians have developed with the departments by pushing the repository (ID 58). To the repository director, it seems that one third of the librarians are seeking content and holding training sessions, one third have found no interest from their departments, and one third are just ignoring the repository (ID 46).

Faculty on campus have expressed doubts about the repository as well. One faculty member, active in open access publishing, asked why the library was doing this and did not think the repository was relevant to him or the library (ID 60). Another was concerned that people would put up “rants” in controversial topics, like evolution and intelligent design (ID 61).

#### **4.1.7 Project Outcomes**

The content recruitment went enormously better than I anticipated, but you have to understand that I have pretty low expectations. (ID 80)

The repository director believes that not limiting to peer-reviewed journal articles was the right step and that actively pursuing dissertations was the correct choice (ID 81). While the repository has been “meeting some people’s needs,” however, nothing can be generalized from the experience thus far (ID 80).

## ***4.2 Purdue University***

We're talking to people about data and we're doing projects related to data and in terms of all this other stuff we're no further than you guys are – trying to find resources, trying to do the assessment, trying to go back and actually build the thing to be successful. (ID 374)

### **4.2.1 Background**

Founded in 1869, Purdue University has a system-wide enrollment of 69,098 students at five campuses and numerous teaching and research sites. Located in West Lafayette, the main campus has an enrollment of 38,712 students served by twelve library “service locations” across campus (Purdue University Facts Online, 2007).

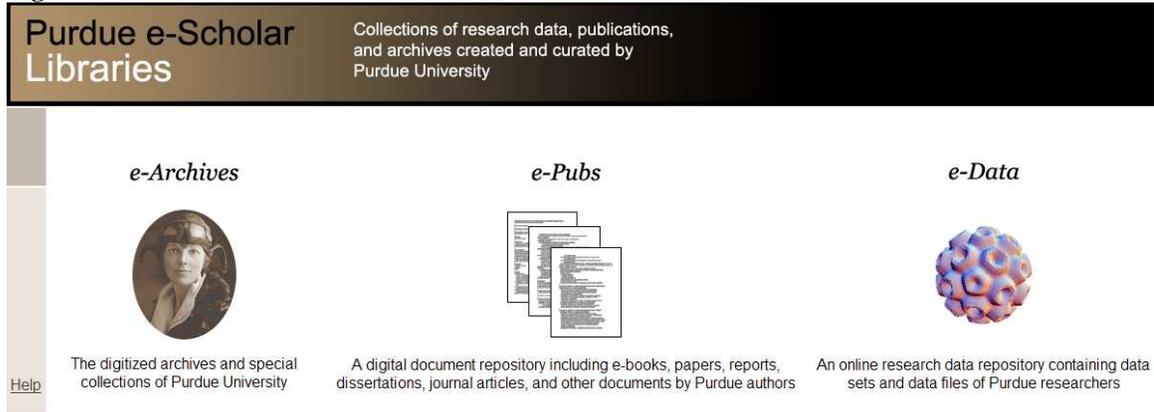
### **4.2.2 Repository Development**

We started talking to people and we started talking about data, and we asked what if we had a data repository... if we are talking about data, it might not be in just one archive, it might be in different places, because we are actually looking at the data grid, an informational TeraGrid, and the possibility of accessing files there, and the possibility of accessing files in more of an institutional type of repository. But the crackpot idea was: what if it's not one repository, what if it is different instances and different things, and that became the distributed institutional repository initiative. (ID 356)

Purdue University Libraries initiated repository development as a research project to create a production system to imbed digital objects into the library system (ID 355). Under the auspices for the Associate Dean for Research, a team was developed to look at “what we were going to do,” and a 2/3 time coordinator, who “becomes this go-to person for actually setting things up and creating the metadata,” was hired (ID 350). The project is a collaborative effort between the IT department, collections, and technical services with subject liaisons “finding out what is there, what the people have, what they need, what they do with it” (ID 350, 355). The project is changing the way librarians think about their roles on campus; they are doing “something that is a slightly different thing than we have done before, but not that different” (ID 355).

The Purdue Libraries developed a “distributed institutional repository” (ID 331). The repository, branded e-Scholar, contains three types of collections: e-Archives, running CONTENTdm; e-Pubs, a Digital Commons collection; and e-Data, an environment to access and archive datasets (Figure 7; ID 331). Under the current model, if content does not conform to the three existing collections, the repository team “would try and figure out where it would fit, and create the overarching interface to allow people to find things, hopefully, seamlessly” (ID331). This distributed approach allows the Libraries to “add the pieces that we need and repurpose the visualizing within the larger [umbrella]” (ID 341). Not only does this approach provide for easy expansion, it allows the library to choose software that is appropriate for the content. For example, the e-Archives collection is in a CONTENTdm environment which is more flexible for archival items and allows the library to “make the best use of the functionality that’s there” (ID 331).

**Figure 7. Purdue Libraries e-Scholar Portal**



By conceptualizing the project as an umbrella for campus content, the library is able to incorporate archives, special collections, digital collections, digital libraries and other digital media initiatives into a repository project, rather deploying resources as they become available (ID 332). This “project perspective” has created a different approach to digital collections focusing resources on data archive projects rather than digital library collections (ID 332).

The repository began as a live pilot project using Digital Commons. The University Press had used Digital Commons to manage “a couple of online journals” and the “press director has found that if you make e-books available, people still want to buy them, and they found that they sell more print copies when they actually make it available freely” (ID 329). There has been discussion about collaboration in the future, but the repository is not mining the university press for content (ID 329).

The repository group has made an effort to discuss the repository with people on campus “whose stuff goes in there,” identifying what is useful to faculty and researchers, not librarians (ID 376). The group has to be careful when it identifies campus needs, it can create “trouble” when the wants and desires do not match with functionality of the products available (ID 376).

While the repository project has been announced, there is not a “strategy for recruitment” (ID 338). Instead, the developers are looking for willing partners on campus and finding “ways to help people solve their problems” (ID 338). There are many organizations and research groups on campus that “have no real, efficient system for putting up on the web” (ID 338). The developers are identifying these groups and “pointing out how [the repository] helps them” (ID 338).

So, our first level approach, having looked at the articles and seen the “if you build it they will come” approach doesn’t work. We set out to identify a couple of communities that would be, what’s a good word, somebody who would really work with us and quickly put something up there so we could demonstrate the capabilities and the power in the system. (ID 338)

Digital Commons only allows submissions from Purdue email addresses. There are three levels of access: all viewable, Purdue viewable and embargoed (ID 342). There has been some interest in developing collections that include researchers at other institutions working on projects with Purdue scholars (ID 342).

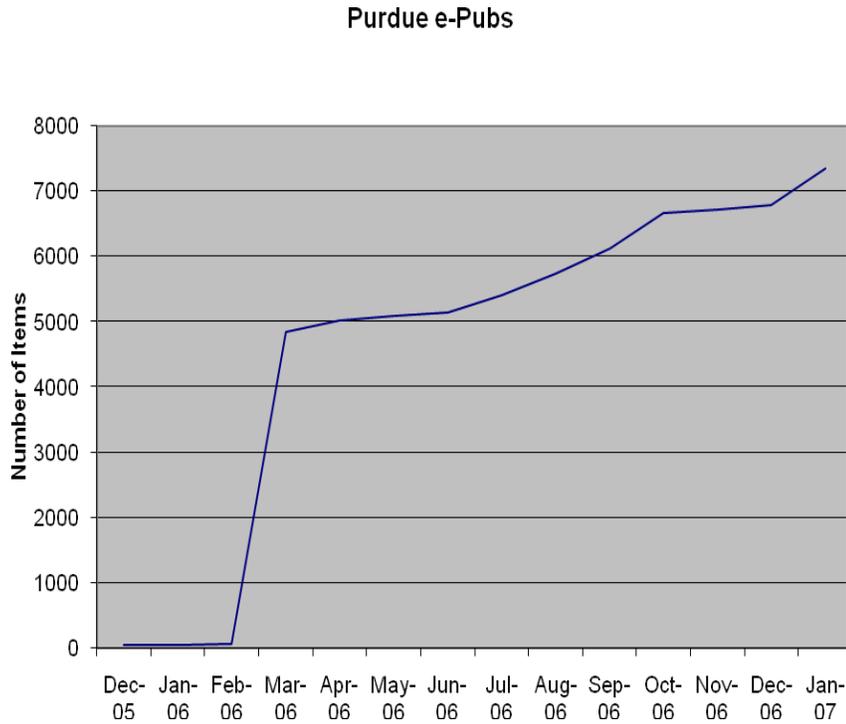
### 4.2.3 Content

I liken it to the movie Jurassic Park where the guy looks down and suddenly sees these ripples in this puddle. What's causing those ripples? I think it's suddenly going to open up when everybody realizes the power of this. (ID 347)

A ProQuest product, Digital Commons provides access to a university's theses and dissertations in Digital Dissertations. Over 4,000 records were added to e-Pubs from Digital Dissertations (ID 360). Dissertations are available full-text to campus users and Committee on Institutional Cooperation (CIC) members. Users without a valid username from Purdue or a CIC university are only able to access dissertation citation information and abstracts (ID 347, 449). The ability to quickly add dissertations to the collection provided developers with enough content to cause a "ripple" (ID 347).

In addition to research documents, Purdue repository developers are trying to create tools to archive datasets. For a project involving soil surveys, for example, a librarian is working to incorporate scanned maps with resources, but the project involves "build[ing] some interactivity between the text and the maps" which may not be possible within ProQuest (ID 340).

**Figure 8. Purdue e-Pubs Growth from Inception to Case Study Visit.**



#### **4.2.3.1 Planning**

An implementation team at Purdue was established and the project started with the idea that “we’ll figure things out on the ground” (ID 361). Digital Commons “was a cheap, almost a turnkey kind of solution” (ID 360). The dissertations were accessible to the campus community and the library could “repurpose the metadata from them to do other things” (ID 360). The team “put up something as a test, and it is open for the world to see” (ID 361). The site has been live ever since.

#### **4.2.3.2 Policies**

The development team was still identifying problems and issues related to policy development, though they already had a case where something was withdrawn from the collection (ID 367). The repository is using a community model to organize content around research groups and departments (ID 335). While student output is not being captured other than theses and dissertations, a “department or a center is kind of responsible for at least articulating what the community is about and what goes into it and, maybe, how it goes into it” (ID 335).

Discovery Park, a group of research centers on campus, has discussed mandatory participation. It is unclear how this can be enforced, one option would be to withhold funds, but it may be of interest to other institutions or groups trying to mandate participation (ID 368).

#### **4.2.3.3 Recruitment**

Initial recruitment began with the “current intellectual output of the university, research output, papers, reports, pre-prints, post-prints and the like of faculty” (ID 334). Much of the content is technical reports in engineering; “it’s fairly grey literature, so it is fairly useful to make it available and searchable” (ID 335). The difficulties lie in finding the “administrative structure” in the departments responsible for managing technical reports and reprint collections (ID 337). The Physics department used to have somebody who managed their reprint collections, and, since that person retired, “they are trying to figure out how to replace that function” (ID 337).

#### **4.2.3.4 Marketing**

Like many libraries, Purdue is unsure how to market the repository. Developers think “going out to specific communities and working with them is helping [to] determine the best ways to market this, to say this is the value added that we can do to this, this is the efficiencies and the maintenance” (ID 369). They are also working with “groups to get testimonials and supporters who will do some of the talking for [developers] (ID 369).

#### **4.2.3.5 Preservation**

Do you know anyone who has an actual Atari that is up and running, so that if you had an Atari disk you could go there? It turns out there are emulators. (ID 372)

The implementation team is more concerned with the data itself, with “holding on to this thing” (ID 372). The focus is preserving an intact file that may be read with another piece of software rather than identifying formats that will be moved to newer versions of a software package.

#### **4.2.4 Assessment**

We are going to send somebody away to assessment training to figure out how to assess this. That’s my hope. (ID 363)

Development started with a “hypothesis that if we were going to work with people in this sort of interdisciplinary research way, it would have something to do with data, data organization, data management, archive, preservation, that kind of thing” (ID

359). The repository can “help people solve problems” and that, rather than populating the repository, is the primary goal (ID 359).

#### **4.2.5 Benefits**

It’s sort of like when you talk to people about their data needs and that kind of thing, they’ll also then see it as part of the bigger picture of what libraries do for all kinds of things, whether it is documents or data or whatever, and then they see the libraries as being able to do more with regards to formats. (ID 373)

Developers were “surprised” by the number of downloads, but the repository has added “another dimension to how people can find something” (ID 364). By collecting “the intellectual output of the institution,” the repository is “providing access to stuff” (ID 373). Through the repository, librarians are able to “value-add and provide access to everything [to] show what [Purdue does]” (ID 373). The repository also allows the library to help campus communities solve problems associated with datasets, to work with people in “new ways” (ID 352, 373).

#### **4.2.6 Challenges**

My concern is that the umbrella can’t account for all the crazy machinations that are going on and perhaps obviously, it’s because of geo-spatial data. Dealing with crazy data, you have to have crazy things happen to make it usable. (ID 351)

The primary concerns of the group are “the accessibility, the sustainability, and the extensibility” of the repository” (ID 351). It is unclear how people are going to use

the collections or how the metadata within it can be repurposed to create “tracking mechanisms for the kinds of outputs people want for annual reports or some of these, they have a responsibility to their partner, to the university, to their granting agency” (ID 351). Value-adding will allow others to use the data and resources collected in the repository in different ways to solve community problems.

There are still obstacles for the repository. Without people dedicated to Digital Commons, it is difficult to identify all of the features (ID 362). A coordinator was hired to spend 2/3 of the positions time on the repository and is “just figuring it out as she goes” (ID 362).

#### **4.2.7 Project Outcomes**

The Office of the Vice President has now seen the proof of concept that we had – the distributed repository – and they are interested in seeing how things go. I think that the next step will be to talk about what we refer to institutionalizing the institutional repository. (ID 358)

The Purdue Libraries started an initiative to “look at some of the research administration with the research administered at the library to come up with some guidelines and templates so people can have those kinds of tools” (ID 354). Repurposing the metadata associated with repository items has sparked some interest. The ability to link faculty members and dissertations has been identified as a “great recruitment and promotion tool. This shows what Purdue does, so we’ll be able to show which department within which school and which professor has done that” (ID 358).

### ***4.3 Indiana University Purdue University at Indianapolis***

Our institutional repository is basically just that, it is defined as works that are contributed by groups or people that are affiliated with the university or campus in this instance, in some way. So, that has worked to be both for electronic theses [and for]...faculty sponsored conference proceedings. There has to be some kind of a relationship with the campus. (ID 4)

#### **4.3.1 Background**

Indiana University Purdue University at Indianapolis (IUPUI) was created in 1969 as a joint campus for Indiana and Purdue Universities in Indianapolis, and is managed by Indiana University. The campus currently serves over 29,000 students and has a satellite campus with a library in Columbus, Indiana. The main campus location in the state capitol has allowed the library to develop unique print and electronic collections in conjunction with state-wide organizations (ID 119). In order to manage those collections as well as the scholarly collections created on campus, the university libraries are using DSpace and CONTENTdm, a digital collection management package, to provide access to various electronic collections.

### 4.3.2 Repository Development

Discussions about an institutional repository began in 2002 with the support of the Dean of the Libraries. The Dean supports digital collection development and is trying to ensure the library has a central place in the campus' scholarly environment (ID 130). In addition to this support, the library staff was motivated by the Open Access Movement, and wanted to increase awareness of IUPUI and to preserve electronic documents (ID 139, 140).

The organizational culture at IUPUI supports the rapid adoption of new technology and its use. The librarians see themselves as Early Adopters of DSpace, which was installed in the library approximately a month after it was released by MIT (ID 131,132). Once the software was operational, they started adding content; the "pilot phase was installing it and seeing if we could submit something to it" (ID 41).

Currently, there are two instances of DSpace running on the IUPUI campus: IDeA<sup>15</sup>, the IUPUI Digital Archive developed by the library, and eArchives<sup>16</sup> developed by Archives/Special Collections. The IDeA project collects works contributed by groups and people affiliated with the university in some way and works as an institutional repository. This allows for the inclusion of electronic theses as well as faculty-sponsored conference proceedings. eArchives contains the digital materials collected by the IUPUI University Library's Special Collections and Archives, electronic versions of university documents and records (ID 31). The University Archives is using DSpace to capture electronic versions of university documents and records and as a digital library for special

---

<sup>15</sup> <https://idea.iupui.edu/dspace/>

<sup>16</sup> <https://archives.iupui.edu/dspace/handle/2450/1>

collections (ID 31). Collections within the eArchives include: General Collections, IUPUI University eArchives, and Philanthropy eArchives (ID 10). Although the sites were developed separately, the two are complementary, and staff members plan to work together in the future to market the services to departmental organizations and faculty.

The IDeA Working Group, contains representatives from Digital Libraries, Special Collections and Archives, Bibliographic and Metadata Services, the Medical Library, IUPUI-Columbus (Working Group for IDeA, 2007). eArchives, on the other hand, is the primary duty of one person. Content development for IDeA is done in addition to other library duties (ID 125).

### **4.3.3 Content**

The IDeA Working Group debated what types of materials were going to be collected, primarily whether to collect scholarly works or works in general. In the end, content in both collections needs to be connected to the university or related to the university in some way (ID 31). Groups created within DSpace control their own content; however, the library has control over group creation (ID 107, 111). The Libraries at IUPUI have developed relationships with many state agencies and organizations, and IDeA includes content from the State Library and other statewide organizations (ID 119).

The projects have similar problems with DSpace. Both collections include PDFs and HTML documents, but have not included multimedia content. While both plan to include the content, there are additional issues with media formats that have not been resolved. Some problems have occurred in the eArchives, older PDFs are not always full-text searchable and some of the university documents collected have security passwords

(ID 123). The DSpace interface also has problems which actually make it easier to search for materials with Google. The eArchives is working with a programmer to develop a better interface for one of its collections (ID 117). The interface problem also makes it difficult to access content from certain items. *Umbrella*, an image-laden modern art periodical, was added to both IDeA and ContentDM (ID 120). A survey has been developed to identify which interface the users prefer; the survey results will inform decisions about future graphical additions. The librarians believe that graphically based items will be more likely added to CONTENTdm (ID 452).

**Figure 9. Screenshot with links to Umbrella in CONTENTdm and IDeA**



#### 4.3.3.1 Planning

DSpace was initially chosen for the project because of its flexibility (ID 145). Once DSpace was operational, the library community on campus was used as a test group to “develop some different processes and workflows” (ID 145). The IDeA Working Group did not spend a large amount of time on planning, but has developed policies for Content Guidelines, Community Guidelines, and Format Support Policies (ID 144).

#### **4.3.3.2 Policies**

IDeA policies allow the faculty members in a group to define a collection and decide what it considers scholarly material; promoting the concept “you get to control what you want and where” (ID 109). The IDeA Working Group has been more involved with submissions than anticipated. Its goal is to see authors submitting their own works and applying their own metadata. At this point, the Working Group is adding files and some metadata to the collections. The Working Group is taking a more active role in identifying collections, contacting the necessary people, and adding the content.

#### **4.3.3.3 Recruitment**

[It is] not us deciding what we want to take, but us taking what people will give.  
(ID 175)

IDeA began with the publications of the library staff and expanded to include a bioethics collection, *Indiana Authors*, and the *Conference Proceedings of Midwest Research to Practice* (ID 166). eArchives is able to mine the IUPUI content as part of its mandate, prioritizing the sites on campus with content that needs to be collected (ID 174). The eArchives and IDeA developed as separate projects; however, the repository developers now plan to provide information about “both options” when meeting with faculty, increasing the visibility of both projects (ID 153).

In addition to the faculty groups and organizations, the IDeA group is working with the Graduate Office to develop an Electronic Theses and Dissertations (ETD) collection. The process took two years to identify appropriate graduate offices and

develop work flows, since some IUPUI graduate from the Bloomington campus and others from the Indianapolis campus (ID 148). The Graduate Office was supportive of the project the entire time; their only concern was the reliability of the system and its availability (ID 167). The process is not mandatory for graduate students yet (ID 157).

#### **4.3.3.3 Content Examples**

- FOLIO (FOundation LIterature Online) is a joint project between the IUPUI University Library and the Foundation Center that captures the publications and reports produced or funded by foundations and makes them permanently accessible. The collection is part of the eArchives.
- National Council on Public History (NCPH), part of the eArchive General Collections contains a scanned volume of *Public History News*.
- *Indiana Libraries*, a professional journal for librarians and media specialists published by the Indiana Library Federation. An IUPUI person, while not always the editor, will be involved with the magazine. The issues are located in IDeA (ID 107).

#### **4.3.3.4 Marketing**

The ongoing education of faculty and librarians has been key to IDeA's marketing effort (ID 135). Members of the IDeA Working Group are meeting with faculty members, groups and departments across campus to demonstrate what is possible (ID 160). The IDeA website includes background information and policies, providing information "which is on all the time" (ID 161).

#### **4.3.3.5 Preservation**

We anticipate having to put some [preservation] procedures in place, but we really haven't had to do that yet. (ID 123)

While the DSpace servers are regularly backed up, neither repository has a long term preservation plan. The files can be archived, “but archiving a file and making it function are two different things” (ID 115).

#### **4.3.4 Assessment**

I guess we'll know if [the repository] is *not* successful, rather than successful.

We'll know if it is not, well obviously, if three years later, we disband the whole thing and give everyone back their submissions. (ID 13)

The IDeA project currently has no assessment measures in place, however, the library advertised for an Assessment/Imagineering Librarian, responsible for developing an assessment program for the library (ID 150). Surveys have been conducted on smaller collections in the repository, like the ETDs, to examine the perceived ease of use (ID 150).

### **4.3.5 Benefits**

Developing the repository has given libraries an opportunity to become involved with “something that is new to the library world in every word” (ID 169). The librarians have been able to interact more with faculty and develop solutions to the changing scholarly communication landscape (ID 169).

### **4.3.6 Challenges**

Several challenges were perceived by the repository developers, among them: educating the faculty, technical support, and copyright. Educating faculty and librarians is a large part of the project. The team meets formally and informally with faculty across campus. There is also an ongoing effort to educate the librarians, especially the subject specialists, about the repository and open access issues in general (ID 109). The subject librarians do not seem confident with the repository, and the team has scheduled talks to demonstrate collections and options so “they can talk to their faculty a little bit more intelligently” (ID 109). While the IDeA Working Group was able to gain some support from the librarians, they have not received the technical support needed. Participants from both IDeA and eArchives identified the lack of commitment from technical support as being a larger challenge than anticipated (ID 168). Archiving student projects, especially multimedia projects, concerned the group. The library staff “paused” at copyright issues until they spoke with Kenneth Crews at the Copyright Management Center on campus (ID149). Finally, the DSpace software itself creates challenges; “while it offers some really great benefits, it also is kind of clunky, [and] ...has a barrier to use” (ID 110).

### **4.3.7 Project Outcomes**

If I were to look at what we have now and go back to when I was initially involved, I would say we haven't accomplished very much at all. But, having gone through the two years and seeing the struggles, not only of us, but everyone who is doing this, I feel like we are okay. We are not where I had hoped [to be] at the start of this process. (ID 12)

On the whole, there were some positive outcomes from the project. The librarians involved have been active in the Open Access and Open Repository movements, sharing “early experiences with others” (ID 147). The library developed an Electronic Theses and Dissertation (ETD) collection (ID 147).

## ***4.4 University of Washington***

I think it's just a repository, a digital repository. And I think we have shied away from the institutional repository. (ID 28)

### **4.4.1 Background**

Founded in 1861, the University of Washington has a system-wide enrollment of 44,023 students at three campuses (University of Washington, 2007). Located in Seattle, the main campus has an enrollment of 39,646 students served by more than twenty libraries and collections on campus.

### **4.4.2 Repository Development**

The Digital Initiatives Program has been developing an online repository for over four years (ID 393). Requested by library administration, a repository was developed to support electronic publishing at the University of Washington (ID 29, 379, 381). With no campus-wide electronic records management system, the repository may also be able to meet the requirements of archives as a preservation tool (ID 402). The D-Space repository was developed by four or five people within the library, and in 2005, an instance of D-Space was working with the campus authentication system (ID 379, 400).

The D-Space environment is just one of several projects on campus that “has repository connections to it” (ID 400). Other campus projects, like the Digital Well<sup>17</sup>, have been built to “acquire, collect, classify, store and deliver large collections of digital media over IP-based networks” (DigitalWell - Executive Summary, 2007). However,

---

<sup>17</sup> <http://digitalwell.org/>

policies and procedures have not been developed for the project, and there is no support for the Well. The university also has developed collections in ContentDM, which are linked in the online catalog (ID 404). Conceptually, the D-Space repository can be used as a “fulfillment tool,” providing access to documents found through Google (ID 404).

While the Digital Initiatives Program is responsible for outreach and marketing of the instance of DSpace, Information and Technology Services (ITS) is responsible for batch uploading, importing and technology issues (ID 378). Faculty or librarians interested in developing a collection meet with the head of Digital Initiatives to “understand [the project] in context” (ID 378). Other people, like programmers, are added to the project as needed (ID 397).

While the project was done “in parallel with other scholarly communication efforts” on campus, it is unclear if the faculty has a “good sense about open access” (ID 382). The repository is “largely invisible,” not a “campus issue at this point” (ID 402). The repository is working to “support the ongoing business processes of the library” which can create ties to the larger campus community (ID 402). For instance, there is campus interest in a common file sharing space or an authenticated collaborative workspace (ID 381).

### 4.4.3 Content

In addition to electronic publishing activities, University Archives was included in the project “because of need;” the archivists needed to store digital only collections from departments (ID 383). The repository is not only for post-print and e-print articles, but also includes the digital learning objects and student produced research papers (ID 403).

A repository can be an open access project or preservation project depending on what the campus is trying to accomplish (ID 396). Conceptually, anything that needs to be preserved long term can be placed in the repository (ID 403). The different dimensions of a repository collection make it difficult to send a consistent message to the campus (ID 396). In this case, developers did not think that a reprints repository made sense at the university level, since it is not the first place that researchers would go for the information (ID 394). The DSpace developers were not sure if a reprints repository has a value on campus. If researchers identify first with their disciplines, not their institutions, then repositories at an institution may have to redefine what the repository is used for (ID 394). Not focusing on reprints has allowed the developmental group to try some unique pilot projects.

Initial interest in the project by the Applied Math faculty failed when the department “wasn’t willing to give us the materials” (ID 384). Applied Math had already created its own metadata and was looking for a portal to the collection (ID 384). Individual faculty members in the sciences were interested and started individual and departmental collections (ID 384). DSpace is also used to manage internal collections. A pilot collection of “admin tech reports” has been added and DSpace is also being used as a collaborative workspace for a Buddhist manuscript project (ID 403, 380).

#### **4.4.3.1 Content Examples**

The “first real collection” included in the University of Washington’s D-Space was the Early Buddhist Manuscripts Project (ID 380). The project needed a place to store manuscript TIFs and collaborate on associated metadata (Figure 10; ID 380). The project is not a live collection, but a workspace until a corresponding book is published (ID 380; Early Buddhist Manuscripts Project, 2005).

#### **4.4.3.2 Planning**

The library did not create any planning documents prior to development. More planning is needed, but it is unclear whether D-Space should be presented as a tool or a service option (ID 408, 396). Initially, the repository was planned to house pre-prints (ID 416). The development group found this plan to be unsustainable, and is identifying additional areas of collection development. They are considering new directions and are starting conversations with users on campus (ID 416, 453).

#### **4.4.3.3 Policies**

Policies are negotiated with individuals and groups as collections are developed, with the library trying to retain access to all of the materials (ID 389). For example, the Buddhist Manuscripts are currently open to three or four members of the research group, but will become publicly available once the print edition is published (ID 389, 390). The library is also working on policies for undergraduate theses. While the University of Washington does not have an Electronic Thesis and Dissertation collection, they are using D-Space to store and preserve materials that accompany theses (ID 409).

#### **4.4.3.4 Recruitment**

Aside from campus publications, UW needs a central university system for coping with electronic institutional records on a decentralized campus (ID 418). The repository is providing the library with a location to store these materials and attract additional materials to the repository (ID 418).

Currently, the DSpace developers on campus will take any format and use the opportunity to “figure out how to put it up” (ID 387).

#### **4.4.3.5 Marketing**

The head of Digital Initiatives has been working with liaison librarians. Similar to other institutions, some liaisons are receptive, some are interested in what will happen and others aren't interested at this time (ID 401). Those involved with different projects are also giving presentations and lectures about how D-Space is being used to manage digital collections on campus (ID 401).

#### **4.4.3.6 Preservation**

A preservation plan had not been developed at the time of the interview, but a preservation consultant visited in early 2007 (ID 391).

#### **4.4.4 Assessment**

While no assessment measures have been developed, if the repository meets the institutions business needs, then it can be considered a success (ID 30). At this point, the development group does not think that there is “enough sustainable material to understand what [the data collected] means” (ID 407).

#### **4.4.5 Benefits**

The repository group identified several benefits that have developed from the project. The repository is a place for the “hybrid and digital stuff that needs to be preserved” (ID 412). It is providing the librarians with “a way to talk about issues that are difficult to talk about in an abstract sense,” like open access and preservation (ID 412). It has value as a “physical place,” a central place to archive and preserve digital content (ID 412). Finally, the repository can provide the university with a product that identifies the scholarly contribution that the school makes (ID 395).

#### **4.4.6 Challenges**

The repository does not have a “present clear message” to share with faculty and staff (ID 396). This has created ambiguity around the project, which makes it difficult for the liaison librarians to work with faculty (ID 396). Finding uses for DSpace is also a challenge; the repository has developed as a tool-based project rather than a service that meets a particular need on campus (ID 396). The library has not identified the problems that they are “going to address” with the product (ID 396). The challenge for the repository group is to “push back on why we are doing this, [and] spend more thinking about it;” letting the conversations between library liaisons, faculty and staff identify the “programmatic needs” of the campus (ID 399).

In addition to some of the campus difficulties, discussions with faculty have identified areas of concern with content. In one case, a humanities faculty member was concerned about excellence and the development of communities run by committees (ID 386). During a preliminary conversation with some faculty in the Department of Communication, “they were appalled that there was even a mechanism for review” and thought it was “morally objectionable” that such a process would exist (ID 386). Several factors come into play in this situation.

Negotiating policies with different groups may offer them more services than needed (the development of committees and a review process). These policies do not have to be adopted by individual groups, and stressing these features as part of the tool, rather than campus policy, may be beneficial. Faculty may be unfamiliar with D-Space and open access repositories, especially with the community based software that can

incorporate services developed for specific institutions or collections. The difficulties with the open access concept are further illustrated by a conversation the project developers had about a collection of audio files. When informed that the repository would be harvested by Google, the faculty member wanted to give every file the same metadata because the unique I.D. works with the campus finding aids (ID 405). The faculty member did not understand that how an item looked “in the new environment” was as important as the content the collection contained (ID 405).

#### **4.4.7 Project Outcomes**

One of the things we should talk about is our campus publication, departmental and college newsletters, campus papers, and things that don't have a formalized publication channel and otherwise would disappear in filing cabinets or e-space. Being sort of an adjunct for the archival half of it, we are able to collect these e-publications that are unique to the campus and not available anywhere else. (ID 398)

In addition to the e-publications on campus, the repository has been successful digitizing the library's collection of materials that accompany theses and dissertations (videos, datasets, maps or images) (ID 414, 417). Prior to the project, the materials were kept in a locked file cabinet and were not easy to access (ID 417). Although the collection is not available to the public, if people request the materials, the library “will burn a CD for them” (ID 417).

## ***4.5 University of British Columbia***

When you start to look at [collections], indirectly libraries handle published articles and that sort of thing; archivists similarly handle personal papers. There is a whole lot of stuff that falls into the middle somewhere, and...it's in this grey lit stuff that can be represented [in the repository], so ultimately it's the student output, it's faculty output, it's institutional output – all those sorts of things altogether and anything, any areas where those areas touch. (ID 17)

### **4.5.1 Background**

Established in 1908, the University of British Columbia is one of the top forty public universities in the world, with a system-wide enrollment of over 45,000 students at four campuses (University of British Columbia, 2006). 43,579 students were enrolled at the main campus located in Vancouver. The Library is the second-largest research library in Canada with over twenty branches/divisions (University of British Columbia, 2006).

### **4.5.2 Repository Development**

The UBC Library began looking at institutional repositories around 2004 or 2005 when the eLibrary Committee developed two working groups on campus (ID 204). In order to get as many people as possible involved in the discussion, one group looked at institutional repositories and a second worked on scholarly communication (ID 204). Although, considered by some to be an “artificial distinction,” the eLibrary Committee

thought the separation was necessary (ID 204). The scholarly communication group was exploring e-journal hosting on campus and open access issues, while the repository group developed a recommendation to set up an institutional repository (ID 204).

The Institutional Repository Working Group (IRWG) included a cross section of librarians from the campus, with the conceptual goal of “exposing [UBC] research, in particular [UBC] research papers, to a broader community” (ID 97, 84). The University Archivist was chair of the eLibrary committee which allowed him to “inject kind of an archival perspective on things and to begin to look at stuff that hasn’t been caught either in libraries or archives,” like data sets, the publications of faculty members, university publications “the whole breadth and scope of the output of an institution” (ID 182).

In the fall of 2005, the IRWG and eLibrary Committee submitted a position paper to the library administration recommending the creation of an institutional repository at UBC. Around the same time period, an Interim Deputy University Librarian was appointed to manage the Library’s day-to-day operations, and the University Librarian supervised the completion of the Irving K. Barber Learning Centre (Report of the University Librarian to the Senate 2005-2006, 2006). Under new management, the position paper was re-worked as a project charter by the AUL Science Libraries and the University Archivist, and presented to the library in the Spring of 2006. A part-time project coordinator was finally hired in February 2007.

Prior to 2005, the University Librarian did bring someone from MIT to speak about DSpace with the President’s Advisory Council on the University Library (PACUL) and the library staff (ID 97). The repository project was also presented to the Faculty Alliance for Technology in Education (FATE) and received a positive response (ID 97).

Response on campus was generally positive, and faculty saw the value of the project:

We had an interesting conversation with a fellow who is the just past acting dean for science – his area of research is climate change. So, when we talked to him about this he said “oh, this will be fantastic because I have had data on Churchill [Manitoba] and it would be significant to have that data now and compare it against what we have now. It could really show some of that climate change data.” So he immediately recognized the value of access to previous data and being able to use it. And I think, we paid for that data as taxpayers, and now we lost it. (ID 84)

Theoretically, the repository would collaborate with whatever digital initiatives exist on campus and incorporate a federated search for all library collections (ID 177, 212). DSpace has been adopted, but the library is open to other options and believes that “it is really important that you wind up using the tool that is best suited to what it is you want to organize and provide access to” (ID 177). The repository may incorporate “multiple tools to both do and create [a] repository, [a] virtual repository as opposed to a big central single storehouse of stuff, but yet you are able to put things in a box” (ID 177).

#### **4.5.3 Content**

The line between library and archives in a repository is increasingly grey. The need to have categorizations breaking things down – you see the mergers of a

bunch of institutions where the library is merging with archives or the archives is merging at a national level, and so we are all...is there a real value in maintaining the distinction? There is tons of stuff that happens on the periphery, and you have to think about how best to manage it. Does it really matter per se if it is published in [the repository] or unpublished? (ID 203)

The project charter identifies several collection areas for the repository. In order to create a collection associated with the campus, the group is looking at ways to incorporate datasets as well as grey materials, archival content, conference proceedings and student theses (ID 83). There is also interest “in the work that is going to come out of the Olympics and that might be from the researchers here, but it could also connect to other research related to the 2010 Olympics in BC” (ID 83). The repository is seen as a place to collect “the productive research and knowledge creation that happens at an institution” (ID 83). The current funding is for a one-year pilot project, and the project coordinator finds it unrealistic to focus solely on peer-reviewed publications (ID 218).

#### **4.5.3.1 Planning**

The IRWG, in collaboration with the Science & Engineering division of the UBC Library, developed a survey on scholarly communication (ID 184). The survey asked questions about repositories and helped the working group “get some sense of who had made use of these things at the [subject level] and just in general who knew about open access” (ID 184). The survey provided the group with some preliminary feedback and identified faculty members interested in developing collections (ID 184).

The project charter outlined the role of the repository, a project plan, and goals and objectives. The detailed project plan identifies collections to develop for the pilot, identifies linkages with other local, provincial and national repository groups and services, and identifies key stakeholders. The IRWG forgot to include, at this stage, the School of Library Archival and Information Studies (SLAIS). Group members plan on including SLAIS in the future (ID 98).

**Figure 10. Project Deliverables from the UBC Institutional Repository (IR) Pilot Project Project Charter**  
**Project Deliverables**

Major Project Deliverables & Milestones
<ul style="list-style-type: none"> <li>○ Determine what metadata requirements we need i.e. elements and who assigns those elements</li> <li>○ Develop plan to add metadata sets and a range of subject areas, communities, and file formats to the showcase/demo including: <ul style="list-style-type: none"> <li>➢ 2010 Legacies Now (Olympics research)</li> <li>➢ UBC Bioinformatics Centre</li> <li>➢ Faculty of Arts Image Repositories</li> <li>➢ UBC Documents</li> <li>➢ World Urban Forum 2006</li> </ul> </li> <li>○ Develop policies related to responsibilities for depositing resource/department and the library <ul style="list-style-type: none"> <li>➢ Collection criteria</li> <li>➢ Licensing and copyright</li> </ul> </li> <li>○ Map the processes and document the procedures for creating and ingesting meta data <ul style="list-style-type: none"> <li>➢ define, test and implement processes to capture the material and add meta data</li> <li>➢ develop end user guide</li> </ul> </li> <li>○ Develop a draft marketing/communication plan for IR that addresses needs of Library staff (raise level of awareness) and the contributors and consumers of IR information/content</li> <li>○ Determine what type of information people need to participate i.e. incentives</li> <li>○ Define an exit strategy for the pilot project</li> <li>○ Identify resource implications and develop business case and scaling up plan for IR Phase II as required: <ul style="list-style-type: none"> <li>➢ People, dollars, IT re. hardware/server requirements etc.</li> </ul> </li> <li>○ Recommendations regarding organizational governance for IR post pilot: including <ul style="list-style-type: none"> <li>➢ Reporting structure</li> <li>➢ Steering Committee</li> <li>➢ Liaisons with academic community</li> <li>➢ Other TBD</li> </ul> </li> <li>○ Final Report</li> </ul>

Starr, L. and Hives, C. (2006). Institutional Repository (IR) Pilot Project Project Charter. Unpublished document.

#### 4.5.4 Assessment

Quantitative measures, after the pilot, may not provide enough information for assessment (ID 216). Instead critical success factors were identified in the project charter.

**Figure 11. Critical Success Factors.**

##### **Critical Success Factors**

- That the IR pilot project is aligned with the Library strategic plan/direction
- That we are able to achieve buy-in for the pilot project from the Library branches/ staff
- That we have enough participation/content to achieve critical mass re. support the IR pilot
- That we are able to clearly identify what the obstacles are to overcome for Phase II
- That we manage to engage interest from senior academic team at UBC i.e. President, VP Research and all other senior administration team members
- That we have IR champions from the faculty/academic side of the house
- That we have people willing to serve on the IR implementation team
- Utilize recognized IR and metadata standards
- That we have the processes for creating, ingesting and storing metadata defined, tested and implemented and these processes are straight forward and transparent
  - Capture the material
  - Add metadata
- IR pilot project is profiled through the Library Public web site

Starr, L. and Hives, C. (2006). Institutional Repository (IR) Pilot Project Project Charter. Unpublished document.

#### 4.5.5 Benefits

The repository should be there to support learning, research, teaching the business of this campus. I think that there are huge opportunities there, especially if we can recruit some really excellent content. (ID 210)

The repository systematically makes the intellectual output of the university more widely accessible than discipline specific sources (ID 103, 194). Repository development is providing the library with an “opportunity to have a leadership role outside the normal scope” of its activities, creating a proactive role for the library (ID 84, 202). This will

create a showcase for the work that happens at the university and “what makes this university compete with other universities... it will support particularly the research efforts and hopefully the teaching and learning as well” (ID 220). This new role can lay the groundwork for future faculty/librarian projects involving digitization and research (ID 94).

#### **4.5.6 Challenges**

I would have liked for my own institution to believe it might be useful. But, you know, that’s life in an institution, whenever you try to change the action a little bit, introduce new ideas, it takes awhile for it to work its way through the system. (ID 187)

Because of the administrative changes taking place within the library, the biggest challenge to the project has been getting the IR Coordinator position “off the ground” (ID 104). It took almost a year after the charter was accepted for the IR Coordinator position to be filled (ID 104). Personal lobbying was one of the things that the librarians could have done to push the project along at a faster pace (ID 93). Now that it is starting, the library has to make new connections within the pilot framework to move the project ahead (ID 93).

#### **4.5.7 Project Outcomes**

The working group participants gained experience with ETDs and D-Space during a concurrent project developing an environment for theses and dissertations (ID 200).

## ***4.6 Simon Fraser University***

An institutional repository is a set of services that a unit on campus or multiple units, typically the library, establishes to allow faculty and staff and students to aggregate documents of scholarly interest – that might be of interest both to the university’s user base and to people outside the university. So, basically, the services take the form of a website, but I don’t want to minimize, this might come out later, I don’t want to trivialize the effort that goes into an IR by calling it a website, it’s not just a website it’s all sorts of things, it’s policies, it’s preservation, it’s support – so, there is a website face, but there are all kinds of services going on. (ID 20)

### **4.6.1 Background**

Located in Burnaby, British Columbia, Simon Fraser University opened to students in 1965 as a second campus available to students in the Lower Mainland (SFU Calendar, 2007). Located less than fifteen miles from the University of British Columbia, SFU serves over 30,000 students at three locations in Burnaby, Vancouver and Surrey (SFU Calendar, 2007). Unlike UBC with multiple libraries and reading rooms at several campuses, SFU has a main library on the Burnaby campus and smaller collections at the Vancouver and Surrey sites.

The library is home to many provincial and national projects including the Public Knowledge Project, a joint research initiative with UBC, which “seeks to improve the scholarly and public quality of academic research through the development of innovative online environments” (Public Knowledge Project, 2007).

#### **4.6.2 Repository Development**

In November of 2002, DSpace was released, the first version, and not long after that we set up an instance and started playing with it to try and get a feel for how we would again present it and what kind of services we could wrap around it to present as an IR package to our university community. So, once we were feeling comfortable with all the stuff and developed documentation and so forth, we began our IR by populating it with content that we had already gotten from faculty and conferences and so forth. (ID 225)

Repository development began at SFU after a request from a faculty member at the Surrey campus to deposit research at the library (ID 279). The library had previously installed EPrints, but the software “wasn’t terribly fast and responsive or flexible” (ID 225). A copy of DSpace was downloaded and the institutional repository was started in 2003 (ID 279). The repository was developed under the direction of the University Librarian by two librarians from the Burnaby campus and one from the Surrey campus (ID 231). The project was located within and driven by the Library Systems department (ID 287).

The library's focus is to provide access to licensed and free content, and the repository provides access to the materials generated by the campus (ID 251). An early collection from the annual Digital Games Research Association (DiGRA) meeting held in Vancouver provided an opportunity to take advantage of some of the services available in DSpace (ID 232). It was important to the development group to use open access software. The Library and Archives of Canada aggregate all theses and contract with ProQuest for microfilming and digitization (ID 227).

The developers were able to take advantage of the digitization center on campus, and the IR has become a "container or a home for some of the stuff" (ID 251). No formal assessment measures were used prior to development, but there were several collections identified by developers (ID 281). Instead, developers chose to "move in a more agile way and get a core that we could show people" (ID 281). While formal assessment plans have not been identified, there are plans to develop a "more systematic approach to populating the IR" by identifying the faculty members publishing in open access journals (ID 281).

Developmentally, the repository group initially intended to use departmental assistants or secretaries as gatekeepers for groups (ID 293). Professors could begin the process by uploading the articles and adding a minimal amount of metadata. The departmental assistants or secretaries "could polish it off and commit it to the IR" (ID 293). This model has not been successful.

### **4.6.3 Content**

In addition to the DiGRA collection, theses were an early addition to the repository. The library has been the “gatekeeper for the theses,” and a library employee worked with grad studies to define “the protocol for submitting the theses (ID 227).

#### **4.6.3.1 Recruitment**

The DiGRA content was acquired through the work of an SFU graduate student on the conference organizing committee (ID 233). The open access and the IR were important to the student, and “he has actually been a good outlet for us in education and amongst other grad students” (ID 233). The student worked with repository developers to take advantage of the participants submitting their papers to the Open Conference System (OCS), which is an online submission and review system for conferences (ID 233). Conference presenters submitted their papers and all the associated metadata into a database, and Systems librarians wrote a script to export from that database and import it into DSpace (ID 233). The OCS code was modified to include a check box that said “I want my paper in the SFU IR and I grant this permission and I sign off on the copyright in question.” By checking that box before submitting, the paper went into SFU IR (ID 233).

Departments on campus are working with repository developers to create collections as well. Community Health Online Digital Archive Research Resource (CHODARR) is a collaborative effort between the Department of Health Science and the Library to create an online digital archive of research materials related to health and

social welfare, with an emphasis on housing, gender, aboriginal issues, HIV and mental health (ID 262).

Repository developers are considering the use of ROMEO to identify faculty who are publishing in open access journals (ID 247). Since the content is already open access, the developers could “have a project to take all that stuff and put it in the IR” (ID 247).

#### **4.6.3.2 Marketing**

Repository developers have produced a brochure, but would like a graphic artist and professional writer to revise the content (ID 295). It could be something the development team or liaison librarians take to department meetings (ID 295). It is unclear if the developers have “really figured out what it is we’re marketing” (ID 313). The developers have “come up with a new product here, and I don’t really think that we know exactly what it is or how we want to deal with it” (ID 313).

#### **4.6.3.3 Preservation**

No one in Canada has a preservation plan. (ID 241)

The Library and Archives Canada is developing a preservation plan for digital materials (ID 241). Once that is completed, the developers may follow its guidelines.

#### **4.6.4 Assessment**

The developers believe that the quantitative assessment measures are “relatively straightforward,” examining the number of hits or downloads (ID 322). They are still thinking about how to “build in some of the qualitative ones” and are considering something like an impact factor that provide a more qualitative assessment (ID 322). One option is “to look at the reasonably well defined metrics for the impact factors for, let’s say, the tier one published output there and see if they can be repurposed to some extent, and start making that sort of some application to IRs” (ID 322). Another option is to use the target groups, in particular faculty and researchers for focus group research (ID 233).

#### **4.6.5 Benefits**

Even though [the repository] shouldn’t be taken as a comprehensive bibliography of everything published at SFU, it can be useful to raise the profile of the university in certain ways. (ID 248)

Developing the repository provided those involved with experiences that can be applied to other projects. In addition to project management experience, the team learned about copyright, DSpace software, and access and preservation issues (ID 269).

The repository also provides the university with a central space to find digital materials (ID 297). Not only is it providing access, but it is ensuring access in the future, something that is not true of a document stored on a personal web page (ID 297). The repository “provides a relatively safe home for digital material” with the library committing resources to the long-term storage and preservation of the content (ID 248).

IR also provides visibility for the research of an individual faculty member or graduate student, as well as for the institution. For example, a master's thesis, published in the IR, was accessed by the Hamilton, Ontario Police Department. The police contacted the student for permission to print and distribute the theses to a group within the department (ID 248). The theses would not have been accessible to the interested group had it not been added to the IR; instead, it would have been sitting on a shelf in the library with a low chance that the document would be retrieved (ID 248).

#### **4.6.6 Challenges**

One thing we haven't overcome, that we need to work on is if we want more faculty to use our IR we need a promotional plan, a marketing plan. That is our current challenge. And because we are not putting a lot of money into it, we can't have one unless we allocate some resources to it. (ID 246)

Copyright is one of the biggest challenges facing the IR developers. With over 5,000 theses digitized, the librarians are waiting for permission from individual authors before the projects are moved to the repository (ID 253). Not only is the author's permission needed for submission, but in some cases, theses require the permission of the authors quoted within the theses (ID 238). Canadian copyright law does not allow for the "fair use" of materials for study and criticism and limits the amount of text that can be cited (ID 238). There are quantitative guidelines in the copyright law, and the librarian working with the theses often requires students to ask for permission to use the content in print format and, now, electronic format (ID 238).

The development group is “trying to get it populated, trying to get people to accept it and to agree to put their stuff in it” (ID 298). It is still unclear whether faculty will “give their research product to [the repository]” (ID 268). Resources for marketing and educating the faculty and librarians about the repository may be necessary to fully institutionalize the repository and ensure its growth. There are reservations on campus about open access and what it means:

You know I think one of the big surprises is the people who don't think open access is saintly; they have reservations about it and I understand the reservations. So it is funny to see us in a row saying: we're going to do this anyway. I mean people just...I think as librarians we just take open access and putting whatever that information is into an IR is a given, and lots of people don't. (ID 301)

One librarian on campus “thinks open access takes money away from vendors... her point is that nothing is ever free if people are not paying for access,” even though the money is coming out of the collection budget, her fear is that “publishers are going to stop publishing” (ID 235). Educating both librarians and faculty about open access publishing and its ramifications to the publishing industry are important aspects of collection development for SFU.

Resources are another challenge, while the librarians can “figure out how to bulk load more easily” and run the repository more efficiently, there is still a need for financial support from the university (ID 246, 268). Like many libraries, SFU is “not putting a lot of money into it,” additional resources are needed for marketing and securing copyright permissions (ID 246).

#### 4.6.7 Project Outcomes

The success actually has been on the user end, where people find these things in Google. It's like oh, wow - you're on the reference desk and it used to be before the theses was checked out – if we had actually had it in paper – or it's in microfilm in Ottawa waiting to be processed or whatever. Usually students' faces light up when it's in full-text. I think that has been the biggest success for us and quite a gratifying one. (ID 300)

SFU IR development started around the theses project, which one librarian thinks was “inspired” (ID 325). The content was already “under the control and responsibility of the library,” and the developers could “do what we needed to do” (ID 325). On the other hand, the developers avoided dealing with the university community, particularly faculty (ID 325).

Although the response from faculty has not been great, the repository is now home to over 1200 theses, as well as content from the British Columbia Electronic Library Network, Canadian Centre for Studies in Publishing and Multicultural Canada (Figure 13).

**Figure 12. SFU Institutional Repository Home Page (Screen Capture March 2, 2008).**

**Simon Fraser University Institutional Repository**

[Learn more](#) about how you and your colleagues can [benefit](#) by participating in the Simon Fraser University Institutional Repository Project!  
See our Institutional Repository [Policies and Guidelines](#) for further details.

**Search**

Enter some text in the box below to search DSpace.

**Communities in the SFU Institutional Repository**

Select a community to browse its collections.

- [ACTION for Health](#)
- [BC Electronic Library Network](#)
- [Canadian Centre for Studies in Publishing](#)
- [Community Health Online Digital Archive Research Resource \(CHODARR\)](#)
- [DiGRA 2005: Changing Views: Worlds in Play, 2005 International Conference](#)
- [Multicultural Canada](#)
- [Simon Fraser University Library](#)
- [Simon Fraser University Linguistics Graduate Student Association](#)
- [Simon Fraser University Theses](#)
- [SSHRC-Funded Projects](#)

## ***4.7 Subject Repositories***

Repositories at the subject and institutional levels share many of the same challenges. Interviews were conducted with faculty and staff involved with the IUBio Archive at Indiana University, the Centre for the Study of Historical Consciousness at the University of British Columbia, and DSpace at the University of Washington Health Sciences Libraries (HSLIC) to identify areas of similarities.

### **4.7.1 IUBio Archive**

Since 1989, the IUBio Archive collects *Drosophila* research data; molecular biology data; and software, as well as a mirror for GenBank (IUBio Archive for Biology, n.d.). Unfortunately, it has not “worked out as a self-serve archive of software,” the coordinator and his students are the only people populating the collection (ID 438). While institutional repositories may compete with traditional publishers or subject repositories, the biology community has begun to use SourceForge to distribute software (ID 439).

### **4.7.2 DSpace at the University of Washington Health Sciences Libraries (HSLIC)**

Registered in December of 2004, the DSpace repository at the University of Washington Health Sciences Library has more than 200 items. Unfortunately, the librarian initially responsible for the project is no longer with the library and no updates have been made since February 2007. The repository is currently maintained by the Systems department at the library, but, with staffing shortages, it is uncertain what will happen with the project (ID 440). Even though the repository was linked to the HSLIC home page and registered in ROAR, many people at the library did not know the

repository existed. Like faculty projects, DSpace projects run by a single person or smaller libraries are not necessarily sustainable. There may be an opportunity to combine collection with the one at the main library.

#### **4.7.3 Centre for the Study of Historical Consciousness (CSHC)**

Initially funded by the Canadian Foundation for Innovation, the director of the Centre for the Study of Historical Consciousness followed the advice of John Willinsky, who “suggested that an ‘open archive’ would fill the need for distribution (ID 441, 447). The repository provides “easy access and communication among scholars” and increases the profile of the CSHC (ID 442).

The director chose to use PKP software and handles technical difficulties, which produce most of the challenges for the repository (ID 443, 444). PKP software allows users to submit items through a web form and add the appropriate metadata (ID 445). The records are then reviewed and accepted by the director (ID 445). The Centre requests that conference presenters post their papers, but “there has been very little activity” (ID 446).

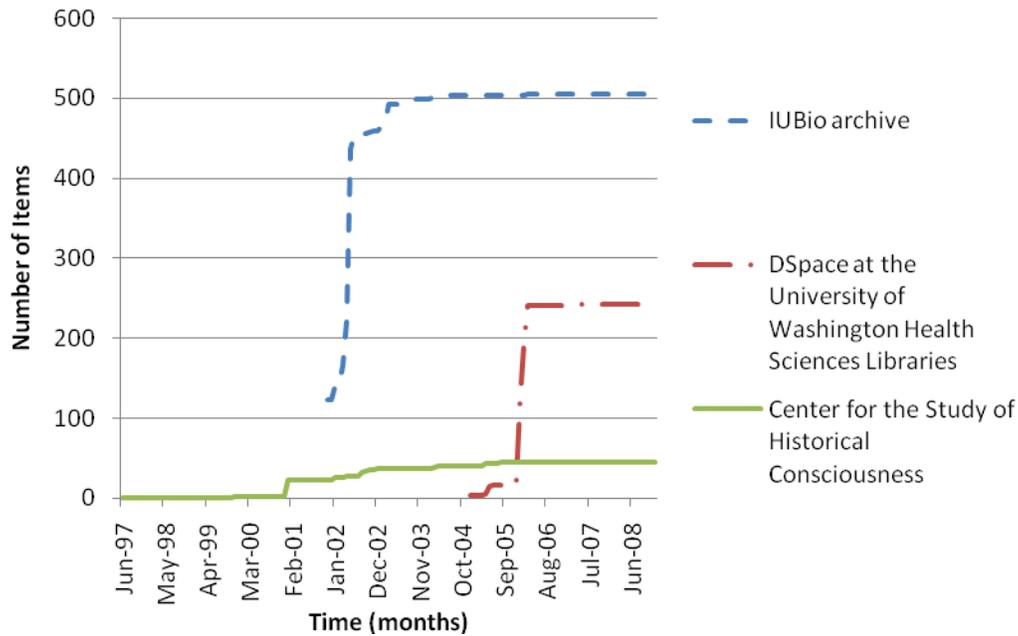
There is no interest in marketing the repository to faculty or motivating participation at this time (ID 446). The repository is sustainable at this level, but would require more resources to expand to “its full potential” (ID 448).

#### **4.7.4 Comparisons with Institutional Repositories**

The subject repositories face many of the same problems that are occurring at the institutional level. The repositories were chosen for the study because they were working within specific research areas or faculty groups. Even with close interpersonal ties to

researchers in the field, it is still difficult to motivate faculty to participate. After initial interest in submitting materials to the repositories, participation slows or stops (Figure 14). Usually run by a single person, the repository developers do not have the resources to market the repositories to interested researchers.

**Figure 13. Comparison of subject repository development**



Getting researchers to submit items is difficult, even with web forms. If developers have to load all the documents, it may take too much time.

Repository function has to be a defined part of someone's job, but the job cannot be done in isolation. Losing the sole person responsible for repository development can effectively kill a project, or, at the very least, make it very difficult for the next administrator to continue.

## **Chapter 5. Analysis of Cases**

### ***5.1 Examining Core Functions***

During the interviews, questions about the core functions of repositories and factors influencing the development of repositories were developed from the literature and initial case studies. Many of the functions and factors identified in the case studies are developmental issues that had not been considered by development teams at the time or were not important to development.

Many of the core functions discussed by Gibbons (2004) were identified as part of the software package and not of major concern to the development.

The research questions that guide this section of the study are:

RQ1. How are IRs developing?

RQ2. What are the factors influencing IR development?

RQ3. How should IRs develop?

#### **5.1.1 Materials Submission**

Materials submission is relatively standard across the cases. Units and faculty formed groups that were set up by the repository managers (ID 43, 113). The groups decided what content was appropriate for their collections (ID 49, 111). Anyone with a campus email address can submit, but some institutions are looking at allowing non-institutional research access to collections that re developed around a grant or research project (ID 432). Unless part of a group created by a department or faculty member, students may only submit dissertations and theses. However, some institutions were

considering the submission of fourth year undergraduate papers. The question then arose, if you take fourth year, then what about the third year and so on (ID 101). The sites took individual and batch submissions; batch loading was preferred, to some extent, especially if the metadata was consistent. Individual faculty members at most sites had the choice of adding records through the web interface or having someone involved in the project add them (ID 44). Materials from research centers and departments are typically batch loaded. In some instances, repository developers have agreed to scan and submit items for groups and departments, but it is not a practice the developers plan to provide for the long term (ID 63, 143, 251, 339).

### **5.1.2 Metadata Application**

Most of the sites used some sort of modified or qualified Dublin Core for the repository (ID 36, 88, 345). Both author and library supplied metadata is used (ID 44). Ideally, all the metadata should be author supplied, most participants believed that the more work required of faculty, the less likely they were to participate (ID 10, 11, 110, 293). The author supplied metadata was more prevalent in “quasi-library” collections, like the Archives of Traditional Music at Indiana University (ID 48). Most of the batch loading done by the libraries had some associated metadata (ID 388). Many sites were relying on the full-text search function rather than metadata to aid discovery (ID 434).

In many ETD collections, the digital copy was the copy of use. As such, call numbers are not a necessity and author supplied keywords or subject headings, in conjunction with full-text searching, provide points of access (ID 228). Metadata does have value in open access collections. Purdue University is using the metadata associated with its dissertations to create links to identify the dissertations from a particular department or chair (ID 358, 360). Sites are looking at repurposing metadata to develop reports from various research centers on campus or to add content to faculty activity reports (ID 37, 345, 351)

### **5.1.3 Access Control**

For the most part, the materials within the repositories are open access, however, there are some exceptions (Table 6). On several campuses, groups have been created that are closed to the public, with the content embargoed for a limited time period or forever. Purdue is considering the use of permanent embargo, the record would exist in the database but would not be publicly available. By placing an embargo on documents, the information on the item is still available or findable; there is still a record of publication that would not exist if the item was deleted (ID 367). Other institutions have embargoed dissertations and theses, either because they were going to be published elsewhere or they contain propriety information about a particular industry or company (ID 253, 276). Other institutions have used the DSpace environment for collaborative work projects, like the Buddhist manuscript project at the University of Washington. The manuscripts are closed to the public at this time, but once a book about the manuscripts is published, the collection will be open access (ID 380, 389). Software choice also limits access, the full-

text of the dissertations available in Digital Commons is only available to the host institution, or in the case of Purdue, to the host institution and consortia partners in the CIC, unless Digital Dissertations are available to the searcher (ID 426). The first twenty-four pages of the dissertations are freely available.

**Table 6. Comparison of Access Control**

	Indiana University	Indiana University Purdue University at Indianapolis	Purdue University	Simon Fraser University	University of British Columbia	University of Washington
<b>Limits on Access</b>	Open and Closed Collections	Open	Dissertation text available at Purdue and other CIC campuses	Open (Dissertations may be embargoed)	Still in pilot	Open and Closed Collections

#### 5.1.4 Discovery Support

Other than adding the repositories with different registries, like ROAR, most institutions rely on the repository software to make things openly accessible, allowing the content to be accessed by search engines (ID 432). The search function within DSpace does have problems; IUPUI is planning to develop its own interface for the eArchive to avoid some of the problems (ID 435). Other developers plan to incorporate some sort of federated search tool into the repository or larger university collection (ID 53, 177, 356, 436). The sites are interested in developing or finding a federate search tool that will allow users the ability to search all digital collections on campus (ID 433).

### **5.1.5 Distribution**

Distribution, as identified by Gibson (2004), is the “mechanism by which a copy of the digital file can be provided or displayed to the user” (p. 8). None of the sites involved considered this to be an issue, the software serves the files and it is up to the individual user to have the appropriate tools to access the information (ID 424). It was pointed out that while one did not need to have the actual software to run the files, one nevertheless needed some sort of emulator that would allow access (ID 372, 425). At this time, most of the files are PDF documents but, if collections start to add multimedia content or datasets, then access with more than MIME type will have to be addressed (ID 425).

### **5.1.6 Preservation**

Although developers acknowledge that a plan for long and short term preservation is an issue, none has developed a comprehensive preservation plan (ID 53, 90, 123, 348). The IUScholarWorks repository is covered to some extent by the preservation plan developed by the Digital Libraries Program, and once the repository is moved to Fedora, it will be part of its systematic plan (ID 53). Other sites have hired consultants, but a public plan has not been developed yet. Throughout the interviews, migration was identified as a possible solution, since librarians have been able to migrate data in the past (ID 427, 428, 429). It was also pointed out that “archiving a file and making it function are two different things” (ID 430). With many file types, the repository can archive the files, but that does not mean that the content is accessible.

### 5.1.7 Resources

According to Bailey (2006), institutional repositories cost at least \$100,000 per year. Even with reallocated staff, the interview participants believe that the estimated staffing expenses involved in the project alone would exceed \$100,000 (ID 91, 244, 333, 349, 370). From the interviews, it was clear that allocating a single person for at least 50% full-time equivalency (FTE) was optimal for development. A dedicated staff person was needed to manage not just the development, but content recruitment (ID 207, 243, 283).

**Table 7. Resource Comparison**

	<b>Indiana University</b>	<b>Indiana University Purdue University at Indianapolis</b>	<b>Purdue University</b>	<b>Simon Fraser University</b>	<b>University of British Columbia</b>	<b>University of Washington</b>
<b>Funding</b>	Reallocated Staff	Part of the materials budget	No Budget	No Budget	No Budget	No Budget
<b>Staffing (FTE)</b>	.5	As Needed*	2/3	As Needed	.5	As Needed
<b>Graduate Assistants</b>	15hrs/week	20hrs/week	-	-	-	-
<b>Programmer</b>	30 hr/month	.1	-	As Needed	.25	As Needed

\* The eArchive at IUPUI has 1 full-time librarian (ID 125).

If you want a big IR, a visible IR, you need to have a dedicated staff person. I would say no less than ½ FTE. (ID 243)

While the reallocation of staff and funding to develop the repository is a cost effective way to start a program, it is unclear how this will affect long term development. Subject repositories are having some problems, especially those started with a grant. In order to grow, a repository needs staff dedicated to identifying collections, assisting with submissions and marketing the service. Reallocating librarians for repository development may work in the short time, but a dedicated person is needed to maintain the repository and identify the appropriate people needed for additional assistance when needed.

### **5.1.8 Responsibility**

In just about every academic university and many academic sites, the concept of institutional repositories kind of came out of nowhere, it seemed, and D-Space had just released their software. There was that aspect that everybody was looking at it and thinking, “Gee, I guess we’d better get one of them there institutional repositories,” without thinking too clearly about what are we going to do with it. It seems kind of like a service, so maybe the library is a good place to put it. I don’t think there was a lot of deep thought that went into why a lot of institutions got involved with institutional repository project. They thought it should be situated in the library, and then sat down and had to figure out how to do something. (ID 305)

Out of the six repositories running at the university level, four of the six projects were initiated by library administration (Table 8). Administrative support did not mean

that the projects would be successful, but it does support projects that would not necessarily occur if the project was championed only by librarians. Administrative support was important for the repositories developed at the grassroots level. At SFU, the repository was successfully created in response to a faculty request, but with the support of the Library Director (ID 231, 279). At UBC, a new Dean of Libraries, unfamiliar with open access and repositories, did not initially provide the resources to develop the repository (ID 86, 93).

**Table 8. Responsibility Issues**

		<b>Indiana University</b>	<b>Indiana University Purdue University at Indianapolis</b>	<b>Purdue University</b>	<b>Simon Fraser University</b>	<b>University of British Columbia</b>	<b>University of Washington</b>
<b>Project Initiator</b>		Dean	Dean	Library Administration	Faculty	Faculty/Librarians	Assistant Dean
<b>Location of Project</b>		Collections	Digital Libraries	Research	Systems	Archives	Digital Initiatives
<b>Key Current Participants</b>	<b>Archives</b>		√		√	√	√
	<b>Systems</b>			√		√	√
	<b>Public Services</b>		√		√		
	<b>Technical Services/Cataloging</b>		√				
	<b>Digital Initiatives/Projects</b>			√			
	<b>Collections</b>	√					√
	<b>Library Administration</b>			√		√	√

Repository projects are located throughout the library, and no one area seems to be better suited than any other (Table 8). In the case studies, repository development was centered outside of public services (Systems, Digital Initiatives, Digital Libraries, Archives, Collections and Research). While the participants all planned to use subject liaisons to aid content recruitment, the liaison librarians were not actively involved in the

development process. Purdue University development team, which included the Head of the Physical Sciences, Engineering and Technology Division and a Geographical Information Systems Specialist, worked closely with faculty. Even with a public services presence on the team, developers are still attempting to embed the repository into the library and to “change the way we think about things and our roles” (ID 355). UBC and UW also included public services librarians in the working groups for the repositories, but were not educating the liaisons. Since the liaisons were not involved, many IR developers believe that these librarians are still unsure what the project really is (ID 109, 226, 250, 281, 431). Without a clear project definition, librarians are unsure whether the project is about preservation, open access, or electronic publishing (ID 431). The ability of the library to use repository software for multiple purposes has not created a project that is easy to market to faculty or librarians.

## 5.2 Examining Development Factors

### 5.2.1 Adoption

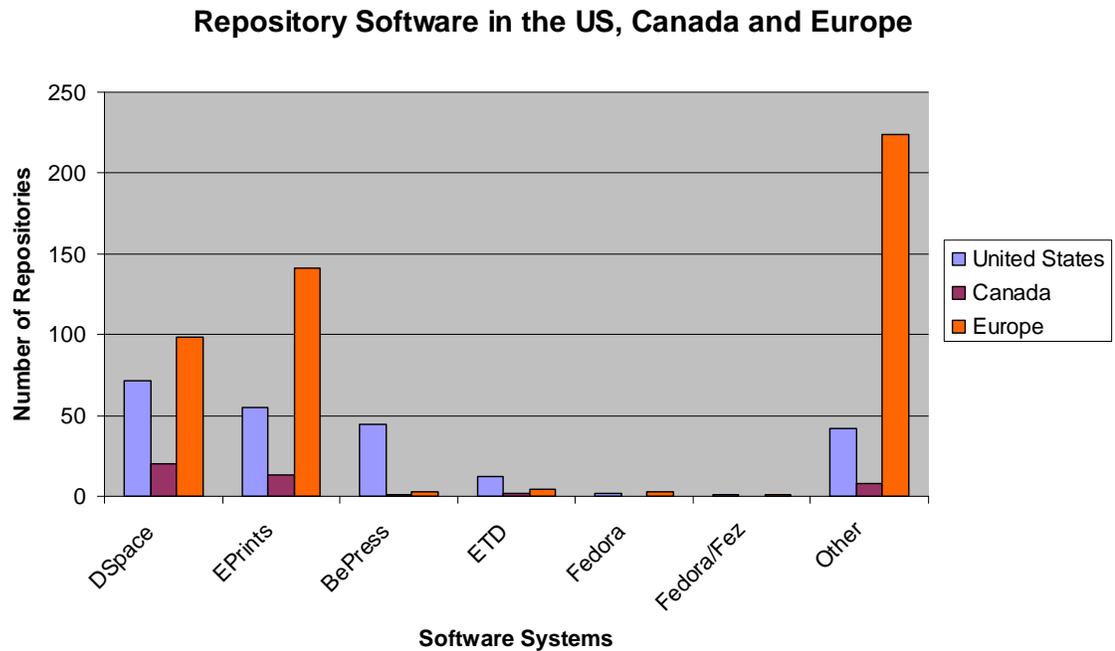
**Table 9. Adoption**

	Indiana University	Indiana University Purdue University at Indianapolis	Purdue University	Simon Fraser University	University of British Columbia	University of Washington
Early Adopter		√				
First Majority	√		√	√	√	√

When asked about IR development, most of the interview participants believed that they were not Early Adopters, but in the First Majority (Table 9; ID 106, 132, 166, 281). Swanepoel (2005) placed IR development at the Early Majority phase. This

placement was similar to that defined by the sites participating in this study, though the interviews took place approximately two years later than Swanepoel’s survey. This may be explained by the European leadership role in open access; not only does the European Commission have a proposed Open Access Self-Archiving Mandate, as of March 2008, there were over 450 repositories in Europe. Canada, on the other hand, has 44, while the United States has 227 registered repositories (ROAR, 2008).

**Figure 14. Repository software instances\***



\* Data from ROAR, March 3, 2008.

Although not specifically identified during the interviews, the slow development process may lead librarians to identify as Early Majority, since most sites began working with DSpace within months of its release. When examining repository growth according to ROAR statistics, five of the six repositories developed collections prior to 2006 (see Figure 18). U.S. repository development is centered around DSpace, while more

European institutional repositories use EPrints software (Figure 15). European and Canadian repository development may be in the First Majority stage. There are ninety-two Canadian public and private not-for-profit universities and university-degree level colleges, and forty-four repositories for ejournals, e-theses, subject repositories and institutional repositories (Table 10; Association of Universities and Colleges of Canada, 2007). U.S. institutions, dependent on DSpace, are at the Early Adopter stage.

**Table 10. Types of Repositories\***

<b>Geographic Area</b>	<b>Total Number of Repositories</b>	<b>Other</b>	<b>Database/A&amp;I</b>	<b>Demonstration</b>	<b>e-journal/publication</b>	<b>e-theses</b>	<b>Research cross institutional</b>	<b>Research institutional or departmental</b>
<b>Europe</b>	474	66	13	5	34	42	51	263
<b>Canada</b>	44	4	0	1	6	5	6	22
<b>United States</b>	227	33	5	13	11	20	17	128

\* Data from ROAR, March 3, 2007

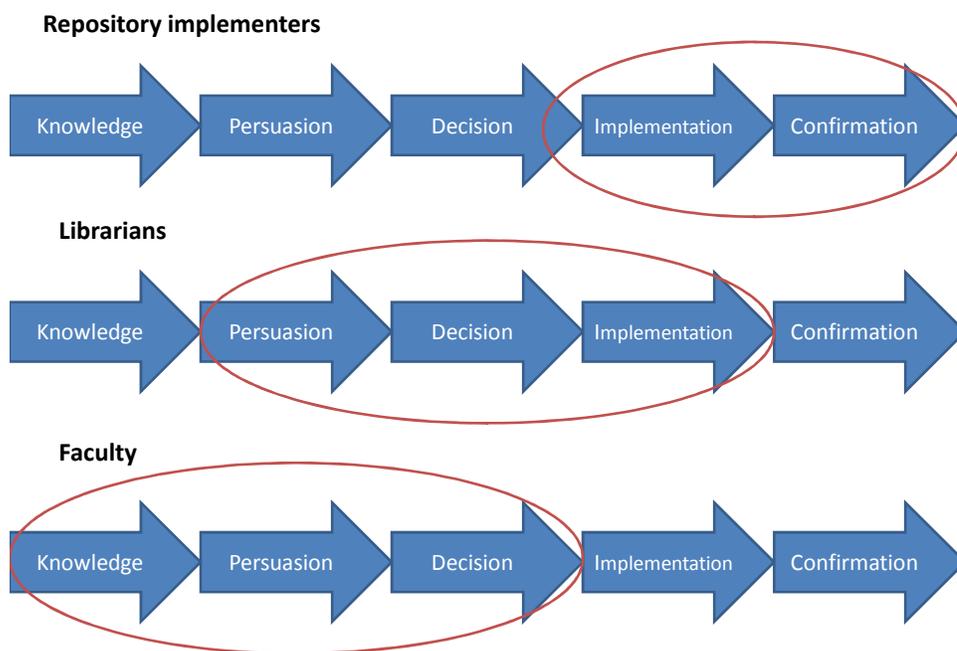
While most librarians working on repositories believe major development work is over; not all faculty agree. One faculty member at the University of British Columbia, identified by repository staff as a champion for open access on campus, believes repositories are at a much earlier stage:

It's in flux, it's in development, maybe it's not going to work, maybe people should try different things more or less. It's what happens when you are on the edge, maybe the bleeding edge. It's just, you have to take some risks and try some stuff – some stuff doesn't work and some stuff works. I think if you are clear in your overall end product, which is to have access to more information, will mean...greater chances of capturing knowledge and making discoveries. (ID 437)

Other faculty members do see this at the Early Majority stage (Jackson, 2007). But, like librarians who are not actively supporting repository developments, this may have more to do with the state of open access on a particular campus. Repository development at UBC is finally in a growth stage, despite developing a project plan more than three years ago.

At the university level, there are three separate groups working through the adoption process: implementers, librarians and faculty. Repository implementers in the case studies were all librarians, but should not be limited to librarians. The innovation-decision process is the path an individual or unit passes through during the adoption process (Rogers, 2003). The process includes: Knowledge, Persuasion, Decision, Implementation and Confirmation. Implementers have developed repositories and are now trying to figure out how to use them at the individual institutions (Figure 17).

**Figure 15. Innovation-Decision Process for Repository Implementers, Librarians and Faculty**



While the software and structure may be in place, the implementers are unsure if the repositories will be adopted by the larger campus community. Librarians and faculty are at slightly different stages in the process (Figure 17). In the process, librarians span several stages depending on familiarity with the open access movement and the librarian's personal relationships with faculty on campus. Librarians familiar with open access and who have close ties to faculty groups on campus may be Early Adopters and may be at the implementation stage. Librarians unfamiliar with open access or with fewer ties to faculty may still be waiting in the Persuasion stage, observing how the repository develops prior to making a decision. Faculty also span several stages. Early Adopters on campus may have already developed collections while other faculty members are still unsure why the library wants to collect papers and datasets.

Interpersonal relationships may also affect adoption rate because interpersonal networks and relationships increase diffusion (Rogers, 2003). Those developing the repositories, liaison librarians identifying content, and faculty may or may not have relationships with each other. This may be especially true if the repository project is located in technical services or a systems department that does not regularly interact with faculty members (ID 234). Repository developers from these areas need to develop relationships with other librarians and faculty; relationships that may not have been necessary before IR development (ID 134). Hiring a repository coordinator outside of the library may also be a detriment since someone new to the institution will not have the same relationship or trust level with faculty members (ID 58).

Faculty may be unwilling to submit items to the repository based on perceived attributes (Rogers, 2003). The DSpace menus can be complex and difficult to understand and the organizational system difficult to follow. At many sites, repository administrators are loading items for faculty and easing the process (ID 32, 44, 116, 291, 385).

Finally, explicitness can affect adoption (Rogers, 2003). The goals and rationale of the repository may not be clear to faculty or librarians. While the advantage for faculty is higher citation rates, these benefits may not be observable at the outset (Harnad and Brody, 2004). The motivating factors for developing a repository are not very explicit, and few sites have developed a narrative to explain what the repository is.

### 5.2.2 Motivating Factors

Four of the sites visited participated in the ARL SPEC Report on Institutional Repositories which asked respondents to identify which motivating factors applied to their institution's repository development (Bailey, 2006; Table 11). These motivators include implicit values of librarianship; collecting and preserving an institution's scholarship and ensuring that it is accessible.

**Table 11. What Motivated Your Institution to Establish an IR? (Bailey, 2006, p. 25)**

	Total N=67		Have <sup>18</sup> N=37		Planning <sup>19</sup> N=30	
Preserve institution's scholarship	62	93%	35	95%	27	90%
Increase global visibility of the institution's scholarship	62	93%	36	97%	26	87%
Provide free access to an institution's scholarship	61	91%	33	89%	28	93%
Collect and organize institution's scholarship in a single system	58	87%	33	89%	25	83%
Respond to requests for an IR from faculty, staff, or students	28	42%	14	38%	14	47%
Other motivating factor	15	22%	8	22%	7	23%

<sup>18</sup> Have a repository.

<sup>19</sup> Planning a repository.

It is unclear who responded on campus to the ARL survey, but, from the case study interview responses, it seems that many of the respondents may have been in library administration rather than directly involved in developing the repository. During the interviews, participants were asked what motivated the founding of the repository (Table 12). Extrinsic motivation in the form of interest from library administration in repository development was the most frequent response. Respondents were also asked about any other motivating factors. Many of the factors were the same as those identified by Bailey: access to the materials, faculty requests, visibility of the institution, and preservation; however, there were several other intrinsic factors that should be considered as well: participating in the scholarly communication process, archival support, and the development of a collaborative workspace.

During the interviews, motivation was not explicitly discussed by the interview participants. Participants did not elaborate on motivating factors; they identified outside requests for an IR, benefits of IRs, and philosophical issues for development. Many of the philosophical issues are technological frames (Orlikowski & Gash, 1994). Access and preservation are core values of library science. Because access and preservation are core values, librarians prioritize collecting and cataloging information to make it available to a specified user group if not the general public. With the development of digital collections including institutional repositories, technology is applied to these core values. According to Orlikowski and Gash (1994), the technological frame includes the “nature and role of the technology”, as well as the “conditions, applications, and consequences of that technology in particular contexts of use” (p. 178). The librarians are applying a technological frame by preserving print and electronic materials on campus and making

them accessible through the institutional repository. Further, this frame is not held by all librarians, some librarians are concerned about the potentially negative effects of open access on the scholarly publishing market, copyright or the peer review process (ID 182, 235).

**Table 12. Motivators for Developing a Repository**

	Indiana University	Indiana University Purdue University at Indianapolis	Purdue University	Simon Fraser University	University of British Columbia	University of Washington
Response to administrative interest	√	√		√		√
To provide open access to materials		√		√	√	
To preserve scholarly material on campus		√				√
To participate in the scholarly communication process					√	√
Response to requests from faculty				√	√	
To support the Archives						√
To develop collaborative workspace/file sharing space						√
To increase the visibility of the institution		√				

Both the SPEC report and the librarians involved in the case studies identified extrinsic and intrinsic motivations for IR development. The extrinsic factors in both cases are responses to requests from library administration or faculty. The intrinsic motivation is tied to the values representing the core values of the institution and librarianship.

During my discussions with librarians, faculty and administrative requests were often the

major reason for developing a repository at a university, while intrinsic factors motivated the individual librarians involved with the projects. These intrinsic factors are related to the values identified by the participating libraries in mission statements or strategic planning documents and by the American Library Association (ALA).

In 2004, the ALA Council adopted a set of core values to inform the profession (Core Values of Librarianship, 2004). Among these values are access, preservation, intellectual freedom, service and social responsibility (Core Values of Librarianship, 2004). These values correspond to several of the motivators listed in Table 12, specifically providing open access to materials and preserving scholarly materials. Other motivators can be linked to core values. Participating in the scholarly communication process and increasing the visibility of the institution can relate to the values of intellectual freedom, service and social responsibility. One of the more interesting factors is the participation in the scholarly communication process. This is moving the library from a passive role collecting information to a more active one, with the library identifying and distributing materials that may fall outside of the publication process. This active role is tied more closely to the values identified by the ALA Council (Table 13).

**Table 13. Library Goals and Values Associated with Repository Development**

University	Document	Goals and Values
University of British Columbia	Strategic Plan <sup>20</sup>	<ul style="list-style-type: none"> <li>• Advance the preservation of and access to institutional knowledge</li> <li>• Improve and expand access to collections</li> <li>• Provide access to and promote use of new methods of scholarly communication</li> </ul>
University of Washington	Strategic Plan	<ul style="list-style-type: none"> <li>• Enriches the quality of life and advances intellectual discovery by connecting people with knowledge</li> <li>• Values service, scholarship and respect and resources</li> </ul>
Purdue	Strategic Plan <sup>21</sup>	<ul style="list-style-type: none"> <li>• Values democracy, inclusiveness, and accessibility</li> <li>• Preserve the intellectual and cultural record</li> </ul>
IUPUI	Mission Statement <sup>22</sup>	<ul style="list-style-type: none"> <li>• To create unique scholarly resources</li> <li>• To enhance the availability of scholarly information</li> <li>• Values diversity, opportunity, and academic and intellectual freedom</li> </ul>
Indiana University	Mission Statement <sup>23</sup>	<ul style="list-style-type: none"> <li>• Become active partners in research</li> <li>• Emphasize the role of the IUB Libraries as the university's primary provider of educational and scholarly resources</li> <li>• Preserve digital resources</li> </ul>
Simon Fraser University	Strategic Plan	<ul style="list-style-type: none"> <li>• Enhance scholarly communication</li> <li>• Expand virtual library</li> </ul>
	IR Mission Statement <sup>24</sup>	<ul style="list-style-type: none"> <li>• Capture, preserve and communicate digital intellectual output of SFU faculty and researchers</li> <li>• Provide long-term storage and preservation of digital products</li> </ul>

<sup>20</sup> *The University of British Columbia Library's Strategic Plan: Furthering Learning and Research 2004 – 2007*. University of British Columbia Library. Retrieved 9/25/2008 from [http://www.library.ubc.ca/home/planning/UBC\\_SP\\_04-07\\_FA.PDF](http://www.library.ubc.ca/home/planning/UBC_SP_04-07_FA.PDF).

<sup>21</sup> *Purdue university libraries strategic plan 2006-2011*. Retrieved 9/25/2008, 2008, from <http://www.lib.purdue.edu/admin/stratplans/plan2011.html>.

<sup>22</sup> *IUPUI university library mission*. Retrieved 9/25/2008, 2008, from <http://www.ulib.iupui.edu/about>.

<sup>23</sup> *IUB libraries: Mission*. Retrieved 9/25/2008, 2008, from <http://www.libraries.iub.edu/index.php?pageId=5429>.

<sup>24</sup> *Mission*. SFU Library. Retrieved 9/25/2008, 2008, from [http://www.lib.sfu.ca/about/projects/institutional\\_repository/about/mission.htm](http://www.lib.sfu.ca/about/projects/institutional_repository/about/mission.htm).

While similar to preservation, the idea of stewardship was also identified as a motivating factor within the desire to preserve access to scholarly communication on campus. More frequently discussed in terms of land or natural resource management, stewardship, conceptually, combines sustainable management with preservation and protection of long-term integrity (Stewardship, 2007). Conceptually, the definition is similar to the mission of institutional repositories to preserve and make available the scholarly output of a particular institution. It also reflects definition of *kuleana* previously discussed in terms of responsibility. Both stewardship and *kuleana* represent values promoted and upheld by most librarians and combine the ALA core values of preservation, public good and social responsibility. While providing long-term preservation and access to materials located within the repository is the *kuleana* of all librarians at the institution, not just those directly responsible for the development, though not everyone incorporates these values into their work (ID 46).

As librarians are developing the repositories and working with faculty on various collections, the repository software can also become a collaborative workspace. Collections of digitized materials can be placed in a closed collection while metadata is added and the items preserved until a project is complete (like the Buddhist manuscripts project at the University of Washington), thus enabling librarians to provide a new service to the university community.

### **5.2.3 Planning**

Project planning was an area that did not get much attention from the project teams. No formal assessment of faculty interest was done on campus prior to repository development. Two sites did, however, work with faculty prior to repository development.

Purdue University began the process by speaking with faculty members and departments about their scholarly communication needs. They first identified what faculty were interested in doing with their scholarly output and then identified the tools that were necessary to meet the needs.

The University of British Columbia conducted a faculty survey on scholarly communication (ID 184). The survey contained questions about open access publishing and will be used to identify faculty who already publish in open access sources and may turn into Early Adopters on campus. The survey took place about the same time that the library held an informational session about open access publishing. They were able to discuss what open access means and clear up some of the misconceptions associated with open access at a forum with more than seventy faculty members.

Identifying both faculty needs and interest in scholarly communication provided the librarians at each institution with information that was used to make decisions about software and content recruitment.

Although many of the sites began with pilot projects, none of the sites developed assessment measures and Indiana University was the only site to develop a business plan.

Writing that business plan was torture. It was really horrible, and nobody made me do it. I just couldn't figure out any way to move ahead, and the directors needed something to approve. You know I couldn't find any other structure to present this stuff, although a lot of people said they didn't like it because it didn't answer the questions that they had about the project. They didn't think the report was structured in a useful way. So, maybe I could have found a different form, but we did a business plan and an action plan, and it got approved – it was great and it got us off the drawing board, which we had been sitting on for a long time. (ID 72)

Although the business plan framework was not intended for this type of project, the exercise helped identify what the issues were, including the resources and staffing needed for development of the repository (ID 71). Once it was accepted, the business plan was used to identify the project's missions and goals and to set up a timeline for development (ID 71). This attempt by UBC personnel to conduct planning prior to launching an IR was atypical.

#### **5.2.4 Stakeholders**

In addition to the problem of a general lack of planning, stakeholders on campus were not always identified prior to repository development. Often, the repository development was managed by a single person or team that did most of the content recruitment and management. While having a single person responsible for the repository was recommended by IR developers, stakeholders should be incorporated into all aspects

of the project. The Health Services Library at the University of Washington is running its own instance of DSpace that contains collections from the medical school on campus. Unfortunately, the only librarian working on the repository left. Nothing has been deposited in the repository since February 2007 (J. Araki, personal communication, July 19, 2007). Stakeholder groups on campus include liaison librarians, faculty, administration and students.

Campus information sessions were scheduled for faculty interested in open access, but not necessarily open or marketed to the librarians (ID 46). Not only is this a new product for librarians and faculty, but content recruitment is a new role that librarians may not be comfortable with (ID 94, 109). After UBC hired a .5 FTE coordinator for their repository pilot, the coordinator has worked with campus librarians to develop working groups for different areas. More than thirty-nine librarians on campus are actively working to develop policies, metadata standards and to recruit content (S. Taylor, personal communication, July 19, 2007). The IUPUI repository staff began holding informational sessions for librarians, the first of which occurred before classes began in January 2007 (ID 109). However, if the liaison librarians had been included in repository planning, content identification and recruitment may have developed faster.

While librarians will be included in repository development at some point, university administration is often left out of early development discussions. Repository developers with ETDs found that developing relationships with graduate divisions was key to a successful project. It was easier for librarians to develop ETD projects if they were already part of the print dissertation process.

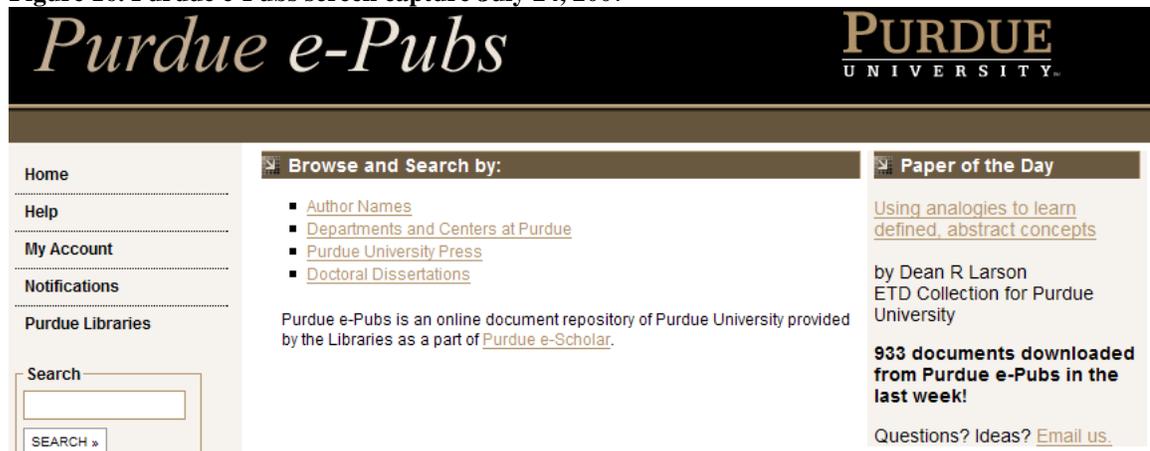
The faculty is another stakeholder group that is often overlooked in repository development. Repositories that developed without stakeholder input are often marketed as a tool for scholarly output. This tool-based approach does not assist faculty with their scholarly publishing needs. Repositories, like Purdue’s eScholar, that were developed after identifying faculty needs, have garnered more interest from departments and faculty than ones that began as pre-print and post-print archives.

### 5.2.5 Assessment

One of the difficulties with assessment identified by participants was the changing nature of the repository. Not only are the repositories still under development, but those involved are still identifying “sustainable material” (ID 407). One librarian noted that it was too “early in the process to develop good qualitative measures” and that the repositories need more than a year or two of growth (ID 216).

Institutions have been looking at the quantitative data to identify trends and usage statistics (ID 73, 150, 216). Many sites, like the Purdue e-Scholar, post the number of documents downloaded from the site for a particular time period (Figure 17).

Figure 16. Purdue e-Pubs screen capture July 24, 2007



IUPUI has identified the need for both qualitative and quantitative measures (ID 150). Although they have yet to identify measures, they are planning to hire an assessment librarian for the library (ID 150). IUPUI is also developing end user surveys for the ETD project, they want to determine which content provider, DSpace or ContentDM, is preferred for certain e-journals (ID 120, 150). Other libraries identified the need for further assessment training and plan to send someone on staff for training (ID 363).

**Chart 1. Assessment Measures**

	<b>Current Assessment Measures</b>	<b>Potential Measures</b>
<b>Indiana University</b>	None	<ul style="list-style-type: none"> <li>Collecting statistics for trend analysis, but project is changing too much to develop measures (ID 73)</li> <li>Ask faculty if there has been any noticeable or measurable impact (322)</li> </ul>
<b>Indiana University Purdue University at Indianapolis</b>	Surveys (ID 363)	<ul style="list-style-type: none"> <li>Need both qualitative and quantitative measures (ID 150).</li> </ul>
<b>Purdue University Simon Fraser University</b>	None	<ul style="list-style-type: none"> <li>Faculty feedback (ID 365)</li> <li>Develop systematic approach to populating the IR (ID 230)</li> <li>Examine log files/requests (ID 23)</li> <li>Changing impact factors</li> <li>Development of impact factors associated with the repository</li> <li>Comments from different users (271)</li> <li>Feedback forum, e.g., an associated wiki</li> <li>Social software</li> <li>End user comments (ID 300)</li> </ul>
<b>University of British Columbia</b>	<ul style="list-style-type: none"> <li>None</li> </ul>	<ul style="list-style-type: none"> <li>Usage</li> <li>Peer reviewed material (ID 216)</li> </ul>
<b>University of Washington</b>	<ul style="list-style-type: none"> <li>None</li> </ul>	<ul style="list-style-type: none"> <li>Developed an assessment working group</li> <li>Meets the business needs of the university</li> <li>Develop core of sustainable material</li> </ul>

While assessment measures are still being developed, participants found that the number of peer reviewed publications may not be a true measure of success (ID 6, 14, 19, 81). Most institutions found it easier to develop quantitative measures of success. Those measures include the number of requests, the amount of peer-reviewed materials, and impact factors of some sort.

Qualitative measures from user feedback and faculty feedback were also identified. Articles in repositories, especially on specialized topics, are finding a new audience. One author mentioned that he has been contacted about articles available in IUScholarWorks (J. Jackson, personal communication, July 3, 2007). Finally, repository working groups are attempting to develop processes that will provide both quantitative and qualitative measures for assessing the repository. Repository developers are attempting to develop approaches to populate IRs, identifying areas where the IR can meet the business needs of the institution by providing a shared workspace, linking materials to faculty and department annual reports (ID 30, 37, 47, 351).

### **5.2.6 Defining Success**

When directly asked about success, the institutions replied that they were in the process of developing qualitative and quantitative assessment measures. At least one institution was not entirely convinced that enough sustainable content had been collected to develop outcome measures (ID 407). It was also noted that the process of developing the repositories and collecting the content, especially with ETD collections, took more

time than anticipated and that the repositories were not as far along as initially anticipated (ID 12, 289). Librarians were able to identify potential measures of success, including faculty involvement, the uniqueness of content and whether the repository met the needs of faculty and the university (Chart 2).

**Chart 2. Definitions of Success**

<b>Indiana University</b>	<ul style="list-style-type: none"> <li>• Identifying relevant content on campus</li> <li>• Preserving the relevant content</li> </ul>
<b>Indiana University Purdue University at Indianapolis</b>	<ul style="list-style-type: none"> <li>• Implement continuous improvement</li> <li>• Provide the service</li> </ul>
<b>Purdue University</b>	<ul style="list-style-type: none"> <li>• Problem Solving scholarly communication issues of faculty (ID 359)</li> </ul>
<b>Simon Fraser University</b>	<ul style="list-style-type: none"> <li>• Stable access to quality material</li> <li>• Faculty involvement</li> <li>• Author awareness of rights with publishers</li> <li>• Routine self-archiving</li> <li>• Provide a safe home for the content</li> <li>• Sensible and feasible preservation practices in place</li> <li>• Provide visibility for the author(s) and institution</li> </ul>
<b>University of British Columbia</b>	<ul style="list-style-type: none"> <li>• Number of hits</li> <li>• Quantity of material</li> <li>• Uniqueness of the material</li> <li>• Knowledge of the library staff</li> <li>• Faculty participation rate</li> <li>• Content representative of the university</li> </ul>
<b>University of Washington</b>	<ul style="list-style-type: none"> <li>• Fitting the business needs of the university</li> </ul>

While assessment outcomes and measures have not been identified at the case study sites, several developers identified measure that identify failing projects. Lack of faculty participation was identified most readily with failure; however, the content that is being collected and preserved in the repositories is not necessarily faculty dependent (ID

12, 13, 189). It was noted that there are benefits, but the difficulty lies with defining them, describing them, and actually confirming that they are happening. By identifying the more subtle types of linkages or relationships or outcomes, than quantitative of indicators of success, will allow the repository developers to assess the impact of a repository, not just its size (ID 324).

### **5.2.7 Content**

The opportunities, first of all, I don't think we've begun to really properly inventory or identify potential resources out there. And I think it's because we've opted to approach it from a pretty dull strategy that somehow we pop up an institutional repository and individual faculty will just somehow contribute to it because it is a good thing. I think the ones that have been successful are the ones that have done a little in the sense of market research (ID 312).

Content in repositories does vary considerably by institution. Although many primarily collect peer-reviewed faculty output, electronic theses and dissertations were also frequently collected (ID 334). Theses and dissertations provide repositories with a unique base collection. While many schools have had difficulties changing the process for submitting dissertations and theses into a repository, four of the six sites have developed a relationship with the graduate division on campus. Since there is a process in place to collect the print dissertations, many universities have been able to edit the process to include electronic submission or offer students the choice of depositing the paper. Not only does this provide greater access to the content, most dissertations and

theses were only available in print on campus through University Microfilms (UMI), and now through ProQuest's Digital Dissertations. In many cases, masters' theses are only available from departments on campus and the content is not available elsewhere. Adding the content to a repository makes it more accessible and showcases a university's scholars.

In addition to ETDs, repository groups are identifying digital documents and collection across campus that need to be preserved. Identifying these collections is difficult, they are often projects that faculty are working on or archiving on a personal webpage or departmental server. This is especially true of campuses without large digital library collections. Indiana University is focusing on ETD, pre-prints, post-prints and journals because other types of content are better served by the Digital Libraries Program.

**Table 14. Content**

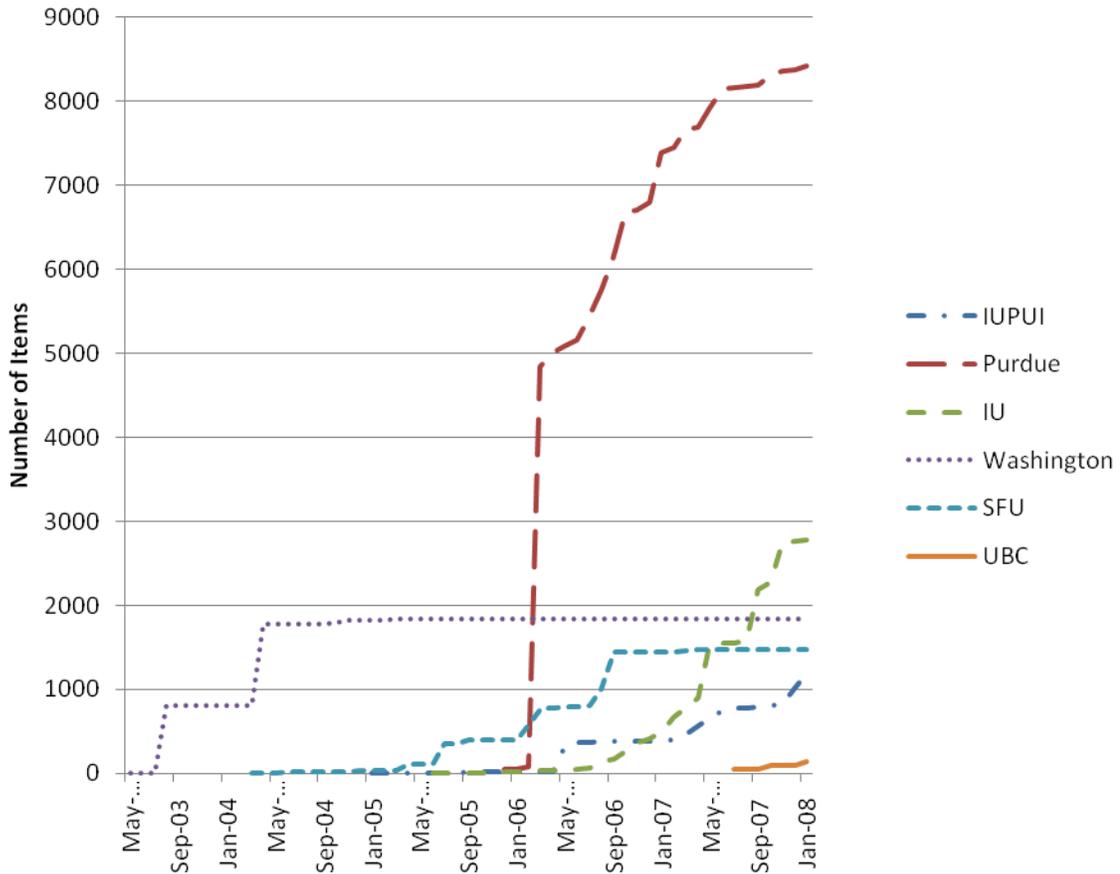
	Indiana University	Indiana University Purdue University at Indianapolis	Purdue University	Simon Fraser University	University of British Columbia	University of Washington
ETDs	√	√	√	√	Initiated as a separate project	Accompanying Materials
Pre-prints	√	√	√	√	√	√
Post-prints						√
Digital Learning Objects						√
Conference Proceedings		√		√		
Faculty Presentations		√				
Administrative Reports			√			√
Archival Materials		√	√			√
Technical Reports		√	√			
Related to Statewide or Provincial Organizations		√		√	√	
Journal Archive	√	√				

Sites without robust digital library development are using the repository to house materials that may not otherwise be archived in a print collection. Not limiting the collections to faculty and student scholarship may create a repository that can adapt to the needs of the campus.

#### **5.2.7.1 Content Recruitment**

Many repository collections have been growing since case study interviews took place between December 2006 and March 2007 as sites market the services to faculty and further define the repository's role on campus (Figure 18). cIRcle, UBC's information repository, was officially launched in the spring of 2008 (cIRcle, 2007). It is interesting to note that UBC has chosen to call their DSpace environment an information repository, rather than an institutional repository. The initial scope of the UBC project included a plan to archive documents from the 2010 Winter Olympics in Vancouver (ID 83). Using the broader term information may help the repository acquire regional collections of interest to scholars.

Figure 17. Repository Growth (data from ROAR)



For the most part, the repositories in this study have been slow to add content. In addition to setting up the software, branding the repository and identifying collections to incorporate, librarians may have to verify copyright, digitize documents and add the appropriate metadata. Content is often batch-loaded, creating small monthly increases. While IUPUI, IU, UW, SFU and UBC have chosen DSpace, Purdue is using Digital Commons. A ProQuest product, Digital Commons incorporates Digital Dissertation records from the institution. Purdue gained more than 4,000 dissertation records, but only has full-text access on campus. Other institutions, like SFU, have been scanning these,

but have to wait for copyright permission before the full-text theses are available to the public. In fact, SFU has developed a successful scanning procedure for theses and was contracted by the University of Victoria to digitize 500 theses (ID 255).

As the repositories grow, it seems that there is a critical mass of articles needed in the repository to before the repository has been truly adopted on campus. By adding content from ProQuest's Digital Dissertations, Purdue added over 4,000 items to the repository. In examining Purdue's sustained growth, this may have created a critical mass of documents and led to greater interest from the campus community. Incorporating ETD collections into the repository provides a model for other groups on campus and adds to the critical mass.

### **5.2.8 Policy Development**

There are some differences in policy development between institutions. Three models were identified in the case studies:

- The repository group creates base policies on group formation and acceptable content.
- The repository group lets individual groups create policies.
- The repository group sets no policies.

Faculty at many of the institutions expressed fear that the librarians were going to become the gatekeepers for acceptable content. Some repository developers believed that policy development would restrict submissions and interest. If the repository develops too

many policies about metadata standards, document removal, and content format, the repository may not be flexible enough to archive content that needs to be preserved. It should be noted that the only content the repositories accepted directly from students was dissertations and theses; other student content was allowable if part of a collection developed by a faculty member, department or school.

### **5.2.9 Marketing**

None of the repositories have developed long-term marketing plans. Developing a brochure or a handout is helpful for both librarians and faculty, but a larger effort is needed to ensure repository growth (ID 59). Creating a larger stakeholder group during repository development also creates a larger network for diffusing information about the repository.

### **5.2.10 Sustainability**

While many participants were not convinced that DSpace is sustainable, they are sure that the collections developed can migrate to another platform if needed (ID 186). Those involved with collection development are more concerned with identifying materials that should be preserved before they disappear in print or electronic forms than with the limitations of the chosen repository software.

### **5.2.11 Perceived Benefits**

I believe, and believe is the operative word – that there are benefits there, but how we actually define them and describe them and actually confirm that they are happening, it's going to take even just the passage of some period of time and I think an ability to look for much more subtle types of linkages or relationships or outcomes, than...more straightforward indicators of success. (ID 324)

While many of the perceived benefits respond closely to the motivating factors identified by Bailey (2006), repository implementers also identified other benefits. These benefits included changes in the role of librarians on campus, creating a dialogue with faculty, and becoming more involved in research (Chart 3).

**Chart 3. Perceived Benefits**

<b>Case Study Site</b>	<b>Perceived Benefits</b>
Indiana University	<ul style="list-style-type: none"> <li>• Scholarly communication discussions between faculty and librarians</li> <li>• Content accessible to researchers</li> <li>• Providing a home for undergraduate research</li> </ul>
Indiana University Purdue University at Indianapolis	<ul style="list-style-type: none"> <li>• Interacting with faculty</li> <li>• Being involved in something new to the library world</li> <li>• Centralized space for particular content – e.g., Conference proceedings</li> </ul>
Purdue University	<ul style="list-style-type: none"> <li>• Access</li> <li>• Collects the intellectual output of the institution</li> <li>• Value adding to content (not just a storage spot)</li> <li>• Problem solving (id 373)</li> <li>• Working with faculty and departments in new ways</li> <li>• Making existing resources more usable/accessible</li> </ul>
Simon Fraser University	<ul style="list-style-type: none"> <li>• Experience with digitization projects</li> <li>• Central space to store items</li> <li>• Creating a “permanent” archive (ID 297)</li> <li>• Provides a “safe” home for digital materials</li> <li>• Visibility for authors</li> <li>• Visibility for the institution</li> <li>• Providing open access to materials</li> </ul>
University of British Columbia	<ul style="list-style-type: none"> <li>• A chance to make discipline specific material more accessible</li> <li>• Being able to “put stuff out there in a systematic way”</li> <li>• Make the intellectual output of the university accessible</li> <li>• Creating a leadership role for the librarian in sc</li> <li>• Support the learning, research, teaching of a university</li> <li>• Showcase for work that happens at a university</li> <li>• Building bridges across campus (faculty, IT services...)</li> <li>• Taking the library in a new directions</li> <li>• Becoming more involved with research on campus</li> </ul>
University of Washington	<ul style="list-style-type: none"> <li>• Identify the scholarly output of the institution</li> <li>• Store and preserve hybrid and digital content</li> <li>• Central system</li> <li>• Provides a “space” for and access to collections that would not be easily accessible elsewhere</li> <li>• Aids dialogue about abstract concepts</li> </ul>

While most of the benefits identified deal with the long term access and preservation of scholarly materials at an institution, the individuals interviewed also identified the perceived benefits as intrinsic motivators. Malone and Lepper (1987) identified factors that promote intrinsic motivation, what people will do without external inducement, as challenge, creativity, control, fantasy, competition, cooperation, and recognition (Table 15).

**Table 15. Intrinsic Motivation, Factors and Benefits of IR Development**

<b>Factors Identified by Malone and Lepper</b>	<b>Perceived Benefits from the IR Case Studies</b>
Challenge	<ul style="list-style-type: none"> <li>• Problem solving</li> </ul>
Creativity	<ul style="list-style-type: none"> <li>• Being involved in something new to the library world</li> <li>• Value adding to content (not just a storage spot)</li> <li>• Working with faculty and departments in new ways</li> <li>• Taking the library in a new directions</li> </ul>
Control	<ul style="list-style-type: none"> <li>• Centralized space for particular content – e.g., Conference proceedings</li> <li>• Aids dialogue about abstract concepts</li> </ul>
Cooperation	<ul style="list-style-type: none"> <li>• Scholarly communication discussions between faculty and librarians</li> <li>• Interacting with faculty</li> <li>• Support the learning, research, teaching of a university</li> <li>• Becoming more involved with research on campus Building bridges across campus (faculty, IT services...)</li> <li>• Make the intellectual output of the university accessible</li> </ul>
Recognition	<ul style="list-style-type: none"> <li>• Creating a leadership role for the librarian in scholarly communication</li> <li>• Showcase for work that happens at a university</li> </ul>

By associating the intrinsic motivating factors with the perceived benefits, repository development is viewed as a challenging and creative opportunity. It allows librarians to cooperate with faculty and to create access to scholarly materials. It may also provide librarians with a leadership role on campus.

## **Chapter 6. Conclusions and Future Research**

### **Recommendations**

Though they differ in size and scope, institutional repositories and subject repositories pose many of the same problems with faculty participation, resources and marketing. While repositories in both the United States and Canada are developing similarly, there appears to be an added emphasis in Canada on making publicly-funded research publicly available.

This chapter revisits the research questions that guide this study:

1. How are IRs developing?
2. What are the factors influencing IR development?
3. How should IRs develop?
4. What are the best practices for IR development?
5. What is the best model(s) for IR development?

The previous chapter focused on the development of IRs and the factors influencing them. This chapter identifies additional factors from the case studies, identifies best practices and presents a model for IR development.

The primary findings of this study relate to the importance of incorporating these components into IR development:

- The use of a narrative approach;
- The addition of project management practices; and
- The development of campus communities.

These factors identify some of the issues and concerns associated with institutional repository development.

### ***6.1 A Narrative Approach to IR Development***

I don't think most universities have really figured out exactly what they want to do to define what the institutional repository for their particular campus is. (ID 307)

Throughout the case study interviews, participants often had difficulty explaining what an IR is and why faculty should deposit scholarly works into it. The development of a narrative and the use of storytelling can provide a foundation for librarians to use when discussing the IR with other librarians, faculty and administrators. The narrative explains what influenced the development of the institutional repository, unlike a strategic plan or a collection development plan which focus on what should happen in the future.

### **6.1.1 Narrative Development**

According to Czarniawska (2004), “a narrative is understood as a spoken or written text giving an account of an event/action or series of events/actions, chronologically connected” (p. 17). The narrative provides context to events and links intention with setting. The use of narrative acculturates societies and provides legitimacy to events or actions. A narrative can be used to discuss the development within the library and with the greater campus community to promote IR development. The development of an institutional repository is a strategic change that shifts the role of the library from an information storehouse to an information provider. The changing natures of scholarly communication, information technology and software development have made this shift possible through the development of Internet publishing tools and software systems. Within librarian communities, there is uncertainty around the change from a passive to active role in scholarly communication. The use of a narrative during a strategic change creates understanding by portraying events surrounding the change to the campus community (Dunford & Jones, 2000). Announcements of changes and new policies may cause defensive attitudes within an organization. Narratives provide a non-threatening alternative to broadcast new initiatives.

According to Pentland (1999), a narrative “describes the process, or sequence of events, that connects cause and effect.” Pentland (1999) identified the features of a narrative as having a sequence in time, focal actors, an identifiable narrative voice, an evaluative frame and other indicators of content and context (pp. 712-713). A narrative for institutional repository development provides a context for developers to explain the

library's involvement in scholarly publishing and the importance of librarian and faculty participation. In terms of the institutional repository, the narrative, similar to a project scope, defines the project, outlines why the project is important and can identify outcomes and assessment measures. The narrative encodes the data that shapes IR development and should address the importance of the IR to faculty and the institution. It also provides librarians with a plan to present to faculty and identifies the types of information that are initially wanted for the collection, providing the answer to the question "what is this for?" (ID 420).

Creating a narrative for IR development limits the number of competing accounts of change that may occur at an institution. Those working directly with the IR had similar narratives and most were aware of competing accounts that did not show the repository in a positive light. At Simon Fraser University, there were reservations on campus about open access and what it means to scholarly communication (ID 301). Developing a narrative that discusses the open access movement and what an institutional repository can do should allay those reservations. Many librarians were unsure about what to tell faculty members when asked why someone should participate (ID 420). A narrative identifying why an IR is important to the campus and how it can help faculty members would equip a librarian who is doing content recruitment with a ready response to hesitant stakeholders (ID 60). Finally, identifying the types of collections the IR should contain can identify assessment measures. For example, an undergraduate institution may wish to focus on student research projects rather than faculty publications, especially at an institution that is focused on teaching. By identifying an evaluation frame, such as student research, content recruitment can focus on departments or groups on campus with

the appropriate research projects. Once the groups are identified, qualitative and quantitative statistics can be collected; the number of projects from a school, department, or class can be used as an outcome measure.

By identifying motivation and benefits prior to content recruitment, repository implementers are able to develop a narrative to share with other librarians and faculty. Outlining the benefits of open access and what that means to the faculty and the institution, the narrative can create an explicit story to illustrate the value of participation.

### **6.1.2 Storytelling and organizational change**

Stories have been employed to convey a message or teach a lesson through allegory, fables and propaganda. Stories encode information and allow for individual listeners to apply perspectives and develop insights (Gargiulo, 2006). According to Gargiulo (2006), “stories work best in relation to one another” (p. 7). Using a story to explain changes in scholarly publishing and the importance of open access while discussing institutional repositories can be “more persuasive” and provide a “tool for thinking” (Gargiulo, 2006, p. 7). The stories contextualize change and create a positive atmosphere by addressing fear, resentment or difficulties associated with the new process (Gargiulo, 2006, p. 9).

#### **6.1.2.1 Storytelling within the Library**

Denning (2001) identifies storytelling as a way to transmit new concepts within an organization, enhancing or changing an individual’s perceptions of change by supplementing abstract analysis. The use of a “springboard story” is provided to listeners as a visualization tool (Denning, 2001). The story provides a framework for individuals

to contextualize change. In developing an institutional repository, the library is moving from a passive storehouse of scholarly communication to a proactive publisher. Denning (2001), while working at the World Bank, asked, “What does our organization do?” to develop a springboard story. The World Bank is not just a financial institution, but an information organization that has to develop tools to share its knowledge outside of the organization (Denning, 2001). In a similar vein, the Internet is changing the role of the library and the way its users access information. By asking “What does our library do?” librarians may develop a story that articulates the change internally to create support for an institutional repository and develop goals.

#### **6.1.2.2 Storytelling with Faculty**

Stories can both encourage acceptance and promote understanding within an organization. Inviting faculty to talk to librarians and other faculty members about experiences with the IR allows others to see how the IR may affect them (ID 369). The stories can be beneficial to librarians who get feedback about the process and how individual faculty members are using the IR.

Gargiulo (2006) developed a table to help individuals capture stories; the stories do not have to be long or complicated. Although the examples in the text are personal, the table can be applied to organizational settings and may assist the development of a narrative. The table consists of three columns: the description of the story, the trigger, and the connection-relationship to other stories and possible applications. Institutional repository development contains multiple stories: the changing nature of scholarly communication, serials pricing, librarian-faculty collaboration and open access

publishing to name a few. Applying the three characteristics identified by Gargiulo (2006) may assist librarians by identifying stories to include in conversations with faculty. The table briefly describes the stories, identifies triggers, and lists the connections that can be associated with IR development (Table 16). The trigger is a word associated with the story, something Gargiulo (2006) identifies as an index to the story. In examining the stories around IR development, it can be a conversational trigger for a librarian to market a repository. The connection-relationship is how the story relates to other stories.

**Table 16. Examples of Stories Associated with IR Development.**

Description of the story (ies)	Trigger	Connection-relationship to other stories and possible applications
<ul style="list-style-type: none"> <li>• The Internet and scholarly publishing.</li> <li>• IR as a tool for scholarly communication.</li> </ul>	Scholarly publishing	Share stories with faculty about open access, the ability to develop an online journal, the ability to share data, or making grey literature available.
<ul style="list-style-type: none"> <li>• IR benefits for faculty</li> </ul>	Tenure	Share stories with new faculty and/or graduate students about IR benefits to encourage use.
<ul style="list-style-type: none"> <li>• Faculty/library collaborations</li> </ul>	Grants	Share stories about the IR as a home for grant projects, a platform for research, and an opportunity to collaborate on a project with library faculty.
<ul style="list-style-type: none"> <li>• Showcasing what a college or university does</li> </ul>	Legislature	Share stories with administrators about the IR as a showcase for the scholarly output of the institution.

## ***6.2 Project Management***

Components of a project plan that can be used by other institutions for IR development were identified through the case studies. The components include:

- A working group;
- A project planning;
- Resource allocation;
- A marketing plan; and
- Assessment.

These components can be used to develop a charter for project management and measurable outcomes for assessment.

### **6.2.1 Working Groups**

One key aspect of repository development is a working group that explores the repository's role on campus and develops a narrative for the project. Working for a common goal, the group should include stakeholders from outside the library as well as librarians. Diversity in the working group should assist developers by identifying potential challenges and blocks before the repository is live. The working group can assess the need for an IR on campus and start developing other key components such as a project plan or identify digital collections that may be suitable for the repository. The

IRWG at UBC surveyed faculty, developed a project charter, and identified collections to incorporate into the repository. Creating a large, diverse group provides developers with a cadre of champions who can educate librarians and faculty or work on particular aspects of the project as needed.

Some repositories in the case studies included catalogers in the development team, but did not include public services librarians. Repository developers could not always find a role for the catalogers, but there was need for liaison librarians (ID 68, 127, 294). Librarians often see metadata creation as a crucial role during IR development. However, most of the case study sites ended up using author supplied keywords within the department structure and did not use a metadata librarian as much as anticipated. Most content is uploaded by a librarian who can make some subject classification decisions and include all relevant metadata instead of being uploaded by an author or departmental assistant who may skip key fields. Including catalogers in a working group would keep them involved in the project until their input is needed. Based on the case studies, more liaison librarians should be included in the working group, if not on the development team, than catalogers.

## 6.2.2 Project Planning

Developing a project charter can be associated with the narrative. A project charter outlining repository milestones and deliverables can be used to reinforce the narrative and can link the narrative with specific outcomes. The charter identifies:

- The resources needed to complete the project;
- Who is responsible for what aspects of the charter; and
- How the project is going to be assessed.

The software system should also be agreed upon during this phase and pilot collections should be identified. Electronic Theses and Dissertations (ETDs) can be an easy collection to include in the repository if the library has an established relationship with the graduate school.

I think picking a really good starting point with the thesis project, that, I think was really inspired ...It was a good way to get it started and it's a lot better to go out and talk to people about something when there is something, anything, in there, rather than a big empty container or something (ID 325)

In addition to low-hanging fruit, like ETDs, the working group can identify faculty that have collaborated with librarians on other projects or are known as Early Adopters on campus. It is never too early to start discussing the repository with faculty, departments and centers.

Of the case study sites, only UBC and IU developed any kind of plan before repository development (ID 71, 98). A business plan and an action plan were developed

at IU which outlined the IR mission and action plan and provided a development timeline (ID 71). While the IU timeline was used primarily for administration, it has project management potential. The UBC Project Charter outlined the role of the repository, a project plan, goals and objectives, and identified collections to develop (ID 98). Rather than attempting to talk to everyone, the planning documents identify Early Adopters and collaborators on campus who may be more likely to participate in this project. A project plan provides developers with a systematic way to incorporate materials on campus that avoids duplication of effort and makes assessment easier.

### **6.2.3 Resource Allocation**

Once the software is running, staffing must be assigned or reallocated to the repository, even if it is only a percentage of a person's time. To insure success, the IR needs a person to act as a contact for faculty and librarians and to delegate development to other librarians. While it is unclear whether this person needs to be in technical services or in public services, it is necessary to have at least a .5fte librarian working on the project. In either case, the repository needs someone who understands how both the library and the campus operate, who can motivate other librarians, and who can keep the project moving.

A dedicated repository staff member maintains both workflow and narrative consistency. Working on the IR, when other assigned duties are complete, is not as effective and may not take advantage of all the internal resources, primarily people, available within the library. Rather than merely adding records when time permits, a dedicated project manager can create more efficient workflows and processes to upload documents with the appropriate metadata while working with liaison librarians to market the repository.

While some institutions initially identified a project director, the working group may benefit if led by a library administrator. This administrator can provide additional support and importance to the project in meetings and can become a champion for the repository in the future. The Associate Dean for Research at Purdue University Libraries is active with repository development, initiating IR development as a research project (ID 355). Among the Purdue University Libraries (PUL) initiatives are to encourage interdisciplinary research and to increase the number of librarians working with researchers on campus as co-principal investigators (PI) on grants (ID 353). PUL started an initiative to look at grant administration on campus and is establishing guidelines and templates so grant seekers can include librarians and the repository in proposals (ID 354).

#### **6.2.4 Marketing**

A marketing plan is essential, not just to increase faculty awareness, but to increase awareness among administration and staff. Most of the information technology projects the library undertakes are completely internal, such as an online catalog, interlibrary loan services, or citation linking. The repository's success depends on faculty

participation, requiring librarians to interact with faculty outside of the traditional scope of library projects. It is imperative to let faculty know why they should participate and how they will benefit. The narrative contains this information and becomes the basis for all marketing documents and presentations, further enforcing and sharing the intention of the IR within the institution.

While brochures explaining IR development were created at several case study sites, most sites found face-to-face meetings with individuals and groups more effective (ID 46, 160, 295). Education is also a large part of marketing efforts on campus and within the library. The development of an education plan for librarians and faculty opens a dialogue about scholarly communication on campus, which can again hinge on the narrative. While the repository is a showcase for both faculty and student research conducted at an institution, it is competing with an accepted publishing model. Marketing should identify the incentives for publishing in an open access repository; for instance, items available in open access repositories are more findable and have higher citation rates. If the faculty or librarians do not understand what open access is or why the repository is important, there is no incentive to participate, regardless of the story associated with it.

### **6.2.5 Assessment Measures**

There is a certain kind of library culture too. We're so damn finicky and perfectionist in some of these things. We obsess over all the little things that didn't quite work out the way we thought they might or should, and kind of forget to step back and actually think "let's ignore the little bumps" – damn, it's quite impressive what we have done. (ID 326)

The narrative can be used to develop outcomes that assess the repository. While IR development is still in its infancy and assessment measures are still being developed, if development is not assessed, problems will not be identified in a timely manner, and the repository can fail. The assessment measures are not just a gauge of success, but identify areas that need improvement and verify whether marketing and education plans are working. Collecting this information provides a snapshot of where the IR started and how it can grow. It also provides developers with feedback and the realization that something has been accomplished. This opportunity for reflection on the process is an important to the growth and sustainability of the repository.

During this study, the case studies were sent to participants for review. Several participants noted that the case studies let them see how far along they have come in the process (K. Zimmerman, personal communication, March 3, 2008; J. Bobay, personal communication, March 8, 2008; H. Colenbrander, personal communication, March 4, 2008). The case studies gave the participants an opportunity to review what had been accomplished since the IR process began.

### 6.2.5.1 Measuring Success

Quantitative measures of success, such as download and usage statistics, the number of peer-reviewed articles and percentage of faculty participation, may provide some measure of the health of a repository. However, quantitative statistics are not and should not be the only measures. Rather than solely relying on quantitative statistics, it may be more useful to identify objectives from a narrative that identify how the repository serves the business needs of an institution. All of these objectives should be incorporated into the narrative. While repositories can do this by providing a centralized location to store and share materials, they can also identify what kind of research is being done at a particular school. It can increase interactions between faculty and librarians. And finally, it can educate faculty about open access and copyright.

I don't think most universities have really figured out exactly what they want do from whenever to define what the institutional repository for their particular campus is. They haven't then come up with what's the best strategy that we are actually going to implement in order to make it happen. And then part of that gets into consulting with your community that you're going to be providing things. (ID 307)

Developers at the case study sites had difficulty identifying assessment measures other than the quantitative ones provided by the repository software system because they have not defined what the repositories should be to the campus; no one has decided what the repository is to the campus. Instead, developers plan to wait and see what happens. By developing a narrative that describes what the repository is supposed to do, the

developers can create outcomes and then recruit appropriate content. In addition, the narrative can be used to garner administrative support by tying the university's mission or vision statement to the repository by getting input from the campus community.

For example, repository development at a master's level institution could include faculty and/or student scholarly work. If more emphasis at the institution is placed on teaching than on research, some faculty may perceive the repository as a threat if they want tenure or promotion. Including representatives from the campus community is essential to developing a narrative that takes the campus political climate into consideration. Many of these individuals may be Early Adopters, but they can also identify potential challenges within departments. A narrative that highlights the university's mission to facilitate student research and identifies student research as a primary collection could alleviate the faculty's concerns and ambivalence. Faculty content could still be collected, but the primary focus of the repository would be student research.

Until adding content into an IR becomes widely accepted, it may be more useful to identify outcomes that can be tracked for growth. These outcomes could then be used to realign content recruitment with objectives or identify problems with the processes associated with submissions. Assessment would be an iterative process throughout the lifetime of the repository. Before a baseline is established, developers could contact Early Adopters to identify problems. Not only would this scenario provide essential feedback, the changes or improvement should lead to faster adoption.

It's may be a much more positive [view of the world] that allows you to look forward and keep doing things and not think, jeez we better do that because we haven't sorted out all of the 1200 other little details there or we really didn't do that as well as we thought we wanted to – well, so what, we did something, and that's better than nothing. (ID 326)

Key areas of assessment include identifying collections to preserve, repository software ease of use, marketing and policy development. Outcomes, based on objectives identified in the narrative, can be used to measure the development of the repository and identify areas of improvement. The narrative can be used to develop a strategic plan. Both the narrative and plan should be revisited as the repository develops and the context around it changes.

Institutional repositories have been successful at identifying grey literature on campus. The number of items unique to a repository, items not available anywhere else, may be an indicator of how much grey literature on campus is preserved. The number of student research collections is another indicator that can be tied to the narrative. If student research is a priority for the IR, the amount of student research should increase. Since student research and grey literature are not visible on campus, an increase in the number of items identifies some success in capturing it. No growth in either outcome identifies areas that need improvement.

With librarians uploading items to the repository for faculty, ease of use is also a concern. Several sites are developing forms and processes to make depositing material easier. The number of self-archivers and the number of repeat posters can be used to measure the ease of use of the repository. If the process is too difficult, faculty will not post another article, even if the librarian acts as a mediator. If the process is easy and does not take considerable time, then the number of self-archivers should increase. If faculty do not post or self-archive, graduate students who participated in an ETD could serve as an experienced group of users who may be willing to participate in research that would increase the ease of use.

Repository marketing initiatives can be assessed through the numbers and percentages of faculty, departments and groups on campus that are participating. By identifying who is participating, librarians can focus attention on those who are not. A strategic plan may outline the order in which departments and schools are marketed to. The outcomes identify whether the marketing was successful and if it was, where to move next.

Finally, the number of unintended uses of the repository may identify the need for policy change. Initially, repository software was intended to archive pre-prints and post-prints. As the software was adopted, librarians started using it to archive grey literature and student research on campus and have plans to incorporate datasets and multimedia projects. While current repository software was not intended for this usage, local campus interest and demand have created instances of unintended use, the Buddhist manuscript project at the University of Washington is one example. The Buddhist manuscript project is using DSpace as a collaborative workspace to add metadata to images. Strict policy

development would not have made this possible. Librarians have to balance between allowing for creative use of the software and following IR policies. The narrative can play a key role here, reining in unintended uses that do not include a majority or large number of local faculty and allowing unintended usage that strategically fits the narrative.

Regardless of the measures used to assess the repository, it is much easier to assess if objectives are defined before the repository is developed. Creating a narrative document not only identifies these objectives, but provides a context and framework for a project plan, marketing and assessment within a campus community.

### ***6.3 Campus community***

An institutional repository is not just a library project; it involves the entire campus community. The IR serves as the warehouse for scholarly output of a campus and benefits all involved: faculty, with increased access to research; librarians, with alternative publishing models and means to access and preserve materials; students, by exposing theses and dissertations to a larger audience; and administration, by collecting and showcasing faculty and student research. Campus communities rely not just on the ties between members of a particular department, school or organization on campus, but on the ties between these groups and the development of communities of practice within the institution.

#### **6.3.1 Communities of Practice**

In order to succeed in higher education, all academics must develop professional networks that assist knowledge sharing and collaboration. These networks can be used to facilitate research publications in scholarly journals or to identify other academics for

joint or multi-authored publications (Lowrie & McKnight, 2004). While many of these networks are developed through personal ties, others are developed by joining an existing community of practice “of people who share a concern, a set of problems, or a passion about a topic” (Wenger, McDermontt & Snyder, 2002, 4). For scholars, these ties often relate to a discipline rather than a campus community. The development of an IR creates an online environment for scholars to share information and to create a community that may not otherwise be available.

According to Hodkinson (2004), new researchers learn how to “judge research through engaging in the authentic practices of whichever research community they join” (p.13). In some ways, the development of ETD collections within a repository are socializing new scholars to trust repositories as an outlet for scholarly communication, and it is sharing the research done by these scholars to the broader research community. Students required to publish dissertations within a repository may be more likely to publish in repositories as faculty members, especially if they have already seen open access benefits. These open access benefits are not limited to increased findability, but may identify the new scholar to a CoP associated with his or her areas of research.

Important to a community of practice is the domain, or the context of the community. Communities of practice develop context, understanding through a shared language and shared assumptions (Lesser, 2000). Narrative development provides the shared context and purpose to an IR that may not be immediately recognized by the

campus community. By defining domain issues, the community develops its own identity and purpose (Wenger et al, 1991). Having an institutional rather than a disciplinary domain may be a new development on many campuses, but may become easier with an increased interest in institutional collaboration.

Repository software allows groups and departments to develop their own norms and standards creating their own identities. Identities defined within the shared global repository that makes the purpose of the group explicit to the greater community. Defining what is acceptable to a group identifies its assumptions about scholarly communication and can be particularly useful when examining items that are not in traditional research paper format like performances, artwork, and presentations. Explicitly identifying what a group considers scholarly also helps social students and new faculty by identifying shared language and shared assumptions.

While many communities of practice develop through face-to-face meetings, virtual communities centered on a common purpose are also common. Porter (2004) defines a virtual community as “an aggregation of individuals or business partners who interact around a shared interest, where the interaction is at least partially supported and/or mediated by technology and guided by some protocols or norms” (p. 1).

Academic virtual communities create a network of peers for resource sharing and collaboration. They provide the framework for geographically dispersed scholars in a particular field with the social capital needed to become an active member in academia. Subject repositories are often developed by these communities as a way to share research with colleagues and identify collaborators from a dispersed group of scholars.

Institutional repositories can have the same role for a single campus, a multi-site university or a university system.

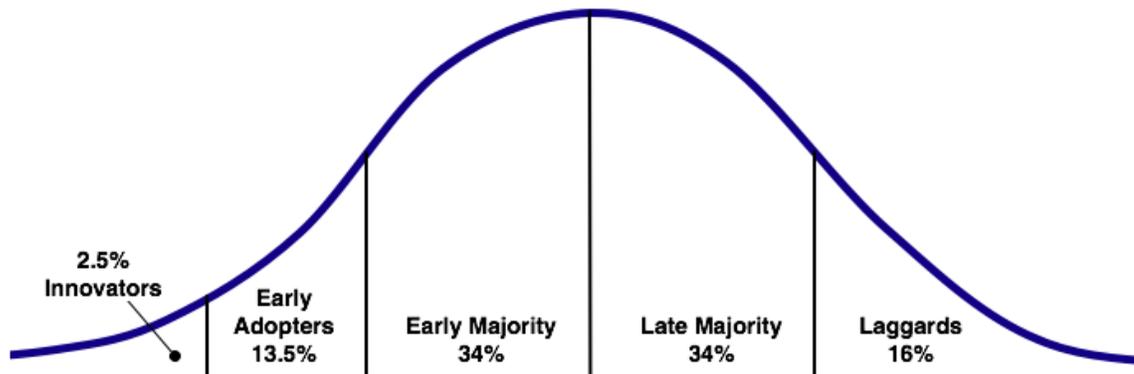
Although the scholars at an institution may be collocated geographically, they are working in distinct units and departments and may not have opportunities to interact with the larger campus. Depositing items in a shared repository may increase opportunities for scholars to identify others within the institution for collaborative research. The repository at the University of Washington has created a collaborative space for researchers working the metadata associated with the Buddhist Manuscripts project, and developers at Purdue have been approached about creating groups that include researchers outside of Purdue University (ID 380, 342). While collaborative research between faculty members through IR participation was not documented in the case studies, Churchman and Stehlik (2007) argue that CoPs within an academic institution can increase diversity by promoting collaboration and interdisciplinary research. As such, repositories are tools to promote knowledge sharing activities and cross-campus collaborations.

### **6.3.2 Diffusion of Innovation**

The development of a campus-wide community of practice provides a network for the diffusion of innovation. The innovation of an institutional repository is perceived as new not only to faculty, but to librarians. Several sites pointed out that librarians were either supportive of the project and wanted to help, interested in where the projects were going, but not sure the project was ready for faculty, or they were not interested at all (ID 46, 109, 170, 208, 226, 235, 362; Rogers, 2003, p.11). Information about the innovation “diffused” to members of the system through various communication channels (Rogers,

2003, p.5). According to Rogers (2003), adoption of an innovation is divided into five categories: Innovators, Early Adopters, Early Majority, Late Majority, and Laggards (Figure 19). Librarians working on IRs at the six case study sites believe repository development is in the second majority of adoption, repository growth and development appears to be in the Early Adopter stage in the United States.

**Figure 18. Diffusion of Innovation/Adoption (Rogers, 2003)**



Repository software at the case study sites was either chosen to meet the needs of the institution or because the software was new and open source. This latter, tool-based approach limited input from the campus community. Throughout the interviews, sites that developed a tool-based repository for scholarly communication found it necessary to meet with faculty members and discuss their needs, making the needs fit into DSpace rather than DSpace fitting faculty needs. Starkweather and Wallin (1999) examined faculty adoption and use of computer technologies in libraries. They found that late majority faculty “preferred to learn about new technologies only when they actually needed them” (Starkweather & Wallin, 1999, p.660). If this preference is true for institutional repositories, some faculty may never participate in a collection because the current scholarly publishing outlets serve their needs.

UBC and Purdue were the two sites that worked with faculty prior to repository development. Purdue worked within the social system, communicating with faculty and organizations to see if there was interest in developing a campus community, assessing needs prior to development. Librarians at UBC surveyed faculty members to identify authors who were already publishing in open access journals and to see if there was interest on campus, thereby developing a campus community of Early Adopters.

Development of a narrative can assist the diffusion of innovation by providing the IR a context within scholarly publishing and within an institution. While institutional incentives for publishing in a repository are rare, using Early Adopters to discuss benefits with other faculty members may be an effective path to diffusion. As such, Purdue is working with groups to get “testimonials and supporters who will do some of the talking for us,” making the repository an institutional project rather than a library one (ID 369).

#### ***6.4 Future Repository Development Plans***

Many sites are relying on the use of a federated search engine to streamline access to the repository (ID 365, 433, 436). A federated search engine would act as an umbrella uniting various software systems and collections through a single search interface. A repository, true to the Lynch definition of a “set of services,” could be composed of multiple systems used to best display or collect various types of items.

There is additional interest in capturing datasets, but a model for storing datasets has not been successfully developed.

## ***6.5 Resolving Anticipated Results***

During the preliminary interviews, several recommendations stood out from the conversations. The case studies confirmed these recommendations and expanded some of them.

IR developers need to define what an IR is and make sure those involved in the project have the same definition. In institutions where liaison librarians are expected to assist with content recruitment, developers must include the liaisons in the decision making process or at the very least discuss IR development with them and explain what their roles will be. Part of this can occur when an IR is defined, when its role on campus is developed, and when a project plan is developed.

Defining content early allows IR developers to identify needed permissions and stakeholders. The incorporation of an ETD collection into a repository requires permissions from multiple parties and a process change within the graduate division. Some of the case study sites took several years to move through the bureaucracy surrounding the theses and dissertations for inclusion. Working groups and departments may have their own systems and protocols in place to manage scholarly communication. Change takes time, and involving stakeholders early in the process will provide early identification of potential problems.

IR developers need to review the IR software and know its functionality and limits prior to discussing options with faculty. The case study sites used software that was on campus already or that was easy to download. Understanding the campus needs prior to choosing the system allows developers to make an informed choice that is able to handle the content effectively. Knowing what you can do is also an advantage, especially

when talking with faculty. It is not unusual to assume a particular functionality in the software and plan to use it to meet a faculty need, only to find out that functionality is not available.

Campus administration should be involved in the project. The development of an institutional repository should include institutional representatives. Including the administration lends credibility and support to the project and creates unique partnerships on campus. For instance, a campus grant writer can suggest that publishing to an IR is included in all grant proposals from the university. From the literature and the preliminary interviews, it was expected that development for depends on who is involved in the project and that multiple models exist. The initial case studies identified problems with several areas that were expected to be resolved through the development process at doctoral institutions.

Marketing the IR leads to increased faculty participation. The case studies revealed that repositories were not being marketed to faculty in a systematic way. Developers were slow to educate the liaison librarians who were expected to identify content. None of the institutions provided a narrative description of what an IR is and does which would have provided a basic understanding of the project.

Project teams have fewer challenges than individuals responsible for IR development. Project teams and individual developers had difficulties, but many of the difficulties incurred revolve around the amount of time people were able to work on the project. While a project team may create a support network for repository development, the team members may be working on several projects simultaneously and may be unable to give

the repository the same attention that an individual solely devoted to the project could. However, repositories with sole developers can stall a project if that person leaves.

There is a relationship between motivation and who is leading IR development. Development models depend on the organizational framework of each institution. Libraries can be centralized, decentralized or located on multiple campuses. The internal organizational structure of the library can be hierarchical or flat. The IRs developed from a need outside of the library or as an internal project around a product like DSpace. Internal projects were both grassroots in nature and requested by library administration. All of these factors affect how the repository develops. Repositories developed without input from public services librarians may find it more difficult to identify faculty willing to participate. The liaison librarians have ties with faculty members and may be aware of departmental collections that should be included in the repository.

The IR may be competing for resources with other digitization projects. While all of the sites had other digitization projects, none were directly competing for resources. The real need with the repositories was not technical, but social. The repository projects are competing for time from the liaison librarians and from the faculty (ID 58). While this may be true at doctoral institutions and sites with large digital library programs in place, it may not be the case at masters or baccalaureate institutions where competition for services may increase if digitization resources are limited. The limited ability to digitize will also restrict the type of content added to the repository to items that are born digital.

## ***6.6 Implications and Contributions***

This study identifies many of the difficulties associated with repository development. Although the case studies took place at doctoral institutions, the lessons from development are applicable to institutions of all sizes. Identifying best practices for development and possible outcomes for success will provide a road map for other institutions developing repositories.

IRs are a growing phenomenon with an increasing number of libraries planning to implement collections. With more than 4,000 degree granting institutions in United States, an increase in the development of repositories is expected with technology and process improvements (Pocket Guide, 2005). The sites involved in the case studies had spent more than two years developing repository collections and were still developing processes to identify scholars and collections on campus. The development of an IR is one of the more complex projects that librarians can undertake, involving a large stakeholder group, technical services, public services and administration. Gibbons (2004, p. 3) identifies “building advocacy, garnering collections, and determining policies” as areas requiring time and effort during IR development. Between November 2006 and March 2008, only twenty-six new repositories were registered with ROAR in the United States. While many are in development and may not have registered a pilot instance, the numbers may also reflect the indecision librarians have about developing repositories. Repository development requires a large amount of resources, particularly time and programming skills. Development also requires librarians to leave the confines of the library and interact with the larger campus community, a community that may not be willing to collaborate without incentives.

This study identified several primary and secondary contributions to IR research. The research provides a view of how an IR is developed that may impact other types of libraries or collections, as IRs can be developed by any organization wishing to share a digital collection among its members.

### 6.6.1 Primary Contributions

Three unique primary contributions to IR research have been discussed from the case study analysis:

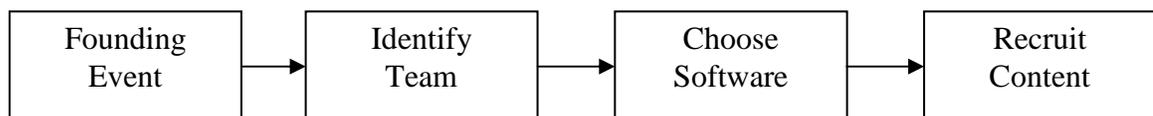
- The identification of best practices for IR development;
- The application of the narrative approach to IR development; and
- The use of storytelling to educate faculty about the IR.

These contributions combined with the case study analyses have been used to develop a model framework for development.

#### 6.6.1.1 Best Practices Framework for Development

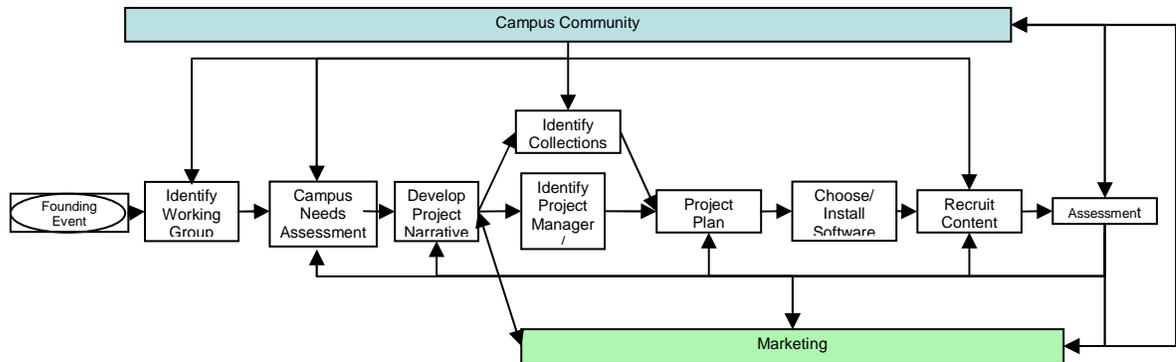
IR software models have been developed, but organizational differences and motivations lead to different developmental practices. All of the case study sites followed the same basic development path (Figure 19). Based on the case studies, a more

**Figure 19. Basic IR Development Framework**



robust developmental framework appears (Figure 20). The best practices framework incorporates the processes from the case study sites and additional factors identified from the case study interviews.

**Figure 20. Best Practices Framework Based on Case Study Analysis**



*Working Group* – Developing a working group that includes stakeholders from the campus community identifies champions on campus. The stakeholders can be early adopters or those interested in open access publishing. The working group can develop a needs assessment and build interest in the IR.

*Campus Needs Assessment* – The working group work within the campus community to identify who has materials to archive, what the materials are and where they are located. The assessment identifies what the repository can provide for the campus. In addition, the needs assessment announces the IR project to the campus and may lessen the need for liaison librarians to recruit content at a later date.

*Development of a Project Narrative* – Using the data collected during the needs assessment, the working group can create a narrative that will define the repository. The narrative is used to further specify the collections that will be relevant to the project and provides the project manager/team with a scope or vision for the repository.

*Identification of Collections* – Using the narrative for guidance, the working group and the project team can further identify collections on campus and create interest.

*Identification of a Project Manager/Team* – While a working group is useful at the beginning of repository development, there has to be a person or team responsible for development and assessment. A manager or team would be responsible for the day-to-day operations of the IR. Working group members or others with special skills could be added to the team as needed.

*Project Plan* – Tied to the narrative, the project plan identifies the resources, people and assessment measure that will be used during IR development. The plan provides goals and objectives for the project manager/team and any librarians working on collection development.

*Software* – Many of the case study sites downloaded D-Space and then identified collections. Rather than force content into the software, the collections on campus can inform software choices. The institutional repository does not have to be a single type of software, but can be a collection of services that provide the best access to the content being preserved.

*Content Recruitment* – There are many collections that can be easily added to an IR on campus, but content librarians and early adopters can also talk to individuals and groups on campus to create more interest. Storytelling can provide a way for the librarians and early adopters to share positive experiences associated with the IR.

*Assessment* – In addition to the quantitative and qualitative assessment measures, assessment can also include reflection on how the repository is developing and whether it is still meeting the needs of the campus community and realigning the narrative and

project plan to meet those needs. Although the framework above stops with Assessment, development of an institutional repository should be an iterative process, with assessment measures linking to content recruitment, collection identification needs assessment and the development of a project narrative.

*The Campus Community and Marketing* – The campus community and marketing initiatives should be present throughout several stages and should encompass repository development in this model; both are needed to create a sustainable repository. It is important to identify key stakeholders in the campus community in the repository process and include them in a Working Group. In addition, the community can provide information about collections on campus and assessment. Marketing creates awareness of the repository and is needed to diffuse information about the IR to the campus community.

### **6.6.2 Secondary Contributions**

In addition to the primary contributions identified above, this study also contains several important secondary contributions to IR research:

- An examination of project management practices for IR development;
- Knowledge of individual IR development; and
- A comparison of development across multiple cases.

All of the contributions further the understanding of IR development and identify political and social changes that occur during IR development on a campus.

## ***6.7 Limitations***

The study captured the early stages of IR development at six universities, each with different missions and organizational structures. Three subject repositories at three of the six sites were also examined. Despite the differences, many of the institutions and subject repositories were having similar problems. Multiple cases strengthen the results by replicating the pattern-matching, thus increasing confidence.

## ***6.8 Directions for Further Research***

During the case study interviews, several areas of further research can include:

1. Research on faculty publishing habits, examining who is publishing in open access repositories. By identifying how scholars publish on a campus, librarians have a better understanding of the materials that can be included in repositories.
2. Research on librarian-faculty relationships. The IRs can identify areas where librarians can collaborate with faculty and the characteristics of successful collaborations that create dynamic partnerships on campus.
3. Research what motivates librarians to participate in open access projects.
4. Research on the IR adoption rates of faculty within an institution or institutions. Questions related to adoption include: are new faculty more likely to publish in an IR. New faculty are more likely to publish a thesis or dissertation within an ETD, the response to that experience may make one more likely to deposit works into the IR.
5. A more in-depth exploration of IR development at an institution, visiting key stakeholder over a period of time during IR development and implementation.

6. The testing of the development framework at a site developing an IR.
7. Research on the knowledge of administrators, faculty and librarians regarding the development of a local IR and about the issues associated with open access and author rights, and to identify faculty already publishing in open access journals. This information will provide librarians with the information necessary to support individual campuses developing institutional repositories. It will also identify a subset of faculty interested in open access publishing for pilot groups, and it will start a dialogue with faculty about the repository. The survey results can provide a benchmark for development of institutional repositories on individual campuses.
8. An examination of the use of narrative in an IR. Collecting the stories that librarians and faculty share about depositing works in an open access repository.

## **APPENDICES**

### ***Appendix A: Interview Guide***

#### **Development**

##### ***Core Functions***

##### **Materials Submission**

- a. How are materials submitted?
- b. Do you have an editor(s)?
- c. How do you judge content?
- d. How do you know if something is appropriate?
- e. How do you know if something belongs in a collection (subdivision)?
- f. Do you take all formats?

##### **Metadata application**

- a. What types of metadata do you collect?
- b. Is it author supplied?
- c. Is it library supplied?

##### **Access Control**

- a. How do you manage digital rights?
- b. Do you authenticate?
- c. Do you support logins?
- d. Do you have to be on campus to access or submit?

### **Discovery Support**

- a. Where can you search for items?
- b. Do you belong to a larger cooperative?
- c. How can you access the IR on campus (linking)?

### **Distribution**

- a. Are there limits on viewing full-text (on-campus; by the author)?
- b. Do you need plug-ins?

### **Preservation**

- a. Do you have a preservation plan?
- b. Do you back up your data?
- c. How often?
- d. Do you convert file formats (i.e. Word to PDF or HTML)?
- e. How often will you update formats?

### ***Additional***

### **Resources**

- a. What are the sources of funding for the IR?
- b. What resources were necessary to develop the IR, such as technology or people?  
Are these resources necessary for the ongoing operation of the IR?
- c. What are the challenges and opportunities you see for the IR?
- d. If you could develop the IR again, what would you do differently?

## **Responsibility**

- a. What experience do you have in libraries?
- b. What are your current roles and responsibilities at the library?
  - i. Who initiated the project?
  - ii. Who is involved on campus?
- c. Is the IR one of many digital initiatives?
- f. Is one person or department responsible for all digital initiatives?
- g. What task roles do the different library departments (technical services, public services, archives) play in IR development?

## **Factors**

### **IR motivation**

- a. What were the events that led to the founding of the IR?
- b. Were there any forces that influenced the founding?
- c. What was the biggest motivating factor?

### **Stakeholders**

- a. What individuals, organizations or resources played key roles in the founding of the IR? What were the roles?
- b. Is there anyone you brought to IR planning after it began?
- c. Why was the person included?
- d. Is there anyone who should have been included initially?

## **Planning**

- a. Did you assess the need for an IR prior to planning? How?
- b. How did you decide what software to use?
- c. Who was involved in the initial planning of the IR?
- d. How could the process been improved?
- e. What obstacles were encountered in the founding of the IR and how were they addressed?

## **Assessment**

- a. How is the IR assessed?
- b. How would you define IR success?
- c. How are assessments being used?
- d. Do you have any examples of the benefits the IR has produced?
- e. What lessons or challenges are faced by IR developers that may impact development?

## **Content**

- a. What formats are collected?
- b. What types of materials?
- c. Who authors the materials?
- d. Is there a collection policy?
- e. Is archival material included?

## **Policy Development**

- a. Have you created any written policies?
- b. Where they created before or after you developed the IR (are they are response to a problem?)
- c. Who can deposit?
- d. What do you collect?
- e. Who can the authors be?
- f. Is there an approval process?
- g. Can things be withdrawn?
- h. Do you support versioning?
- i. Who owns the content?
- j. Is participation mandatory for some groups (dissertations)?
- k. Who checks copyright?

## **Marketing**

- a. Are you marketing the repository?
- b. How?
- c. To whom?

## **Costs**

- a. Does your IR have a budget?
- b. What is it?
- c. What does it include?

**Sustainability**

- a. Do you think this repository is sustainable?
- b. How long will it last?

**Perceived IR benefits**

- a. What are the benefits to the repository?

**Best Practices**

- a. What challenges did you face?
- b. How did you manage them?
- c. What went well?
- d. What would you do differently?
- e. What did you do right?

## ***Appendix B: Preliminary Interviews***

### **Baccalaureate - Liberal Arts (Midwest)**

The college began a pilot project open access repository during the summer of 2004. While there has been some interest in open access publishing, the repository has not received much interest outside of the library. The goal of the project is to develop a place to encourage faculty to publish into instead of journals. There has also been a brief discussion about adding archival materials.

As of March 9, 2006, there are forty-three items in the repository. Thus far, the repository contains student papers. After discussion from various departments, only papers with distinction from students who have passed comprehensive exams are included in the repository. Student permission is also required before a paper is submitted. Current student material consists of papers and a few PowerPoint presentations. Since many students produce multi-media projects, the library is looking at other software environments as well. It is not easy to include streaming video, images or other files in Digital Commons.

Other than the Digital Commons fees incurred, there is no budget. No marketing or training has occurred. There has been some discussion between the library director and department chairs involved in the pilot.

No technical problems have occurred in the pilot, ProQuest houses all the software and materials. Copyright questions have arisen, primarily, the legality of copyrighted images within student papers. It is unclear what responsibility the college has for the images used in student work.

Students involved in the pilot are excited and seem to realize the value of the repository. Without advertising, the repository gets approximately 100 hits and downloads/week. Weekly statistics on March 9, 2006 show 207 hits to the site.

Although it is still in the pilot stage, it is unclear where the project is going or how it should move forward. The priority of the project is unclear, and it lacks buy-in from the college and a clear roadmap for future growth.

### **Baccalaureate - Liberal Arts (Northeast)**

Begun in May 2005, the Digital commons IR is being run by a technical services librarian. IR development is just part of her workload. The college archivist is also involved in the project. The archivist is developing a committee and working on policy issues.

The librarian was brought into the project after the vendor was chosen by the library director. Proquest's Digital Commons software was chosen without review.

Two pilot projects were developed, one for student theses and one for a philosophical journal to see how such projects would be received on campus. No evaluation measures or policies were developed. While no budget was allocated to the project, a few thousand dollars from the Philosophy Department was used for web design.

Currently, records for over 8,000 student theses are available; however, full text is not available at this time. The college's Lecture in Philosophy series has begun digitizing presentations; full text was available in May 2006. The project has not been advertised or officially launched on campus, but the college's archive has gotten several requests for the theses indexed.

Proquest does offer training, but the session the librarian attended off-site was not robust, and more questions developed. The forum set up for Digital Commons users did not provide much information, but some documentation on non-ASCII characters from other users proved helpful. On the whole, the documentation is lacking, a users meeting at ALA Midwinter 2006 provided a forum for complaints over documentation and training. The librarian pointed out that Digital Commons uses XML, but there is no model or template available for users. Unfamiliar with XML, she searched online and found a template from another participating university. Proquest's technical assistance has been satisfactory, but slow to respond. The number of users, in the librarian's opinion, has grown faster than the technical services department. Requests for changes have been submitted, but may take a while. Proquest talks about customization, but one has to be an advanced programmer to really change the environment. Most changes are made by emailing technical support.

The college did run into a few challenges loading records. A database of student theses was already in use, but it could not be easily loaded into the Digital Commons. The data was sorted by department, and communities had to be set up prior to uploading the records. Records also had to be uploaded in batches for each community.

When developing the project, the librarian met with the Philosophy department and the Archivist before Digital Commons was set up. In retrospect, she would have taken more time learning the program before meeting with interested parties, especially since some things she thought would be included weren't.

The project may grow to include photos of student life. They are discussing how the images will work in Digital Commons with ProQuest.

### **Masters' Universities & Colleges I (South)**

A Master's Universities and Colleges I in the south developed a Digital Commons site under the direction of the Assistant University Librarian. There is no budget for the project. Money and staff time come from other areas of the library, primarily technical services.

After beginning with honors theses, a small group met to brainstorm and approved other projects (university documents, curriculum guides). No extensive policies have been developed. There was general administrative buy-in prior to the start of the project; the Vice President in charge of the library was on board, as well as the Academic Affairs offices. Working on an ad hoc basis with the Vice President of Academic Affairs, the group is trying to decide whether theses in the digital commons will be optional or mandatory. This is creating problems with some faculty members because placing the material in a repository may be considered publishing by some journals. This is especially true in the sciences. Several of the theses, in chemistry, for example, are publishable.

No evaluation measures have been developed. The library is waiting until they get more interest on campus. They have gotten informal evaluation from faculty members and library staff. Faculty members have been the least responsive to the purpose. In order to stimulate cooperation, the library used faculty recommendations to identify a couple of good theses. Once these were in the collection, the project was taken back to faculty.

Content policy is developing, and, eventually, the site could be used for anything of campus value. Currently, education curriculum guides, an e-journal and the faculty senate handbook are in the repository. Some of the theses have some multimedia component, and one presentation has been added. Video streaming is not supported in Digital Commons. Photos are in another digital media server.

Digital Commons has been shown at faculty meetings, faculty have been emailed about the project, semester talks with the departments included information on digital commons, though no end user training has occurred. It has also been advertised in newsletters and as a meeting topic.

Staff found that Digital Commons is a good turnkey product, and does not require the staffing that some other repository software packages, like DSpace, require. The site was easy to implement, but it is expensive. They are considering a move to DSpace, if they could find a hosted site.

Students are happy to see their work online, and a campus e-journal is being accessed more. There are usually a few hundred hits a week.

Currently, the staff is trying to decide whether older materials should be in Digital Commons or in the ContentDM server which provides access to the searchable text and a page image.

The library didn't fully understand the campus politics around student theses, more support from the administration is needed. More planning with Academic Affairs would have helped, though issues were not been insurmountable.

The Assistant University Librarian would have spent more time looking at software options instead of changing in the middle of a project; now it is unclear if Digital Commons is right for their content. More time should have been spent on content development.

### **Masters' Universities & Colleges I (Midwest)**

An advisory group was created for IR development. Consisting of members from the university's colleges, grants projects office and the library, the group fluctuated between DSpace and EPrints for depository software, finally choosing DSpace.

The initial concept for the IR was to preserve the intellectual output of the university, as well as share it. By placing items in an IR, the hope was to archive lesson plans and documents faculty might have on floppy disks, as well as current content. Unfortunately, faculty members have been leery of placing materials in the IR and the library has met with faculty resistance.

Student research is also in the IR. Theses and dissertations are included with the author's permission. Student participation in terms of portfolios and research presentations are increasing. Students see the value of including their work, especially as artifacts to show potential employers. The Provost of Academic Affairs has been supportive of the project as well.

The initial pilot project began with the School of Education. Although multimedia content was intended to be included, published papers were easier to include. Slowly, the library went to other departments and schools on campus trying to develop a critical mass.

To publicize the IR, the library planned workshops over the summer for faculty. Faculty visits and department meetings have also occurred. A marketing committee for the library is being developed and it is hoped that they will advertise the IR as well.

No evaluation measures have been developed, but eventually, they will decide what to quantify. No additional budget was created for the project, though the Interlibrary Loan overhead scanner is being used to digitize materials.

So far, the biggest problems have been getting faculty involved and time constraints. In hindsight, more time should have been spent on the budget, marketing and evaluation.

## **Doctoral/Research Intensive (East)**

Prior to developing the repository, two working groups were formed to define what an IR was and how it should be developed. One group was campus-wide and the other library-based. Both groups did collaborative research and evaluated software. The library chose DSpace for electronic theses and dissertation as well as research because of its open software development.

Although no general policies have been developed for the IR, a library working group has been charged with development. Although there are some new administrators at the university, the IR was cleared through Academic Affairs and the legal department.

Four test cases were developed with the College of Fine Arts and Music libraries. However, no criteria for evaluation was discussed prior to development.

The library director sees the IR as a space for learning and collaboration; however only early adapters on faculty are participating at this time.

Working with open source software has been a plus on the project. Since there is a worldwide community working with DSpace software, there is a free exchange of information and improvements.

Reflecting on the development process, the library should have defined the IR conceptually before taking the issue to campus. Without an agreed upon view of IRs and open access, the campus committee was slow to understand the concepts. More strategic planning would have helped. Faculty are not sure what the benefits of the IR are, and more could have been done to get faculty involved initially.

## **Doctoral/Research Extensive (Midwest)**

The current coordinator for scholarly communication was hired in June of 2005 after the decision had been made to use digital commons for an institutional repository. Prior to his arrival, the university had done a study and decided to contract with BEPress. Working full time on scholarly communication, the coordinator is managing and populating the collection.

The repository is divided into two sections, one for dissertations and one with faculty papers and other creative activities. A general content policy exists. The repository will include any scholarly activity associated with the university. Interest in the repository has picked up lately, and more than 1400 items have been added to the collection with research from departments, centers and research groups. It contains more faculty research than student work at this time. Some series are set for undergraduate work, and an e-journal has started, but has not published yet.

Since BEPress is not offered through ProQuest, all of the university's doctoral dissertations available through UMI's Digital Dissertations are included in the repository. Full text is available to anyone searching from the university, and over 9,500 dissertations are searchable.

No evaluation measures were developed prior to the project. Probable measures include, but are not limited to: content, usage (through downloads), faculty participation and faculty responsiveness.

The budget for the project includes the annual digital commons fee, the coordinators salary and a salary for a student worker. The project is run out of the Office of Scholarly Communication in the library.

The coordinator has met face to face with a large number of faculty members. This led to the addition of a dictionary of invertebrate zoology. The manuscript (976 pages) had been completed but the deal was cancelled just prior to publication. The coordinator noticed it during a face-to-face meeting and it was added to the collection. It has been downloaded over 8,000 times. The coordinator has also attended department meetings, presented to individuals and groups, and met with administrators. He also made sure that links to the site were in the appropriate places on the university web site.

It is a challenge to get faculty to participate. No incentives are offered to faculty, but the coordinator will do all the work (saving to PDF, added it to the collection, checking permissions ...). Time constraints, fear/distrust of online documents, and a “loss of control” seem to be limiting faculty involvement. The coordinator has been able to make connections with people across campus that have digital projects, but needed a place to archive the materials.

In addition, he has begun sending monthly download reports to contributors. The reports seem to carry some weight with faculty members, and have brought a positive response.

The coordinator would not do anything differently at this point and is happy with Digital Commons. Since it is off site, ProQuest has to deal with all the technological problems. While Digital Commons has its limits, they can be worked around or taken into account during development.

## **Doctoral/Research Extensive (South)**

The head of technical services at the university library is in charge of this project. The repository started three years ago and has seventy-five to eighty documents, and is expected to reach critical mass soon. Initially, she thought it was mainly a meta-data project.

Although a feasibility study was not done, the library believes there is a market on campus. A faculty committee was created to develop policies and grassroots interest on campus. The committee proposed a policy for collection development in 2004. That policy, although unchanged, is still considered a draft. Scholarly material from anyone associated with the university is accepted in the repository.

At the library, a scenario for digitizing equipment was created, but no formal budget exists. No evaluation measures have developed, but a small library committee was formed.

Workshops and library presentations have taken place. Although the faculty is receptive and members have documents to archive, follow-through and submission is rare. To make it easier for faculty, once a document is uploaded, the library will add all appropriate metadata. The faculty in the humanities have been the most supportive of the IR.

Student work is also included. Honor's theses from the history department will be included when copyright permission is received. Copyright permission from authors of journals published on campus has been more difficult to acquire. Electronic theses and dissertations are submitted through a separate process and are not available to the public at this time.

The project has been interesting for the head of technical services. This is the first time she has dealt with faculty members and finds it interesting to see things from the author's viewpoint. The campus has been supportive of the repository, despite the lack of follow through. The library staff has been accommodating, and she has volunteers throughout the library working on the project, including other catalogers working on metadata issues.

Identifying the projects viability would have been helpful prior to development. As part of technical services, the librarian feels she may not have been the best person to lead the project as she does not have the faculty ties someone in public services has.

## ***Appendix C: Informed Consent Form***

### **INFORMED CONSENT**

UNIVERSITY OF HAWAII AT MANOA  
AGREEMENT TO PARTICIPATE IN A RESEARCH STUDY

TITLE OF STUDY: Factors influencing the developing of institutional repositories

RESEARCHER: Jennifer Campbell-Meier, PhD. student  
1-808-282-6716  
jlcampbe@hawaii.edu

ADVISOR: Dr Rebecca Knuth  
1-808-956-3494  
knuth@hawaii.edu

---

#### **PURPOSE**

You are being asked to participate in a research study. The purpose of this study is to identify and articulate how institutional repositories (IRs) are developed to inform the policies and development of institutional repositories at other universities. This study is being conducted as a component of a dissertation for a doctoral degree.

#### **PROCEDURES**

Participation in this project will consist of taking part in an interview and/or being observed in a meeting. Interview questions will focus on participants' views of how the IR at an institution developed and who was involved. Each interview will last about one hour. Interviews will be audio recorded for the purpose of transcription and later analysis. The interviewer will observe the meeting and take notes. With participant permission, the interviews may be taped to verify interview content.

#### **RISKS**

The researcher anticipates no known risks for participation in this study.

#### **BENEFITS**

There will be no direct or monetary benefits for your participation in this study. However, the research anticipates indirect benefits as a clearly articulated knowledge of how participants develop an IR, who is involved, and what policies are needed. The research will be used to identify models for IR development.

#### **CONFIDENTIALITY**

Information obtained in this research study will be confidential to the extent allowed by the law. Data will be summarized during analysis for reporting. The researcher, the researcher's dissertation committee and the University of Hawaii Committee on Human Subjects Protection will be the only parties to have the authority to review research data and access to your identity. Research records will be stored securely by the researcher in a home office and will be shredded once the legal time limit for maintaining this data has expired.

## INFORMED CONSENT

### RIGHT TO REFUSE OR WITHDRAW

Participation in this research study is voluntary. You may refuse to participate or withdraw at any time without penalty or loss of benefit to which you would otherwise be entitled.

### QUESTIONS

If you have any questions regarding this study, please contact the researcher, Jennifer Campbell-Meier, at 1-808-282-6716 or by email at [jlcampbe@hawaii.edu](mailto:jlcampbe@hawaii.edu)

You may also contact the advisor for this study, Dr Rebecca Knuth, at 1-808- 956-3494, or by email at [knuth@hawaii.edu](mailto:knuth@hawaii.edu)

If you have any questions regarding your rights as a research participant, please contact the University of Hawaii Committee on Human Subjects at 1-808-956-5007.

CONSENT FORM

Participant:

I have read and understand the above information, and agree to participate in this research project.

---

Name (printed)

---

Signature

---

Date

## ***Appendix D: Participant List***

All interviews conducted face-to-face unless otherwise noted.

### **Indiana University**

January 4, 2007

Julie Bobay, *Assistant Director of Collection Development and Scholarly Communication*

March 28, 2007 (telephone, email)

Don Gilbert, *Lab Director, Genome Informatics Lab*

July 3, 2007 (telephone, email)

Jason Baird Jackson, *Associate Professor, Department of Folklore and Ethnomusicology*

### **Indiana University Purdue University Indianapolis**

January 3, 2007

Kevin Petsche, *Bibliographic and Metadata Services, Digital Libraries*

Karen Zimmerman, *Liberal Arts & Music, Special Collections (Archives)*

Kristi Palmer, *Bibliographic and Metadata Services, Liberal Arts*

### **Purdue University**

December 28, 2006

D. Scott Brandt, *Associate Dean for Research*

Michael Fosmire, *Head, Physical Sciences, Engineering and Technology Division*

Christopher Miller, *Geographical Information Systems Specialist*

### **University of Washington**

February 13, 2007

Ann Lally, *Head, Digital Initiatives*

Catherine Gerhart, *Music/Media Cataloger*

John D. Bolcer, *University Archivist*

William (Bill) Jordan, *Associate Dean of University Libraries, Resource Acquisition and Description/Information Technology Services*

July 19, 2007 (telephone, email)

Joy Araki, *Administrative Services*

### **University of British Columbia**

January 10, 2007

Lea Starr, *Interim Director, Information Systems & Technology/Assistant University Librarian, Science Libraries*

January 30, 2007

Christopher Hives, *University Archivist*

January 29, 2007

Hilde Colenbrander, *Institutional Repository Coordinator* (formally hired after the interview)

February 5, 2007

Francis Ouellette, *Director of the UBC Bioinformatics Centre/ Associate Professor in Medical Genetics*

July 26, 2007 (email)

Peter Seixas, *Professor and Canada Research Chair, Director, Centre for the Study of Historical Consciousness*

**Simon Fraser University**

February 27, 2007

Mark Jordan, *Head of Library Systems*

Nina Sakilar, *Systems Librarian*

Ian Song, *Digital Initiatives Librarian*

Brian Owen, *Associate University Librarian - Processing & Systems*

***Appendix E: Themes, Categories, and Sub-categories Developed for Data Analysis***

<b>Themes</b>	<b>Categories</b>	<b>Sub-Categories</b>
Content	Formats MaterialTypes Authorship ArchivalMaterials	
Policy	Written  CreationDate Withdraw Versioning Ownership Participation Copyright	Submission Deletion Collection
CoreFunctions	Submission  Metadata  AccessControl  DiscoverySupport Distribution Preservation	Process EditorialControl ContentEval Structure AuthorDriven DRM Authentication  Limits Plan Conversion Backup Updates
Resources	Funding Staffing Technology	FTE_Postitions
Responsibility	Who Initiator LibraryInvolvement	
Motivation	Factors Founding	

<b>Themes</b>	<b>Categories</b>	<b>Sub-Categories</b>
Stakeholders	CampusInvolvement LateAdditions	
Planning	NeedsAssessment SoftwareEvaluation Involvement	
Assessment	Measures DefineSuccess Usage	
Marketing	How Who	
Sustainability		
PerceivedBenefits	Repository SuccessStories	
PerceivedChallenges	Challenge Solution	
BestPractices		
Content	Recruitment Stories	
Faculty	Participation Motivation RelationshipLibrarians	
Librarians	Participation Motivation	
Adoption	EarlyAdopters	

## References

- Association of College & Research Libraries. (2003). *Environmental Scan 2003*. Shannon Cary et al eds. Chicago, ACRL.
- Association of Research Libraries. (2006). New Models for Scholarly Communication. In *Scholarly Communication*. Retrieved June 23, 2006, from <http://www.arl.org/osc/models/index.html>.
- Alavi, M. & Leidner, D.E. (1997). *Knowledge management systems: Emerging views and practices from the field*. Fontainebleau, France: INSEAD.
- Association of Universities and Colleges of Canada. (2007). *About AAUC*. Retrieved February 13, 2007, from [http://socrates.aucc.ca/about\\_us/index\\_e.html](http://socrates.aucc.ca/about_us/index_e.html).
- Bailey, C.W. (2006). *Institutional Repositories: SPEC Kit 292*. Washington, DC : Association of Research Libraries.
- Bishoff, L. & Allen, N. (2004). Business Planning for Cultural Heritage Institutions: A framework and resource guide to assist cultural heritage institutions with business planning for sustainability of digital asset management programs. Council on Library and Information Resources. Retrieved January 13, 2007, from <http://www.clir.org/pubs/reports/pub124/pub124.pdf>.
- Bock, G.W., Zmud, R.W., Kim, Y.G. & Lee, J.N. (2005). Behavioral Intention Forming in Knowledge Sharing: Examining the Roles of Extrinsic Motivators, Social-Psychological Forces, and Organizational Climate. *MIS Quarterly*, 29(1), 87-111.
- Borgatti, S. (2005). *Introduction to grounded theory*. Retrieved June 19, 2006, from <http://www.analytictech.com/mb870/introtoGT.htm>
- Budapest Open Access Initiative*. (n.d.). Retrieved November 13, 2006, from <http://www.soros.org/openaccess/>.
- Budd, J. M. & Silipigni Connaway, L. (1997). University faculty and networked information: Results of a survey. *Journal of the American Society for Information Science*, 48(9), 843-852.
- Carnegie Commission on Higher Education. (2006). *The Carnegie Classification of Institutions of Higher Education*. Retrieved June 25, 2006, from <http://www.carnegiefoundation.org/classifications/>
- Cervone, H. F. (2007). The system development life cycle and digital library development. *OCLC Systems & Services*, 23(4), 348-352.

- cIRcle: UBC's Information Repository. (2007). Retrieved March 13, 2008, from <https://circle.ubc.ca/>.
- Core values statement. American Library Association.* Retrieved 9/25/2008, 2008, from <http://ala8.ala.org/ala/oif/statementspols/corevaluesstatement/corevalues.htm>.
- Creswell, J.W. (1998). *Qualitative inquiry and research design*. Thousand Oaks, CA: Sage.
- Creswell, J.W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2<sup>nd</sup> ed.). Thousand Oaks: Sage.
- Crow, R. (2002). The Case for Institutional Repositories: *A SPARC Position Paper*. Retrieved November 25, 2005, from [http://www.arl.org/sparc/IR/IR\\_Final\\_Release\\_102.pdf](http://www.arl.org/sparc/IR/IR_Final_Release_102.pdf)
- Churchman, D. & Stehlik, T. (2007). Transforming academic work: Communities of practice in Australian universities. *Journal of Organisational Transformation & Social Change*, 4(3), 263-278.
- Czarniawska, B. (2004). *Narratives in Social Science Research*. Thousand Oaks: Sage Publications.
- Deep Blue Preservation and Format Support Policy. Retrieved December 03, 2006, from <http://deepblue.lib.umich.edu/about/deepbluepreservation.jsp>.
- Denning, S. (2001). *The springboard : How storytelling ignites action in knowledge-era organizations*. Boston: Butterworth-Heinemann.
- Denzin, N.K. & Lincoln, Y.S. (1994). *Handbook of qualitative research*. Thousand Oaks, CA: Sage.
- DigitalWell - Executive Summary. (2007). Retrieved November 24, 2007, from <http://digitalwell.org/execsummary/index.html>.
- Dunford, R. and D. Jones. (2000). Narrative in strategic change. *Human Relations*, 53(9), 1207-1226.
- Early Buddhist Manuscripts Project. (2005). Retrieved November 24, 2007, from <http://depts.washington.edu/ebmp/index.php>.
- Espisito, J.J. (2004). The devil you don't know: The unexpected future of Open Access publishing. *First Monday* 9(8). Retrieved December 9, 2006, from [http://www.firstmonday.org/issues/issue9\\_8/esposito/](http://www.firstmonday.org/issues/issue9_8/esposito/).

- European Research Advisory Board. (2006). *Scientific Publication: Policy on Open Access*. Retrieved June 13, 2007, from [http://ec.europa.eu/research/eurab/pdf/eurab\\_scipub\\_report\\_recomm\\_dec06\\_en.pdf](http://ec.europa.eu/research/eurab/pdf/eurab_scipub_report_recomm_dec06_en.pdf).
- Finholt, T. & Sproull, L. (1990). Electronic groups at work. *Organization Science*, 1, 41-64.
- Gakobo, J. (1985). The role of the special collection in the academic library. *International Library Review*, 17(4), 405-418.
- Gargiulo, T. L. (2006). *Stories at work : Using stories to improve communication and build relationships*. Westport, Conn: Praeger Publishers.
- Genoni, P. (2004). Content in institutional repositories. *Library Management*, 25(6-7), pp. 300-306.
- Gibbons, S. (2004), "Establishing an Institutional Repository," *Library Technology Reports*, 40(4).
- Granovetter, M. (1973). The Strength of Weak Ties. *The American Journal of Sociology*, 78(6), 1360-1380.
- A Guide to the Preparation of Theses and Dissertations. (n.d.). Indiana University Graduate School. Retrieved May 15, 2007, from <http://graduate.indiana.edu/thesisGuide.php>.
- Hamel, J., Dufour, S., & Fortin, D. (1993). *Case study methods*. Newbury Park, CA: Sage Publications.
- Harnad, S. & Brody, T. (2004). Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals. *D-Lib Magazine*, 10(6). Retrieved December 3, 2006, from <http://www.dlib.org/dlib/june04/harnad/06harnad.html>
- Hawkins, B. (2006a). Advancing scholarship & intellectual productivity: An interview with Clifford A. Lynch (Part 1). *EDUCAUSE Review*, March/April, pp. 46-55.
- Hawkins, B. (2006b). Advancing scholarship & intellectual productivity: An interview with Clifford A. Lynch (Part 2). *EDUCAUSE Review*, May/June, pp. 44-56.
- Hildreth, P. M. 2004. *Going Virtual: Distributed Communities of Practice*. Heshey PA: Idea Publishing Group.
- Hodkinson, P. (2004). Research as a form of work: expertise, community and methodological objectivity. *British Educational Research Journal*, 30(1), pp.9-26.

- Honhart, F.L. (1983). The solicitation, appraisal, and acquisition of faculty papers. *College and Research Libraries*, 44(3), 236-241.
- Hoyle, R.H., Harris, M.J. & Judd, C.M. (2002). *Research Methods in Social Relations*. Fort Worth, TX: Wadsworth.
- Indiana University. (2006). Indiana University Factbook 2005-2006 - Fast Facts – *IUPUI*. Retrieved November 05, 2006, from [http://factbook.indiana.edu/~urr/factbook/fbook05/fast\\_facts/fastfacts3.shtml](http://factbook.indiana.edu/~urr/factbook/fbook05/fast_facts/fastfacts3.shtml)
- Indiana University. (2008). *IU Bloomington Campus Profile*. Retrieved February 2, 2008, from <http://www.indiana.edu/profiles/iub.shtml>
- IUB Libraries Institutional Repository Pilot Team. (2005). *Business plan for IUScholarworks*. Unpublished document.
- IUBio Archive for Biology. (n.d.). Retrieved November 5, 2006, from <http://iubio.bio.indiana.edu/soft/About-IUBio-Archive.text>.
- IUScholarWorks*. (2007). Retrieved May 15, 2007, from <https://scholarworks.iu.edu/dspace/index.jsp>.
- IUPUI. (2006). *About IUPUI*. Retrieved November 05, 2006, from <http://www.iupui.edu/about/>
- Johnson, R. (2002). Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication. *D-Lib Magazine*, 8(11). Retrieved October 5, 2006, from <http://www.dlib.org/dlib/november02/johnson/11johnson.html>
- Jones, C.H. (1984). Remarks on the integration of special collections. *College and Research Libraries*, 45(6), 437-441.
- Jones, R., Andrew, T & MacColl, J. (2006). *The Institutional Repository*. Oxford; Chandos Publishing.
- Jubert, A. (1999). "Developing an infrastructure for communities of practice: the Siemens experience." *Online Information 99 Proceedings*. Hinksley Hill: Learned Information Europe, 165-168.
- Kankanhalli, A. Tan, B. & Wei, K.K. (2005). Contributing to Electronic Knowledge Repositories: An Empirical Investigation. *MIS Quarterly*, 29(1), 113-
- Kovacs, D. K., Robinson, K. L., & Dixon, J.(1995). Scholarly E-Conferences on the

- academic networks: How Library and Information Science Professionals use them. *JASIS: Journal of the American Society for Information Science*. 46(4), pp. 244-53.
- Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Lawrence, S. (2001). Online or invisible. *Nature*, 411(6837), 521-22. Retrieved December 21, 2005, from <http://citeseer.ist.psu.edu/online-nature01/>.
- Lesser, E. L. (2000). Leveraging Social Capital in Organizations. In E. L. Lesser (Ed.). *Knowledge and Social Capital: Foundations and Applications* (pp.3-16). Woburn, MA: Butterworth-Heinemann.
- Lincoln, Y.S. & Guba, E.G. (1985). *Naturalistic Inquiry*. Beverly Hills, CA: Sage Publications.
- Lowrie, A. & McKnight, P. K. (2004). Academic Research Networks: A Key to Enhancing Scholarly Standing. *European Management Journal*, 22(4), 345-360.
- Lynch, Clifford. "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age." *ARL Bimonthly Report*, No. 226, February 2003. Retrieved October 7, 2006 from <http://www.arl.org/bm~doc/br226ir.pdf>.
- Lynch, C. & J. K. Lippincott. (2005). Institutional Repository Deployment in the United States as of Early 2005. *D-Lib Magazine*, 11(9). Retrieved October 31, 2005, from <http://www.dlib.org/dlib/september05/lynch/09lynch.html>.
- Malone, T. W., & Lepper, M. R. (1987). Making Learning Fun: A Taxonomy of Intrinsic Motivations for Learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, Learning and Instruction: III. Cognitive and affective process analyses* (pp. 223-253). Hillsdale, NJ: Erlbaum.
- Markey, K., St. Jean, B., Rieh, S.Y., Yakel, E., Kim, J., & Kim, Y.M. (2006, June). *National Census of Institutional Repositories: Preliminary Findings*. Paper at the Joint Conference on Digital Libraries 2006, Chapel Hill, NC. Retrieved November 15, 2006, from <http://www.ils.unc.edu/tibbo/JCDL2006/Markey-JCDLWorkshop2006.pdf>
- McCarthy, R. (2004). Understanding project costs and building costs. *The Bottom Line*, 17(1), pp. 6-9.
- MIRACLE Proposal. (2005). Institutional Repositories: Ensuring Continued Access to *Learning Objects*. Retrieved November 13, 2006, from <http://miracle.si.umich.edu/reports/proposal.html>.

- MIT Libraries. (2004). *Work Sheet: Project Planning Timeline*. Retrieved June 15, 2006, from <http://dspace.org/implement/project-plan.pdf>
- Open Archives Initiative. (n.d.). *Mission Statement*. Retrieved November 13, 2006, from <http://www.openarchives.org/organization/index.html>.
- Orlikowski, W.J. and Gash, D.C. (1994). Technological Frames: Making Sense of Information Technology in Organizations. *ACM Transactions on Information Systems*, 12(2), 174-207).
- Pentland, B. T. (1999). Building process theory with narrative: From description to explanation. *Academy of Management Review*, 24(4; 4), 711-724.
- Pinfield, S., et al. (2002). Setting up an institutional E-Print Archive. *Ariadne 31*. Retrieved October 13, 2006, from <http://www.ariadne.ac.uk/issue31/eprint-archives/>
- Pocket Guide to U.S. Higher Education 2005. Retrieved December 15, 2005 from <http://www.educause.edu/ir/library/pdf/PUB2201.pdf>.
- Porter, C. E. (2004). A typology of virtual communities: A multi-disciplinary foundation for future research. *Journal of Computer-Mediated Communication*, 10 (1). Retrieved July 5, 2007, from <http://jcmc.indiana.edu/vol10/issue1/porter.html>
- Project: Faculty annual reporting*. (2007). University Information Technology Services. Retrieved May 15, 2007, from <http://uits.iu.edu/scripts/ose.cgi?auzy.ose.help>
- Project Management Institute.(2000). *A guide to the project management body of knowledge (PMBOK guide)*. Newtown Square, PA: Project Management Institute.
- Public Knowledge Project. (2007). About the Public Knowledge Project. Retrieved July 20, 2007, from <http://pkp.sfu.ca/>.
- Pukui, M.K. & Elbert, S.H. (1986). *Hawaiian dictionary : Hawaiian-English, English-Hawaiian*. Retrieved October 13, 2006 from <http://wehewehe.org/>.
- Purdue University. *Purdue University Facts Online*. (2007). Retrieved November 30, 2007, from <http://www.purdue.edu/facts/index.html>.
- Reitz, J. M. (2004-06). ODLIS - Online Dictionary for Library and Information Science. Retrieved November 9, 2006, from <http://lu.com/odlis/>.
- Registry of Open Archive Repositories*. (2007). Retrieved July 22, 2007, from <http://roar.eprints.org/>.

- Report of the University Librarian to the Senate 2005-2006. (2006). Retrieved June 17, 2007, from <http://www.library.ubc.ca/archives/pdfs/librarians/2006.pdf>.
- Rogers, E.M. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.
- Rowlands, I., Nicholas, D. and Huntingdon, P. (2004). *Scholarly Communication in the Digital Environment: What do authors want?* Retrieved September 14, 2005, from <http://www.homepages.ucl.ac.uk/~uczciro/ciber-pa-report.pdf>.
- Schachter, D. (2004). Managing your library's technology projects. *Information Outlook*, 8(2), pp. 10-12.
- Schauder, D. (1994). Electronic Publishing of Professional Articles: Attitudes of Academics and Implications for the Scholarly Communication Industry. *Journal of the American Society for Information Science*, 45(2), 73-100.
- Shervington, C. (1982). The establishment of University archives – the Western Australian experience. *Archives and Manuscripts*, 10(1), 20-27.
- Shearer, K. (2002). A guide to setting-up an institutional repository. Canadian Association of Research Libraries. Retrieved July 24, 2006, from [http://www.carl-abrc.ca/projects/institutional\\_repositories/setup\\_guide-e.html](http://www.carl-abrc.ca/projects/institutional_repositories/setup_guide-e.html).
- Simon Fraser University. (2006). *SFU Facts*. Retrieved November 05, 2006, from [http://www.sfu.ca/mediapr/for\\_media/sfu\\_facts.html](http://www.sfu.ca/mediapr/for_media/sfu_facts.html)
- Simon Fraser University. (2006). Simon Fraser University 2006-2007 Calendar. Retrieved November 05, 2006, from <http://students.sfu.ca/calendar/CalendarTOC.html>.
- Simon Fraser University. (2007). Simon Fraser University 2007-2008 Calendar. Retrieved July 05, 2007, from <http://students.sfu.ca/calendar/CalendarTOC.html>.
- Smith, M., Barton, M., Bass, M., Branschofsky, M., McClellan, G., Stuve, D., Tansley, R. & Harford Walker, J. (2003). DSpace: An Open Source Dynamic Digital Repository. *D-Lib Magazine*, 9(1). Retrieved May 13, 2005, from <http://www.dlib.org/dlib/january03/smith/01smith.html>.
- Soanes, C. & Stevenson, A. (Eds.). (2005). *The Oxford Dictionary of English* (revised edition). Oxford University Press. Retrieved November 10, 2006, from Oxford Reference Online.
- SPARC. (2006). *About SPARC*. Retrieved November 12, 2006, from <http://www.arl.org/sparc/about/index.html>.

- Sproull, L. & Faraj, S., 1997. Atheism, sex and databases: the net as a social technology. In: Kiesler, S., (Ed.), *Culture of the Internet*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 35–51.
- Stake, R. (1995). *The art of case research*. Thousand Oaks, CA: Sage Publications.
- Starkweather, W., & Wallin, C. (1999, Spring). Faculty Response to Library Technology: Insights on Attitudes. *Library Trends*, 47(4), 640.
- Starr, L. & Hives, C. (2006). Institutional Repository (IR) Pilot Project Project Charter. Unpublished document.
- “Stewardship”. (2007). *Definitions of Stewardship on the Web*. Google. Retrieved July 12, 2007, from <http://www.google.com/search?q=define%3A+stewardship&ie=utf-8&oe=utf-8&aq=t&rls=org.mozilla:en-US:official&client=firefox-a>.
- Swanepoel, M. (2005). Digital repositories: All hype and no substance? *New Review of Information Networking*, 11(1), pp.13-25.
- Suber, P. (2005). Open access, impact, and demand: Why some authors self archive their articles. *British Medical Journal*, 330, pp. 1097-1098.
- Tague, N. (2004). *The Quality Toolbox*, 2<sup>nd</sup> ed. Milwaukee: Quality Press.
- Tellis, W. (1997). Introduction to Case Study. *The Qualitative Report*, 3(2). Retrieved September 23, 2006, from <http://www.nova.edu/ssss/QR/QR3-2/tellis1.html>.
- Thibodeau, K. (2006, June). What constitutes success in a digital library repository? Paper at the Joint Conference on Digital Libraries 2006, Chapel Hill, NC. Retrieved November 15, 2006, from <http://www.ils.unc.edu/tibbo/JCDL2006/Thibodeau-JCDLWorkshop2006.pdf>
- University of British Columbia. (2007). *UBC Facts & Figures (2005 / 2006)*. Retrieved December 15, 2007, from <http://www.publicaffairs.ubc.ca/ubcfacts/index.html>.
- University of Washington. (2007). *Student Headcount by Campus and Term*. Retrieved January 2, 2008, from <http://www.washington.edu/admin/factbook/>.
- Wasko, M. M. & Faraj, S. (2000). “It is what one does”: why people participate and help others in electronic communities of practice. *Journal of Strategic Information Systems*, 9, pp.155-173.
- Wasko, M. M. & Faraj, S. (2005). Why Should I Share? Examining

Social Capital and Knowledge Contribution in Electronic Networks of Practice. *MIS Quarterly*, 29(1).

Wenger, E., McDermott, R. & Snyder, W. M. (2002). *Cultivating Communities of Practice: A Guide to Managing Knowledge*. Boston: Harvard Business School Press.

Wren, J. (2005). Open access and openly accessible: a study of scientific publications shared via the Internet. *British Medical Journal*, 330, 1128-31.

Widén-Wulff, G. & Ginman, M. (2004). "Explaining knowledge sharing in organizations through the dimensions of social capital." *Journal of Information Science*, 30(5), 448-458.

Wolff, J. (1981). Faculty papers and special-subject repositories. *American Archivist*, 44(4), 346-351.

Working Group for IDeA. (2007). *IDeA: IUPUI Digital Archive*. Retrieved May 13, 2007, from <http://www.ulib.iupui.edu/idea/wgroups.html>

Yin, R. K. (2002). *Applications of Case Study Research*. Thousand Oaks, CA: Sage.

Yin, R. K. (2003). *Case Study Research: Design and Methods*. Thousand Oaks, CA: Sage Publications.