

FACULTY ATTITUDES TOWARDS INSTITUTIONAL REPOSITORIES

Nathan F. Hall

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

December 2014

APPROVED:

Brian C. O'Connor, Major Professor
Scott Warren, Committee Member
Ana D. Cleveland, Committee Member
Suliman Hawamdeh, Chair of the Department
of Library and Information Science
Herman Totten, Dean of the College of
Information
Mark Wardell, Dean of the Toulouse Graduate
School

Hall, Nathan F. *Faculty Attitudes Towards Institutional Repositories*. Doctor of Philosophy (Information Science), December 2014, 122 pp., 2 tables, references, 97 titles.

The purpose of the study was to explore faculty attitudes towards institutional repositories in order to better understand their research habits and preferences. A better understanding of faculty needs and attitudes will enable academic libraries to improve institutional repository services and policies. A phenomenological approach was used to interview fourteen participants and conduct eight observations to determine how tenure-track faculty want to disseminate their research as well as their attitudes towards sharing research data. Interviews were transcribed and coded into emerging themes. Participants reported that they want their research to be read, used, and to have an impact. While almost all faculty see institutional repositories as something that would be useful for increasing the impact and accessibility of their research, they would consider publishers' rights before depositing work in a repository. Researchers with quantitative data, and researchers in the humanities are more likely to share data than with qualitative or mixed data, which is more open to interpretation and inference. Senior faculty members are more likely than junior faculty members to be concerned about the context of their research data. Junior faculty members' perception of requirements for tenure will inhibit their inclination to publish in open access journals, or share data. The study used a novel approach to provide an understanding of faculty attitudes and the structural functionalism of scholarly communication.

Copyright 2014

by

Nathan F. Hall

ACKNOWLEDGEMENTS

I would like to thank my committee for lending me their guidance, experience, and expertise. My advisor, Professor Brian O'Connor has inspired and motivated me for my entire graduate education. Professors Ana Cleveland and Scott Warren helped me organize and communicate my research rigorously and effectively. Thank you also to the faculty of the UNT College of Information for a great program that continues to advance my career. I would not have started this without the inspiration and encouragement of Dean Emeritus Don Grose of University of North Texas, and my grandfather Professor Emeritus Frank DiGangi of University of Minnesota. I received generous support from current and former colleagues and supervisors, including Gail McMillan, Julie Speer, Tyler Walters, Mark Phillips, Cathy Hartman, Don Grose, Anne Lawrence, and Gilbert Borrego. Kiri Goldbeck and Margaret Merrill helped me practice and refine my interview protocol. Thank you to Hannah Tarver, KT Torrey, Katie Garahan, Rosemary Grose, and Ellen Hall for their careful editing. I would also like to thank Christian Matheis, Amy Van Epps, Michael Khoo, George Buchanan, and the JCDL Doctoral Consortium for helpful conversations, and feedback on preliminary findings. This research would have been impossible without the support of my participants. Each one of them generously and anonymously donated their time for the benefit of my education. My parents provided catering and babysitting services, on top of endowing me with the confidence, intellectual curiosity, and perseverance necessary to pursue and complete a Ph.D. My wife Monena was incredibly supportive, allowing me to selfishly pursue this goal for eight years. In the final year she took on much more than her share of parenting our children and managing our household while also pursuing her own professional career.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	iii
LIST OF TABLES.....	vii
CHAPTER 1 INTRODUCTION	1
1.1. Overview	1
1.2. Statement of the Problem	1
1.3. Purpose of the Study.....	2
1.4. Working Definitions	2
1.5. Research Goals.....	3
1.6. Research Questions	3
1.7. Limitations and Assumptions.....	3
1.8. Significance of the Study.....	5
1.9. Dissertation Structure	5
1.10. Chapter Summary	6
CHAPTER 2 LITERATURE REVIEW	7
2.1. Overview	7
2.2. Scholarly Communication	8
2.2.1. Open Access Landscape	11
2.2.2. Research Data Policy.....	13
2.2.3. Institutional Repositories.....	17
2.3. Sociotechnical Factors	19
2.4. Sociotechnical Factors in Library and Information Science	23

2.5.	Information Behavior	27
2.6.	Faculty Attitudes	30
2.7.	Chapter Summary	31
CHAPTER 3 RESEARCH DESIGN		33
3.1.	Overview	33
3.2.	Strategy of Inquiry	33
3.2.1.	Development of Phenomenology	34
3.2.2.	Intersubjectivity in Phenomenology	35
3.2.3.	Limitations	37
3.3.	Researcher's Role	38
3.4.	Data Collection Procedures	39
3.4.1.	Bounding the Study	39
3.4.2.	Setting	40
3.4.3.	Events and Process	40
3.5.	Participants	41
3.5.1.	Participants Background	42
3.5.2.	Research Methodologies	44
3.5.3.	Levels of Engagement	47
3.6.	Data Analysis and Interpretation	54
3.7.	Validation	54
3.8.	Chapter Summary	55
CHAPTER 4 FINDINGS		56

4.1.	Overview	56
4.2.	Ways Faculty Want to Disseminate Their Research	57
4.2.1.	Increased Transparency	60
4.2.2.	Increased Impact.....	61
4.2.3.	Increased Intellectual Property Rights.....	67
4.2.4.	Status Quo.....	69
4.3.	Faculty Attitudes Towards Data Sharing.....	72
4.3.1.	Transparency.....	74
4.3.2.	Research Ethics and Protecting Human Subjects	76
4.3.3.	Context.....	78
4.3.4.	Career Protection.....	84
4.4.	Chapter Summary	92
CHAPTER 5 ANALYSIS AND SUMMARY		95
5.1.	Overview	95
5.2.	How Faculty Want to Disseminate Research.....	95
5.3.	Faculty Attitudes Towards Data Sharing.....	97
5.4.	Recommendations for Repository Services	98
5.5.	Recommendations for Future Research	100
APPENDIX A IRB APPROVAL LETTER.....		103
APPENDIX B RECRUITMENT LETTER.....		105
APPENDIX C INTERVIEW PROTOCOL.....		107
REFERENCES		111

LIST OF TABLES

Table 4.1	Novel Ways Faculty Want to Disseminate Their Research	60
Table 4.2	Attitudes of Faculty Towards Data Sharing.....	74

CHAPTER 1

INTRODUCTION

1.1. Overview

This chapter has ten sections. Section 1.2 states the problem and defines the gap in knowledge that this dissertation research addresses. Section 1.3 discusses the purpose of the study, states the reason for undertaking this research, and identifies the population that will benefit. Section 1.4 operationalizes key terms. Section 1.5 states what the dissertation should achieve. Section 1.6 defines the specific questions that this dissertation addresses. Section 1.7 discusses limitations and assumptions in the study and introduces some epistemological issues that are more fully defined in Chapter 3. Section 1.8 describes the study's significance as well as the gap in library and information science (LIS) literature that this dissertation addresses. Section 1.9 describes the organization of the remaining chapters. Section 1.10 summarizes the key points of the Introduction.

1.2. Statement of the Problem

New media and Internet technologies have transformed many aspects of life and the ways that people create, seek, use, and manage information, but they have not fundamentally transformed scholarly communication (Harley, 2013). Institutional repositories have not been as successful on a national scale in the United States as early advocates had hoped (McDowell, 2007; Salo, 2008). Some researchers question whether the institutional repository model of self-deposit by researchers is broken, especially in comparison with discipline-specific repositories, such as arXiv at Cornell and the CERN Document Server (Thomas & McDonald,

2007), which are experiencing much higher rates of self-deposit into their repositories. Since many university repository developers, managers, and administrators have remained largely uninformed of faculty needs (Foster & Gibbons, 2005), they have a poor understanding of faculty attitudes towards institutional repository services (Xia, 2011). This disconnect is not new in library technology services (Lynch, 2003b).

1.3. Purpose of the Study

This research explores the factors that affect faculty adoption of institutional repositories in order to better understand their research habits and preferences. A better understanding of faculty needs and attitudes will enable academic libraries to improve institutional repository services and policies.

1.4. Working Definitions

Section 1.4 defines key terms used in this dissertation.

- Tenure-track faculty: This study involved faculty with academic tenure, or with contracts that could lead to tenure after a set period. The faculty members in this study had responsibilities for instruction, research, and scholarship. The term “faculty” will be used throughout the dissertation referring specifically to this category, instead of faculty who only have instruction duties, or only have research or administrative and professional duties.
- Attitudes: The feelings towards something, usually as reflected in statements or behavior.

- Institutional repository: A set of tools and services that together form a system for collecting, managing, indexing, and preserving the intellectual and creative output of a university or other research organization.
- Data sharing: The socialized practice within the research community of giving and receiving digital recorded material necessary to validate research findings. See section 2.2.2 for a broader discussion of the policy environment for defining research data.
- Research: The formal academic work useful for judging the merits of a scholar. Blogs, book reviews, and editorials are included in this definition if the scholar's institution or community recognizes these materials as legitimate research.

1.5. Research Goals

The goals of this research are to:

- Produce findings that will inform academic library strategies for establishing and improving institutional repository services targeting faculty.
- Identify areas for further research that advance libraries' understanding of faculty use of institutional repository services.

1.6. Research Questions

1. How do faculty want to disseminate their research?
2. What attitudes do faculty have towards data sharing?

1.7. Limitations and Assumptions

This study uses a phenomenological approach (see Chapter 3 for a full description) to

explore faculty attitudes towards scholarly communication. The nature of phenomenology introduces limitations to the study that are fully described in Section 3.2.3.

There are a few challenges to understanding the needs of faculty in scholarly communication. It is difficult to measure or analyze human behavior without intrusive direct observation. While this dissertation refers to behavioral research, it is important to note that I did not study behavior. This study specifically analyzes faculty self-report about their behavior regarding scholarly communication and data management. The participants in the study were all tenure track faculty members from two large public universities, who studied environmental issues.

Study participants do not always express what they mean (Creswell, 2003), they sometimes do not feel there is a problem, and they sometimes have trouble expressing the problem as a need (Allen, 1996; P. Wilson, 1977). Studies of different communities in different settings using different survey tools can yield very different findings. For example, an anthropological study at the University of Rochester (Foster & Gibbons, 2005) found that digital preservation was not an important factor for faculty, yet a survey of 684 respondents from 17 institutions by Kim (2011) concluded that digital preservation was a very important factor.

With those constraints in mind, a more realistic starting point is an exploration of faculty attitudes to determine what they report they need, while framing their attitudes in cultural, economic, and policy contexts and understanding that the findings might not be the same at peer institutions.

1.8. Significance of the Study

Some existing studies explore faculty deposit patterns and attitudes (Akers & Doty, 2013; Davis & Conolly, 2007; Thomas & McDonald, 2007), but an environmental scan of relevant literature found few studies that examine the sociotechnical aspects of faculty attitudes regarding scholarly communication and data sharing in an institutional repository context, and no other similar studies that use a phenomenological approach. A shifting policy landscape (described in section 2.2.2) impacts the scholarly communication and research data sharing practices of faculty, particularly for federally-funded research, but that does not mean that academic libraries can be complacent about addressing stakeholder concerns in the development and deployment of institutional repositories. Academic libraries have new opportunities to become partners in the campus research environment via institutional repositories. This dissertation uses a novel approach to providing a better understanding of faculty attitudes and the sociotechnical factors that affect scholarly communication and data sharing practices in order help academic libraries improve services.

1.9. Dissertation Structure

The dissertation is divided into five chapters, along with supporting diagrams, and a bibliography. Chapter 1 introduces the research problem, the research goals, and the research questions, and provides functional definitions and context for the research. Chapter 2 explores the literature of information behavior and sociotechnical factors research, and discusses the relationship between the two and how they inform this dissertation research. Chapter 3 reviews literature on phenomenology and presents the research methods for collecting and

validating the data that support the findings. Chapter 4 presents the findings of the study. Chapter 5 discusses the broader impacts of the findings, the limitations of the study, the potential contributions to the field, and ideas for further research.

1.10. Chapter Summary

A multi-case study set at 45 institutions (Harley et al., 2007) concluded that attempts to influence deeply-embedded values held by faculty towards scholarly communication would fail, in the short-term, and that it would be more productive to study the scholarly communication needs of faculty and use those findings to plan for future scenarios and services. This view was shared in the conclusion of an evaluation of Cornell's institutional repository (Davis & Conolly, 2007) which states that different disciplines have different cultural norms and that institutional repositories need to address this diversity.

This dissertation uses a phenomenological approach to exploring faculty attitudes in scholarly communication in order to improve institutional repository services in academic libraries. This chapter described the problem, defined the research questions, introduced the approach and its limitations, and argued for the significance of this research.

CHAPTER 2

LITERATURE REVIEW

2.1. Overview

Chapter 2 provides broader context for the research problem described in Chapter 1 by reviewing the literature to form a basis for a strategy to address the research problem. This chapter illustrates five basic observations. 1) The current policy and technology environment enable radical change to the scholarly communication model, the model has changed little. 2) Sociotechnical systems theory is useful for examining the interaction between technology and the social context in which it operates. 3) Information behavior research sometimes reveals sociotechnical factors even though sociotechnical systems are rarely the research focus. 4) With few exceptions, neither sociotechnical systems theory, nor information behavior research have focused on formal channels of scholarly communication. 5) Recent scholarly communications research has produced data and findings, but with few exceptions has not explored deep-rooted and underlying causes for trends, nor generated theory. Together, these five basic observations illustrate that research in this area would be novel and useful, both for informing academic library scholarly communication services, and for adopting new theory in LIS research.

Section 2.2 provides an overview of relevant contemporary issues, problems, and library technologies in: scholarly communication, the open access landscape, the research data policy environment, and institutional repository services. Section 2.3 defines and describes sociotechnical systems theory, providing several examples of its application. Section 2.4 reviews the use of sociotechnical systems theory in LIS. Section 2.5 identifies sociotechnical

factors that affect attitudes and behaviors in other contexts by reviewing the published research in information seeking behavior and workplace communication. Section 2.5.1 specifically reviews information behavior and workplace communication research that examines faculty attitudes.

2.2. Scholarly Communication

The Association of Research Libraries (ARL) defines scholarly communication as “the system through which research and other scholarly writings are created, evaluated for quality, disseminated to the scholarly community, and preserved for future use” (ARL, 2013). Swan (2006b) describes differences between formal and informal scholarly communication. While technological advances have changed informal scholarly communication through channels like blogs, listservs, wikis, and online forums, formal scholarly communication still occurs through peer-reviewed journals, books, and monographs. Peer review is the essential criteria for formal scholarly communication, though some publishers are using web technologies to experiment with “open review” and other alternatives to formal peer review. Furthermore, electronic publications can enable published findings to be supported by making the research data available online, which was not possible with print journals. This has vastly improved the transparency of research in both science and the humanities. In this regard, research data is becoming part of the formal scholarly communication process, with the Joint Data Archiving Policy (Dryad, 2011) being a prime example.

Communication is an exchange of information between a sender and a receiver via a channel (Shannon & Weaver, 1948). In scholarly communication, the sender is an author or a set of co-authors, the receiver is the scholarly community, and the channel is typically a

scholarly journal, conference, website, or repository. Shannon and Weaver identify three classes of problems in this model of communication. The semantic problem is the precision of meaning conveyed. The technical problem is how accurately the message transmits from the sender to the receiver. The effectiveness problem deals with the degree to which the received meaning of the message affects the receiver's behavior. For this study, I am interested in the latter two problems.

Technical problems in scholarly communication deal with not just accuracy, but timeliness. Scholarly communication systems involve a peer review process, which is often slow and labor intensive. Some journals and conferences limit the number of accepted manuscripts to works that are relevant, rigorous, and valid, but also novel enough to generate a high volume of discourse and level of interest. A consequence of this system is that many submitted articles, which meet high standards of rigor and validity, are never published, or they are published months or years after the research is complete.

This consequence relates to the effectiveness problem. If effectiveness is a measure of impact, and if a channel regularly rejects valid research in order to be more selective, an unpublished article has little chance of affecting behavior of the community that stands to benefit from the new knowledge. Further, the audience for academic journals is quite limited, due to the financial barrier, the challenge of professional terminology, and the lack of marketing to lay audiences. Some universities address this problem through programs like an extension service or technology transfer. These aim to either freely disseminate information to benefit a community, such as agricultural land owners, or to commercialize an innovation and bring it to market for the use of firms making new technology investments.

An important issue resulting from the technical and effectiveness problems described above, in terms of Shannon and Weaver's model, is the signal-to-noise ratio in top-tier journals. Shannon and Weaver's model describes an ideal communication channel and receiver as a system that eliminates noise. Some journals identify some valid research as noise instead of signal, due to a lack of novelty, and eliminate it from the discourse in order to lower the acceptance rate and to appear more selective (*PLoS Medicine* Editors, 2006). The low incremental cost of disseminating information on the Internet has allowed many more journals of varying quality to proliferate, without any significant risk to the general business model of the top-tier journals.

Kling and McKim (1999) write that scholarly communication includes three dimensions: publicity, access, and trustworthiness. Publicity, and the verb "to publish" are both etymologically derived from the same Latin word *publicare*—to make public or generally known. Publicity is effective to the extent that target audiences become aware of the thing that is published. Access is a corollary of publicity, but long-term accessibility is dependent on a commitment of institutional resources devoted to stewarding knowledge in the form of documents. Trustworthiness in scholarly communication systems develops from peer review. Trust is qualitatively and quantitatively measured in scholarly journals through reputation and impact. Critical peer review "provides valuable functions for scholarly publication that are not effectively replaced" by self-publication on a website or repository (Kling & McKim, 1999).

Several studies have identified the culture of one's research discipline as an influential factor in scholarly communication (Akers & Doty, 2013; Kling & McKim, 2000). Fry and Talja (2007) use empirical data to analyze disciplinary differences in informal scholarly

communication. Their study identifies a number of differences in how disciplines create knowledge in online environments. Members of the physics community, for example, often work outside of formal publishing channels through the use of online preprints and disciplinary repositories. In a number of humanities fields however, some online journals have a weaker reputation and are mostly used by junior researchers seeking forums to disseminate their ideas.

The evaluation of quality is an important aspect of scholarly communication. The quality of articles and books submitted for publication is evaluated through rigorous peer review. A researcher's career is evaluated during the promotion and tenure process based in part on the number of publications and the quality (as judged by peers) of those publications. The Institute for Scientific Information also has a system that measures the number of citations per article and gives an objective value to researchers and to journals (Swan, 2006a), though the algorithms can be "gamed" by editors who ask authors to cite the editor's own journal before accepting the article for publication (*PLoS Medicine* Editors, 2006; Wilhite & Fong, 2012).

A relatively recent movement in scholarly communication rose in reaction to perceived problems with the current model of publishing. The Open Access movement challenges the commercialization of scholarship that authors provide to publishers for free.

2.2.1. Open Access Landscape

At the most basic level, open access is a set of principles applied to research literature. The main principle is that research literature is freely available without legal or financial barriers for any user to read, copy, download, and otherwise use for any lawful purpose. A number of

scholarly communities have adopted various statements to affirm open access principles, including the Budapest Open Access Initiative (Open Society Institute, 2002), the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (Max Planck Society, 2003), and the Bethesda Statement on Open Access Publishing (2003).

Open access is sometimes divided into two categories: green and gold. “Green” open access occurs through disciplinary repositories, institutional repositories, departmental repositories, or an author’s personal website. Authors can publish the authoritative version of an article in a peer-reviewed journal, and post a copy in one of those options. “Gold” open access is through conventional, peer-reviewed scholarly journals in which the author pays a fee to offset the publisher’s loss of revenue. Some journals have proliferated to take advantage of this model, though it has been demonstrated that they often do not provide peer review. The lower quality journals often publish sub-standard work, and some authors (often graduate students) with valid research will pay to publish through these outlets, without the benefit of peer review, and with no benefit to their academic careers. Some open access journals do not charge a fee, but are financially sustainable through the support of membership in an academic society, or through some other form of sponsorship, such as through a university or foundation.

While most literature on the subject finds that open access provides authors and articles with a citation advantage (Gargouri et al., 2010; Harnad et al., 2004), many researchers still choose not to make their work available on a website or repository. There are several reasons researchers do not favor open access for their own publications. These include (1) lack of awareness of open access concepts or implications; (2) lack of understanding or appreciation of the issues, often due to misconceptions and misinformation; (3) belief that access to journal

literature, or ability to disseminate publications, is not a problem; (4) lack of familiarity with “gold” open access journals; and (5) anxiety over rights management and time commitment for self-deposit in “green” open access repositories (Swan, 2006b).

2.2.2. Research Data Policy

As noted above, providing open access to published articles presents one set of challenges. Making scientific research data generally available for review presents a different set of problems. Due to developments of data policies by some research sponsors and some high-profile scholarly journals, the practice of sharing research data is an increasingly important factor in scholarly communication. Availability of research data makes peer review more rigorous because reviewers can verify an author’s findings. It also allows other researchers to advance research-based knowledge by building on the earlier datasets.

This dissertation applies a somewhat ambiguous definition for the term “research data.” The ambiguity is due in part to the nature of the current policy landscape. The National Science Foundation (NSF) states that research data covered by the NSF policy is determined by the “community of interest through the process of peer review and program management. This may include, but is not limited to: data, publications, samples, physical collections, software and models” (2010). The Final National Institutes of Health (NIH) Statement on Sharing Research Data (2003) applies to “final research data” which are defined as “recorded factual material commonly accepted in the scientific community as necessary to validate research findings.” The NIH policy excludes “laboratory notebooks, partial datasets, preliminary analyses, drafts of scientific papers, plans for future research, peer review reports,

communications with colleagues, or physical objects, such as gels or laboratory specimens.”

NIH (2004) also notes that “unique data are especially important,” where unique data are defined as “data that cannot be readily replicated” and includes examples such as “large surveys that are too expensive to replicate; studies of unique populations, such as centenarians; studies conducted at unique times, such as a natural disaster; studies of rare phenomena, such as rare metabolic diseases.”

The NSF, as of January, 2011, mandated that investigators share primary data, samples, physical collections, and supporting materials gathered using NSF funding, and that all applications must include a data management plan describing how the project will conform to NSF policy. These requirements were based on the recommendations of the 2005 report, *Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century* (National Science Foundation, 2005).

The White House Office of Science and Technology Policy Memorandum (Holdren, 2013) and OMB Circular A-110 define data as “the digital recorded factual material commonly accepted in the scientific community as necessary to validate research findings including data sets used to support scholarly publications” (Office of Management and Budget, 1999, 36.d.1.2.1).

These definitions reflect how agencies that sponsor research view data and they are the definitions that support the policies to which principle investigators and their institutions must adhere. The limitation on these definitions is that they only apply to research in the domain of “science” (e.g. physical, life, social, applied, formal, etc.), whereas federal policies also govern humanities research. The National Endowment for the Humanities (NEH) Office of Digital

Humanities policy defines data as “materials generated or collected during the course of research” and examples could include “citations, software code, algorithms, digital tools, documentation, databases, geospatial coordinates (for example, from archaeological digs), reports, and articles” (NEH, 2013, 1).

In all of these cases, research data is somewhat defined by its format, but is especially defined by its purpose to the researcher and by the researcher’s community of peers. For the purposes of this dissertation, the term “research data” assumes that the data 1) are digital (and therefore subject to the benefits and limitations of electronic media), and 2) support formal research through documentation and validation.

These federal policies have had a strong impact on libraries, research institutions, academic journals, and scholarly discourse in general. A final report from an Association of Research Libraries workshop on stewardship of digital data sets in science and engineering asserted that digital data archives, scholarly publications, and associated communication need to be closely linked in order to provide better incentives for scientists to contribute to data collections (ARL, 2006). The journal *Nature* (2009) published an editorial titled “Data’s Shameful Neglect,” which asserts that university research libraries are obvious candidates to assume the role of hosting and preserving long term access to digital data and that most disciplines “lack the technical, institutional and cultural frameworks required to support open data access ... leading to a scandalous shortfall in the sharing of data by researchers.” In the same issue, *Nature* also published an article (Nelson, 2009) about the failure of some institutional repositories to populate their collections.

Foster and Gibbons (2005) cite factors like lack of time as a reason why researchers do not deposit their data, and a study by Akers and Doty (2013) found differences in attitudes and practices regarding data management between four categories of research disciplines (arts and humanities, social sciences, medical sciences, and basic sciences). Parsons (as quoted by Nelson) found that another problem is the diversity of data sets between disciplines. Parsons' project to archive data from the International Polar Year "encountered a staggering diversity of incoming information, as well as wide variations in the culture of data sharing." There were marked differences between some fields, such as atmospheric science and oceanography, which have traditions of open access, versus wildlife ecology and social sciences, which do not, according to Parsons. Wallis and Borgman (2011) found that scientists often do not consider issues of ownership of and responsibility for research data. As noted in the *Nature* article cited above, this role of data archiving and curation often falls upon academic libraries.

Libraries have developed a number of services in order to prepare for faculty data. Some offer tools to assist with data management planning. Reilly and Dryden (2013) describe one approach at the University of Houston, and Sallans and Donnelly (2012) describe two inter-institutional efforts. In a separate type of service, the Data Curation Profile project (Witt et al., 2009) developed a tool and a method for librarians to assess the data curation needs of faculty research. The Data Curation Profile tool is an in-depth interview and survey protocol that collects qualitative data from faculty members about their research data in order to help the interviewer and interviewee better understand the decisions and resources necessary for long-term management of research data. Findings suggest that librarians who participated in training workshops to use the toolkit felt more confident about data management and curation

services and in crafting data policies, but their levels of engagement with faculty in data management did not change significantly (Carlson, 2013).

A study by Scaramozzino, Ramirez, and McGaughey (2012) explores faculty attitudes and behaviors about data curation, including a component about the importance of data sharing. They found that “while the majority of researchers believe that colleagues should share their data, only a minority of respondents actually share their own data with individuals who did not help in gathering the data” (p. 361). Researchers are increasingly expected to manage and share research data, especially for federally-sponsored research, but the reasons few researchers share scientific data are poorly understood. Some reasons have been identified, as noted in this section, but there has been little analysis of the rationale for those reasons.

Given that many faculty support the sharing of research data while not sharing their own, and given that many faculty want to disseminate their published work and reduce barriers to access while not engaging in self-deposit of materials, it is essential to examine institutional repositories.

2.2.3. Institutional Repositories

Lynch (2003a) defines institutional repositories as “a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members ... A mature and fully realized institutional repository will contain the intellectual works of faculty and students—both research and teaching materials—and also documentation of the activities of the institution

itself in the form of records of events and performance and of the ongoing intellectual life of the institution. It will also house experimental and observational data captured by members of the institution that support their scholarly activities.” In the previous sections, numerous cited works discussed institutional repositories as a likely solution to some of the issues raised, and many discussed the lack of widespread faculty adoption and use.

Institutional repositories serve a variety of purposes. Buckland (1991) writes, “Noncommercial provision of information sources is of particular interest because it provides empirical evidence of the values and social goals of individuals and groups providing information and of perceptions of the cost-effectiveness of information systems” (p. 182). In other words, a university invests in an institutional repository to demonstrate the institution’s value and impact to the community it serves.

In spite of the benefits afforded by institutional repositories, they have not been as successful in the United States as their proponents predicted (McDowell, 2007; Salo, 2008). Some research suggests that institutional repository development is too often uninformed by the needs of the target user communities (Foster & Gibbons, 2005; Xia, 2011). Some authors advocate for the importance of understanding faculty attitudes on their own terms in order to improve institutional repositories (Harley, 2007). An exploration of sociotechnical factors affecting faculty attitudes towards scholarly communication would provide a better understanding of the target audience and the context in which institutional repositories are deployed.

2.3. Sociotechnical Factors

Sociotechnical factors are circumstances resulting from the interplay of social and technical influences that affect the performance of a system (e.g. technology, organization, or structure). Sociotechnical systems theory is useful for understanding how factors outside of usability and system performance can affect behaviors and attitudes of a population interacting with a system. A number of theories with subtly different names cover the same basic principle. This section covers several examples that authors have described with various labels. Howcroft, Mitev, and Wilson (2004) for example, demonstrate a connection between sociotechnical systems theory and Bruno Latour's actor network theory. They review criticisms of technological determinist views and social determinist views of technology. In separate discussions of sociotechnical networks, Meyer (2006) hyphenates socio-technical whereas Kling (2000) does not. The terms sociotechnical network and sociotechnical system are used almost interchangeably. In spite of the differences in what these theories are called, these examples illustrate different approaches to the study of technology that incorporate cultural, organizational, economic, behavioral, and political contexts.

Several authors have used or reviewed sociotechnical approaches in a variety of research settings. Trist and Bamforth (1951) define sociotechnical systems as a problem-solving approach that examines the interactions between people, structures, organizations, and tools with the goal of optimizing the performance of all of them. Their study of social and psychological consequences of innovation in coal mining techniques was an early application of sociotechnical systems theory to understand how the context in which a technology is applied affects the technology's performance. In their analysis, organizational structure of work teams

and differences in economic opportunity between 1930s Britain and 1950s Britain were important factors that affected production, worker morale, and attrition rates after the implementation of a new method for coal mining.

Trist and Bamforth's analysis examines the combined effects of a particular method of coal extraction, known as "longwall," as a technological system with the intent of maximizing production, along with a social organization of occupational roles used to implement it during a two year period. Trist and Bamforth assumed that the technological system of the coal extraction method and a given method of social organization had interactive psychological effects on the miners.

Bijker (1995) analyzes the social construction of three different technologies (the bicycle, bakelite plastic, and fluorescent lighting) to illustrate three major points. The first point is that technologies are shaped by cultural norms, while simultaneously changing them. The second point is that innovations occur in social and cultural contexts. The third point is that technology simultaneously shapes and is shaped by political forces. An important element in Bijker's analysis is the concept of "relevant social groups" and their diverse views of what constitutes technological problems and technological solutions.

Pool (1997) conducts a sociotechnical exploration of the nuclear power industry to identify the nontechnical influences that shaped the industry's development in several different countries. His main thesis is that "any modern technology ... is the product of a complex interplay between its designers and the larger society in which it develops" (p. 6).

Pool discusses the shift in social science from a technological determinist view, where one sees technology as a driving force of society, to a social constructivist view where society

shapes technology. Some of the non-technical factors cited by Pool that determine whether a technological innovation is widely adopted include market forces and business acumen, psychology, historical trends, economics, culture, and institutionalized groupthink.

Pool retells how Edison harnessed electricity but then, more significantly perhaps, marketed it to a society in the face of public fear and competition from rivals. Lobbying by the gas industry was a major challenge to getting city neighborhoods wired for electricity. Working with urban planning officials was another challenge, as was population density in the initial markets for distribution. The debate over alternating current versus direct current presented other technical, as well as decidedly non-technical, difficulties.

Pool argues that society's relationship with technology has changed since the Industrial Revolution. Material technologies are more complex and their use sometimes has unforeseen consequences, especially when they are put in the service of billions of people instead of just a dozen or even a few million.

Pool writes that 80% of France's electricity comes from nuclear power and it has been safe, cheap, and reliable. In the United States on the other hand, where nuclear power only generates 20% of electricity, the development of nuclear power has come at great cost, environmentally, socially, politically, and economically. Pool claims that both countries use the same engineering, but that the technology operates within different sociotechnical systems.

Pool advocates an approach to sociotechnical understanding that incorporates both a positivist view as well as a social constructionist view. A technology has to work on some level to be useful, and to do so it has to be designed within certain constraints of the physical world, and that is the provenance of positivist knowledge. Yet a technology also has to operate within

the subjective constraints of the social world, and that is the provenance of social constructivist knowledge.

Brown and Duguid (2000) argue that understanding the social context in which technologies operate leads to better designs and uses. They call for a better understanding of resources and constraints. Resources are materials or traits within a system that are available for designers to utilize. Resources can be physical, in the form of matter and energy, like water, food, muscles, and paper. Resources can also be conceptual, like intelligence and charisma. Resources can even be mythologized, like Athene in Homer's *Odyssey*. Constraints on the other hand can be in the form of physical limitations on a tool. A hammer, for example, works best when used in one direction. Constraints are sometimes the target for new technologies. Electronic media addresses the limitations of paper, since it is cheaper to replicate the intellectual content of a digital file enabling many people to simultaneously read, manipulate, copy, or link to it. Brown and Duguid also note, however, that constraints and resources often overlap. For example, geographic regions with rocky soil are difficult for farmers to cultivate, yet builders in areas with rocky soil do not have to go looking very hard for rocks with which to build walls or homes. Rocks are a constraint in tilling the land, but a resource for building walls for fencing in livestock.

Social organizations and structures, like other technologies, have inherent constraints and resources, and sometimes the constraint and the resource are the same thing. A policy for evaluating merit is a resource for justifying a promotion for a person who exhibits extraordinary performance. It can also help an administrator decide which people need more support and which people need to be let go. The same policy might be constraining in that it incentivizes

people to only perform within narrow parameters, because other extraordinary work around a new type of product or service might not fit within the context of the evaluation policy. Therefore, extra work on a useful project that cannot be evaluated within the policy's framework takes away time to spend on projects that can be evaluated.

Another example that is both a constraint and a resource is intellectual property law, which is a resource for rewarding and protecting innovators to incentivize further innovation. Firms with no interest in developing and marketing an innovation, however, can buy useful patents in order to prevent a new idea from competing with their own vested interest.

2.4. Sociotechnical Factors in Library and Information Science

LIS research has long focused on the nexus between people, information, and technologies, which makes it an appropriate area for the use of a sociotechnical approach in the manner described in the examples above. Paul Otlet (1934) and Suzanne Briet (1951) used the term “documentation” to describe a field that covers human behavior in information systems. Patrick Wilson (1996) also describes a field of human sciences practiced in Europe that studies information and the social and behavioral aspects of users. Howcroft, Mitev, and Wilson (2004) assert that while sociotechnical views of technology are “almost an orthodoxy ... it is evident that, aside from a few notable exceptions, this is not the case in [information science] research” (pp. 329-330). They argue that sociotechnical views of technology are useful in the study of information systems and organizations and list a number of factors, including “organizational, political, social, economic, and cultural—that pattern the design and use of technology” (p. 329). While LIS research, such as information behavior, usability, and human

computer interaction, can examine how individuals or populations use an information system, sociotechnical approaches sometimes invert that approach to explore how a technology performs within a social context.

Meyer (2006) asserts that the use of sociotechnical interaction networks is a strategy that “leads to choosing particular methods, to favoring certain kinds of understandings about the world, but maintains the overall social informatics open-mindedness towards a variety of methods, and a preference for multiple methods approaches to research questions” (p. 44).

Rob Kling (1999) defines social informatics as “the interdisciplinary study of the design, uses and consequences of information technologies that takes into account their interaction with institutional and cultural contexts.” Kling asserts that sociotechnical systems theory views technological development as a social process. While performance, functionality, and usability are important criteria in developing software or hardware, they are not sufficient to ensure that a population of potential users will adopt a given tool. Users have additional incentives for behavior stemming from culture, policies, and resources. In LIS, as well as other disciplines, sociotechnical systems theory is one approach for studying the design, use, and impacts of information technologies with an emphasis on economic, cultural, and political contexts.

Sociotechnical theory has been applied in several different domains within LIS research. Levy (2003) demonstrates that libraries (analog or digital) are inherently sociotechnical systems because “ongoing human practices are required to stabilize documents so that they in turn can stabilize us, our practices, and our institutions” (p. 39). Khoo (2006) describes a sociotechnical approach to digital library evaluation, and Khoo and Hall (2010) use sociotechnical theory to explore issues in metadata interoperability.

In digital library research, a sociotechnical view examines the social aspects of technologies, including the complex, dynamic interrelations between users, documents, collections, and institutions (van House, Bishop, & Battenfield, 2003). Marchionini, Plaisant, and Komlodi (2003) describe a multi-case longitudinal analysis to describe a design and evaluation scheme for educational digital libraries based on defined user needs. In one of their case studies, local and state school administrators determined the user needs based on existing educational curricula and assessment. They found that technology and content “are not sufficient to initiate or sustain community in settings where day-to-day practice is strongly determined by personal, social and political constraints” (p. 157). While the system designers employed a user-based approach to creating a digital library, they discovered that there were other sociotechnical factors that inhibited wider use of their system.

Kling, McKim, and King (2003) offer a sociotechnical view of the conditions and activities that support scholarly communication forums. They define a Standard Model which views information technology as tools that are adopted by organizations according to objective measures of technical efficiency. Kling, McKim, and King’s sociotechnical model of scholarly communication views the interaction between people and technology as more integrated.

A study at Dartmouth (Seaman, 2011) used interviews to explore the information needs of humanities faculty in order to inform institutional repository design. Based on his data, Seaman asserts that his participants think in terms of storage problems and the need to market their research, whereas librarians think in terms of scholarly communication, access, and institutional promotion. Incentives were also an issue, and Seaman notes that faculty members

feel that the university does not give credit in the promotion and tenure process for “digital work” in general.

Harley et al. (2010) conducted interviews with 160 researchers, librarians, university administrators, and publishers at over 45 institutions in more than 12 distinct research disciplines to determine faculty values and requirements in the scholarly communication sphere. Fowler (2011) surveyed 600 mathematicians about attitudes and behavior in scholarly communication. The study noted high levels of concern among participants about authors’ rights with respect to journal publishers. Xia (2011) applied an anthropological analysis to a literature review about the open access practices of academic faculty. Xia’s approach examines the difference between the findings of anthropological studies that presented an insider’s view of the data versus an outsider’s view of the data.

Lage, Losoff, and Maness (2011) investigated data curation needs and practices of twenty-six faculty members at the University of Colorado at Boulder and aggregated their findings into eight personas, where each persona represented a different type of “typical” researcher. Some of the factors associated with positive responses to library-assisted data curation include a perceived lack of institutional support for data management, as well as a personal social obligation (for some researchers) to make data available for broader research applications. Factors associated with negative responses to library-assisted data curation include issues surrounding proprietary data, sensitive data (such as human subjects data), system design that is not aligned with faculty needs, or current affiliation with an existing subject-based repository. They also found that researchers in earth sciences had a culture of data sharing in their discipline, whereas researchers in life sciences, physical sciences, and

applied sciences tended to be less inclined to share or curate their data. Some participants cited reasons for not sharing such as a competitive disciplinary culture.

Garcia et al. (2006) conducted a literature review of workplace studies and technological change. Workplace studies use a sociotechnical perspective to examine the social organization of work, and the effects of technologies on work activities.

2.5. Information Behavior

As the previous section demonstrates, sociotechnical approaches have been used in a number of domains within LIS, but it is fairly limited. The literature review below demonstrates that information behavior research regularly produces findings that fall within this dissertation's definition of sociotechnical factors, though in different contexts.

The field of information behavior historically focuses on how people seek, find (sometimes serendipitously), manage, give, and use information, with particular emphasis on academic settings (Case, 2006; Fisher & Julien, 2009; T. D. Wilson, 2000). Human communication behavior, collaboration, and workplace behavior are significant subsets of information behavior research, but there is comparatively less research about how people disseminate findings. Information behavior research has focused less on scholarly communication, yet it does have some insight to bring to sociotechnical factors because of the focus on how people interact with information systems and make informed decisions to accomplish tasks.

Social constructivism is a popular strategy in information behavior since by definition it deals with human subjects. Interviews are the primary data collection technique in information

behavior, though observation, surveys, and participant logs, or journaling, are also common. (Fisher & Julien, 2009).

Ellis, Cox, and Hall (1993) use a grounded theory approach to conduct an analysis of information-seeking patterns of research physicists and research chemists and compared their data with a previous study of information-seeking patterns of social scientists. This study collected data through personal interviews and analyzed data using constant comparative method. Only minor variations were found between the samples and there were no fundamental differences. Ellis, Cox, and Hall find that new technologies—in this case e-journals—are unlikely to replace existing research dissemination practices in the short term, due in part to their lack of status and professional recognition compared to traditional outlets. This is an interesting study to reexamine because in the twenty years (hardly short-term) since those findings were published, e-journals are now accepted and widely used as dissemination platforms for researchers, though many journals offer both electronic and print versions.

Courtright (2007) studies the concept of context in information behavior research and identifies rules, resources, culture, social capital, and social norms, as contextual factors in information practices. She also describes a constructed meaning approach, where contextual factors are examined from the perspective of the actor or participant. Context is important in information behavior research, and this approach provides rich, deep analysis into contextual factors from different perspectives.

Sonnenwald (2007) conducts a review of scientific collaboration and lists a number of motivations for collaboration, including availability of sponsored research funding, increased citation, augmented ability to solve difficult problems, increased reliability, and access to

unique research data. The negative consequences include a diffusion of responsibility. Furthermore, technologies that are incompatible or irrelevant to existing policy and practice will not benefit collaboration. Rogers (1995), as cited by Sonnenwald, asserts that for a technology to be successful, it must be compatible with the values, experiences, and needs of its users, it must be simple to explore and implement, and it should serve a clear purpose. These are relevant findings for exploring faculty attitudes in scholarly communication because of the relationship between collaboration and sharing. Collaborators share resources and responsibilities within a controlled group. Data sharing can be a component of scientific collaboration as noted by Sonnenwald, but data sharing can also occur through public dissemination, as discussed in Section 2.2.3.

Agre (2003) describes conceptual traps in digital library design, including “the trap of inventing a new world” (225). This pitfall arises from technology designers believing that they can substantially change the habits of their users, whereas habits might be shaped by a variety of interdependent factors. Agre’s findings could suggest sociotechnical factors, as well as other types of factors.

In a conceptual analysis of time as a factor in information-seeking behavior, Savolainen (2006) defines three approaches, including time as context, time as a qualifier of information access, and time as an attribute of information seeking. Time as context includes wide variation of circumstances and behavioral patterns in information-seeking or information-gathering situations. These include an individual’s regular media habits (such as how frequently she uses various media) as well as economic factors (such as how far in advance a person plans for the future, which affects the kinds of information she seeks). Time as a qualifier of information

access refers to how quickly certain information is needed and how quickly it can be retrieved. Time as an attribute of information seeking refers to the time frame in which an information-seeking transaction occurs. For example, finding a date for a citation is a relatively short information seeking transaction, whereas designing and conducting a study is a relatively long information-seeking transaction.

Jamali and Nicholas (2006) identified intra-disciplinary differences in information-seeking behavior of doctoral students in physics and astronomy. While this was a preliminary study, it is important to note that doctoral students conducting theoretical research engaged in different behaviors than doctoral students conducting experimental research.

Communication was also an important factor in studies of information-sharing habits of scientific and multi-disciplinary collaboration (Haythornwaite, 2006; Sonnenwald, 2007). Haythornwaite found that researchers shared numerous types of information and learned much from each other about findings, project management, research methods, and technologies.

2.6. Faculty Attitudes

In his master's thesis, James Allen (2005) investigated faculty attitudes towards depositing work in institutional repositories. He surveyed faculty and conducted a collection analysis at twenty-five institutional repositories in the United Kingdom. Humanities faculty concerns about open access repositories included the ease of plagiarism, changes in the nature of their current publishing relationships and practices, and the perceived ephemerality of electronic media.

Judith Palmer (1991) conducted a study using semi-structured interviews to explore certain influences such as personality, discipline, and organizational structure on the information-seeking behavior of a population of scientists. The study demonstrated no difference between the groups for library and document-based activities. Computer use for scientific work, information handling, and degree of enthusiasm for information seeking was different between the groups. The study also determined that discipline, work role, and time spent in subject field were all determining factors in information behavior. This is relevant because studies have found disciplinary differences in faculty attitudes towards data management and curation (Akers & Doty, 2013).

Cronin (2005) considers faculty attitudes toward open access in the context of incentives for professional advancement. He identifies a dichotomy between self-interest and altruism in faculty attitudes and cites the academic reward system as a major contributor to their attitudes about open access and digital communication.

2.7. Chapter Summary

Chapter 2 described a theoretical and empirical basis for initiating a study in sociotechnical factors of faculty attitudes toward scholarly communication. Certain elements in the scholarly communications environment are undergoing some degree of change due to new and disruptive technologies. Several applications of sociotechnical systems theory demonstrated the relationship between technologies and societies, and how it is applied to improve performance. There have been some examples of sociotechnical systems theory applied in LIS research, though information behavior research has revealed sociotechnical

factors in other contexts, such as information seeking and workplace communication. Faculty attitudes have been previously explored in information seeking behavior and library-mediated institutional repositories.

CHAPTER 3

RESEARCH DESIGN

3.1. Overview

This chapter describes the methods used to gather and analyze data. Section 3.2, on the Strategy of Inquiry, describes the theoretical approach. Section 3.3 details my role in the study and my relationship to the participants and setting. Section 3.4 describes the procedures used to gather data. Section 3.5 describes the participants in terms of their research programs, and their levels of awareness of issues in scholarly communication. Section 3.6 describes the methods of data analysis and interpretation. Section 3.7 discusses techniques for validating the data.

3.2. Strategy of Inquiry

This study uses a phenomenological approach and grounded theory in the collection and analysis of qualitative data from semi-structured interviews. It identifies sociotechnical factors that inhibit faculty adoption of institutional repositories at the Institute of Applied Science at the University of North Texas (UNT), and selected colleges at Virginia Tech studying environmental issues. I use phenomenology as a research strategy to understand “human experiences about a phenomenon as described by participants” (Creswell, 2003, 15). In studies that apply a grounded theory approach, the researcher collects data before forming a hypothesis. The data is then coded, categorized, and grouped into themes, which then form the basis of a theory (Creswell, 2003).

Phenomenology is an appropriate strategy for my study because it is not strictly a direct

observation of scholarly communication, data management, or information behavior. I am primarily interested in how researchers understand these concepts, and in their attitudes towards these concepts, with the goal of eventually providing them with better library services. This study specifically analyzes self-reported faculty behaviors regarding scholarly communication.

3.2.1. Development of Phenomenology

The German philosopher Edmund Husserl defined phenomenology in the early twentieth century (1970); Martin Heidegger (1972), Paul Ricœur (1967), and others developed phenomenology further. For Husserl, phenomenology is a philosophical method and a system for descriptive analysis of what we come to know (Velarde-Mayol, 2000), or that which is given. Phenomenology's focus is on what Buckland (1991) defines as information-as-process, as well as information-as-knowledge. In other words, the object of study is one's subjective experience of a thing rather than the thing itself. It is important to note though that phenomenology, as practiced in philosophy, is different from its adaptation in social science and psychology. Philosophical phenomenology is the study of human experience and it addresses basic concepts of the human perception of reality, such as "perception, thought, memory, imagination, emotion ... bodily awareness" etc. (Smith, 2013).

The Austrian social scientist and philosopher, Alfred Schütz, adapted phenomenology for social science (T. D. Wilson, 2002). Whereas Husserl's work was philosophical in nature, and therefore applied to the study of consciousness and the organization of knowledge, Schütz was interested in the social world, and his use of phenomenology was to make social science more

rigorous and valid. Wilson notes that Schütz' rationale begins with the position that in social science, the researcher is dealing with participants who are interpreting the world that the social scientists also wish to interpret. People actively and continuously interpret the world, and researchers observe and analyze that interpretation.

3.2.2. Intersubjectivity in Phenomenology

Phenomenology, as practiced in social science and psychology, is a research strategy that studies human phenomena through the intersubjective lens of human experience. A phenomenological approach assumes humans construct meaning and perceive reality through individual subjective experience. Different communities of practice and cultures live, operate, and construct knowledge within different incommensurable paradigms (Kuhn, 1962).

Intersubjectivity is a critical factor of phenomenology because of the view that humans do not create meaning independently, but through personal information systems which include social networks, selected media, and cultural norms (T. D. Wilson, 2002).

Laura Bohannon's (1966) narrative about sharing Shakespeare's *Hamlet* with the Tiv people of West Africa is a useful illustration of intersubjectivity in phenomenology, and how different cultures construct meaning in order to fit new knowledge into their paradigm. She took a copy of *Hamlet* with her for leisure reading on a trip for field research. At one point, her hosts asked her to tell a story, so she explained *Hamlet* to them in detail. She assumed that there was one universal interpretation of the play, but trouble began immediately when she described the sentries seeing the ghost of the dead king in the first scene. One of the village elders immediately called this impossible and says that it was actually an omen sent by a witch.

Bohannan tried to explain that Horatio was a scholar so he knew how to interpret these things. Unfortunately, Bohannan's closest Tiv translation for scholar was "one who knows things," which also means "witch." The translation becomes increasingly complicated because according to Bohannan's audience, if a man dies, his brother should marry the widow. Throughout Bohannan's narrative, her hosts continued to interrupt with their own interpretations to fit the story into their own cultural context, much to her confusion and annoyance. Eventually the leader of the village told Bohannan that her elders never properly explained the story's meaning to her.

Bohannan attempted to teach the village the "true meaning" of Hamlet, but instead discovered that the people of this culture were able to construct a totally different meaning from the events of the narrative. This anecdote illustrates Kuhn's (1962) assertion that different cultures construct knowledge within different incommensurable paradigms.

Phenomenology is useful for trying to understand a culture or community on its own terms rather than through an imposed external view. This study adopts a phenomenological approach to understand the attitudes of faculty towards scholarly communication on their own terms in order to better understand and provide them with useful services and programs, as discussed in Chapter 1, and called for by Harley et al. (2007) and Davis & Conolly (2007).

Phenomenological research involves prolonged individual contact with a small number of participants to identify common trends. Phenomenology is often characterized by in-depth qualitative data from a very small sample. Its value comes from its emphasis on personal perspective and subjective experience. It can identify deep issues that would be easily overlooked by other methods, such as surveys or log analysis.

In addition, phenomenology is an appropriate strategy for a variety of reasons. It is flexible compared to other approaches, and flexibility is a critical feature since there is no established general tool for understanding user communities in an institutional repository context. Institutional repositories at the two locations in this study are still in the early stages of development, thus the two populations in question are not well understood in terms of their scholarly communication attitudes and practices. Therefore, a study of these groups would be useful to the designers and managers of library repository and scholarly communication services. Because of these factors, phenomenology is a practical exploratory approach for gathering data about how these researchers disseminate scholarly information, and how access to, and management of, scholarly communication tools could be improved.

3.2.3. Limitations

There are limitations to the phenomenological approach. While data collection is richer and more descriptive than some other strategies of inquiry, the nature of phenomenology, which often corresponds with a small sample size, limits the generalizability of the data. Furthermore, the questions asked by the researcher are affected by the current and historically-defined problems and approaches of the researcher's field. Conversely, phenomenology reveals how people ascribe meaning to their activities so that the researcher may understand their attitudes and behavior (T. D. Wilson, 2002).

The use of a phenomenological approach affects the questions asked. Depending on a participant's answer to a question in the interview, the researcher may ask probing follow-up questions that were not standardized throughout the interviews. If the population was better

understood, it might be useful to construct surveys asking specific questions about specific issues. However, it would then be impossible to ask unscripted follow-up questions from individual participants and get additional context in their answers.

A phenomenological approach affects data collection in a number of ways. Data generated from audio-recorded, semi-structured interviews are descriptive rather than numeric. Data collection occurs in a natural setting rather in a controlled research environment and, in this study, timetables were subject to the schedules of the participants. The use of a phenomenological approach affects data analysis because it involves data collected from human sources. Data analysis is subject to nuanced interpretation as a researcher attempts to generalize and code observed behaviors and recorded quotations. Lastly, the phenomenological strategy affects the final narrative by omitting some details that are irrelevant to the research problem. Not all collected data is useful to the stakeholders. A researcher must determine what data trends are relevant based on how clearly they represent participant attitudes and how clearly they relate to the research problem.

3.3. Researcher's Role

Because I collected data from human participants, I needed to obtain permission in advance to conduct the study from the respective Institutional Review Boards at Virginia Tech and UNT. I submitted my list of interview questions to them, and they approved the project (see the IRB Approval Letter in Appendix A).

I am connected to the UNT participants in this study as a former colleague and as a currently-enrolled doctoral candidate, and to the UNT Libraries as a former librarian. I am

connected to the Virginia Tech participants as a librarian and faculty member of that university. I am connected to the VTechWorks repository as the coordinator of that project and I work closely with the software developers. Having these connections to the communities facilitated the process of recruiting participants.

Due to my connection to the two digital library projects, risk of bias on my part was present. I managed the level of bias by avoiding the sources of bias listed by Schensul et al. (1999). The sources I avoided include asking leading questions, failing to follow up on topics introduced by the participant, interrupting or redirecting the participant, failing to recognize participant reactions to my appearance, using nonverbal cues that encourage the participant to give certain answers, or voicing my opinions on any issues.

3.4. Data Collection Procedures

3.4.1. Bounding the Study

This study is limited to tenure track faculty members in the Environmental Studies field. While interviewing researchers in multiple disciplines or including graduate students would produce additional findings, it is useful to limit the research scope in this exploratory phase. Environmental Studies is a suitable field due to its interdisciplinary nature. It includes scientists, humanists, and social scientists. Additionally, Environmental Studies researchers are diverse in methodology, philosophy, tools, culture, and practice. Robust Environmental Studies programs exist at both sites, Virginia Tech and UNT, which means there is a convenient sample. Bounding the study this way narrows the subject of this research to provide useful findings, while the choice of the population increases the diversity of the sample.

3.4.2. Setting

One location for my study was the Denton campus of the University of North Texas. This is a large state university with between 35,000 and 40,000 students seeking degrees. The campus is the site of the Institute of Applied Science, an interdisciplinary research unit focused on human interaction with the environment, and solutions to environmental problems.

The other location was the Blacksburg campus of Virginia Tech. Virginia Tech is a large state university with nearly 30,000 students in degree-seeking programs. Several units at Virginia Tech have faculty in relevant research areas, including the College of Natural Resources and Environment, the College of Liberal Arts and Human Sciences, the College of Agriculture and Life Sciences, and the Department of Civil and Environmental Engineering.

Both universities have faculty and graduate students pursuing education, outreach, and research on environmental problems, sustainability, conservation, and human interaction with the environment and natural resources.

3.4.3. Events and Process

The study gathered data through semi-structured interviews (Cresswell, 2003; Schensul, Schensul, & LeCompte, 1999) conducted between May and September 2012. Although face-to-face interviews were preferred, some interviews took place using a telephone, or a Voice-Over-Internet-Protocol (VOIP) software when faculty members were in other locations, as noted below. I also took descriptive and reflective notes during the interviews.

The eight semi-structured interviews at Virginia Tech were conducted in person. The shortest interview was 18 minutes long and the longest interview was 1 hour and 20 minutes

long. The average interview length was approximately 45 minutes. Seven of the Virginia Tech interviews were recorded with a digital audio recorder and transcribed with a word processor. One participant requested that the interview not be recorded, so I took notes of the participant's responses. Six of the Virginia Tech interviews occurred in the offices of the participants. One interview occurred in a library meeting room, at the request of the participant. One interview occurred at a local coffee shop, at the request of the participant.

Five of the semi-structured interviews at UNT were conducted over the phone. One of the semi-structured interviews at UNT was conducted with Skype (VOIP software), at the request of the participant. These interviews were all recorded through a computer with software called WireTap Studio.

During the interviews, I asked the participants questions about their research and instructional activities, their use of digital libraries and repositories, and their perceived barriers to adopting institutional repositories or to publicly sharing research data. I designed the interview protocol (see Appendix C), and the College Librarian for Agriculture and Life Sciences and the College Librarian for Natural Resources and Environment at Virginia Tech vetted it and helped determine the timing by participating in mock interviews.

3.5. Participants

Section 3.5 and its subsections provide more detail about the participants and their research culture. Section 3.5.2 discusses the participants' research programs and organizes them into different methodologies. Each research methodology has a different way of defining knowledge, asking research questions, and designing research. Some of the participants

engage in multiple types of research. In these cases I coded participants for multiple methodologies. For example, a professor might have one research program that is qualitative, and another research program that uses mixed-methods, so that professor would be coded once in each category. Section 3.5.3 demonstrates how different participants had different levels of engagement with the topics discussed in the interviews.

3.5.1. Participants Background

I interviewed faculty members at Virginia Tech and UNT. The faculty members of the Institute of Applied Sciences at UNT are a target audience of the Scholarly Works Collection in the UNT Digital Library. The faculty members in the College of Natural Resources and Environment, and in the Environmental and Water Resources Engineering research cluster within the Department of Civil and Environmental Engineering at Virginia Tech are target audiences of the VTechWorks Institutional Repository.

Faculty research in Environmental Studies deals with pragmatic problems—access to clean water, clean air, safe food, and stable climates—affecting the livelihoods of people globally. The research that these faculty produce is of great value to everyone, yet the pay walls imposed by commercial journals limit dissemination of findings to people with access to large databases and costly journal subscriptions. The people who are most at risk of the impacts of climate change—the poor—cannot benefit from the findings.

I recruited participants using two methods. Some potential participants were selected by reviewing the research activities listed on their professional or personal web pages linked from their department websites at Virginia Tech or UNT. Other potential participants were

identified through “snowball sampling.” I sent a message to the work email addresses of potential participants to request appointments to discuss their research activities (See the recruitment letter in Appendix B). The email stated that this study was a doctoral dissertation project with the objective of improving the availability of information resources to environmental policy researchers.

Participants were assigned pseudonyms in my recordings, transcripts, and notes to protect their identities. I also recorded certain profile information, such as academic rank and area of research. Identifiable characteristics of the participants, such as name, age, and nationality, are masked in all sections of the report. Given that some participants conduct research in a narrow enough field that they are the only ones engaged in that domain at UNT or Virginia Tech, it was especially important to disassociate fields of research from quotations.

Fourteen people participated in the study. The sample included eight men and six women. Eight participants were from Virginia Tech and six participants were from UNT. At the time of the interviews, four participants were full professors, five participants were associate professors, and five participants were assistant professors. Most participants were American. One participant identified as Canadian. One participant was from a country in Western Europe. All participants appeared to be Caucasian, but I did not ask participants about race or ethnicity. All participants were fluent English speakers.

The participants in the sample conduct research with several different research methodologies, using a variety of methods for data collection and analysis. At least three participants engage in quantitative research. Two participants engage in qualitative research. Eight participants conduct mixed-methods research. Two participants engage in advocacy or

action based research to at least a minor extent, and these two also studied philosophy of science and technology. Three of the participants identified as philosophers, two of which claimed to not collect data. These two participants conduct research through textual analysis.

Most participants had little to no experience with academic digital libraries. With few exceptions, the author had to define digital libraries and institutional repositories, and indicate how these concepts were different from e-books and electronic journal databases. One exception was a participant who was a NSF Fellow several years ago and had some experience with data management policy issues. Another participant was aware of institutional repositories, sees them as useful for faculty research, and is an advocate for them.

The participants benefited from this study by providing input to improve repository services that meet their research needs. The results are reported in the final draft of this dissertation, which is freely available through the websites of the UNT Libraries and Virginia Tech University Libraries. All participants were informed when the study became available.

3.5.2. Research Methodologies

As discussed in Chapter 2, other related studies (Akers and Doty, 2013) highlight the difference between disciplines or departments. This study, by contrast, interviews participants who all fall under one interdisciplinary umbrella and highlights their differences not based on their academic department, but by their approach to scholarship. The participants in the sample engage in five different research methodology categories, and some participants engage in more than one. Most of the fourteen participants conduct mixed methods research involving human dimensions around environmental resources. Two participants conduct

qualitative research. Four participants conduct quantitative experiments and observations. Two participants describe parts of their research program with elements of an action-based or participatory approach. There were also three humanists who mostly work with texts.

Among the four participants engaged in quantitative research, there are two major approaches. One participant uses airborne and space-borne remote sensors and ecosystem process models to study forest ecosystem structure and function and validates the models with collected field data. Another participant's work focuses on improving environmental services, such as storm water mitigation and air quality provided by urban forests, and also focuses on the health of soils in urban forests. The third participant in this category also researches urban forestry with a focus on evaluating arboriculture practices, such as soil amendments, pruning, and transplanting. This participant also conducts forest inventory and assessment with questions about what trees are in urban forests, where they are, and how they function. The fourth participant in this category analyzes land cover as part of a collaborative team looking at environmental changes that correlate with climate change at timberline and high elevations in a national park set in Latin America.

There were four participants with five qualitative research programs. One program examines human behavior in problems and conflicts in natural resource management and organizational psychology and sociology in natural resource management agencies. Another program studies climate change impacts on subsistence farmers in developing countries. A third program, under development, would analyze the effects of shale gas drilling on adjacent populations. A fourth program studies expressions of environmental guilt. A fifth program studies social impact in the oil industry of a West African nation.

Four participants use six mixed-methods research programs. Two participants engage in multiple areas of research. In order to maintain confidentiality I will only report the research projects without stating that two or more specifically relate to a particular participant. One research program studies water quality at the tap by blending qualitative and quantitative methods to detect metals, pharmaceuticals, and pesticides in drinking water. A second research program is in educational institutional development to help institutions build capacity to train and educate people in natural resource management. A third research program examines student retention in undergraduate STEM programs and also explores the trouble for minority and women graduates in finding STEM careers. A fourth research program is in the marketing and utilization of non-wood forest products—such as mushrooms and fibers—for dietary supplements and medicines in some cases, and rubber and other industrial applications in other cases. A fifth research program deals with conservation education and its quantitative effects on behavior change in municipal utility usage. A sixth program focuses on sustainability indicators that measure progress towards policies or goals in communities and institutions. These indicators were for technical measures, such as pollution levels, as well as ethical, social, and value based measures like quality of life, or social justice.

For the activist and participatory research, one participant engages with a community on local politics and ordinances in energy development. Another participant set up a company to assist a village in bringing its goods to market, with all profits going back to the community through a communal bank account. Both of these participants noted that they kept this area of work separate from the more “formal” research that they publish.

Three participants engage in critical reading and thinking, and in producing texts. One of these participants self-identifies as a humanist and creative writer, another identifies as an ethicist, and the third identifies as a philosopher. The humanist has an environmental studies background but conducts research in the philosophy of science and technology with a focus in energy and policy. The ethicist studies sustainability factors. The philosopher studies water issues from an interdisciplinary humanities perspective that includes social, political, cultural, philosophical, ethical, and theoretical questions.

As stated in the beginning of this section, this sample represents an interdisciplinary community and therefore highlights differences in faculty attitudes between research methods rather than disciplines, as highlighted in previous studies (Akers and Doty, 2013; Sonnenwald, 2007).

3.5.3. Levels of Engagement

Participants demonstrated different levels of previous engagement with issues discussed in the interviews. I group participants into three categories based on their levels of engagement with the topic. In the first category, eight of the participants probably had not discussed these concepts and technologies in depth before. In the second category, four participants had some previous level of engagement with the issues discussed. In the final category, the remaining two participants had a thorough understanding of scholarly communication issues and had been involved in setting university policy.

One or two participants were previously unaware of, and remained unconcerned by, the scholarly communication issues raised in the interview questions. These participants thought

that institutional repositories sounded useful, but they accepted current norms for disseminating research. Another group of participants were aware of some scholarly communication issues, such as the rising cost of journals. They were in favor of broader dissemination without pay walls but were unfamiliar with institutional repositories. The participants in the third set were very much aware of scholarly communication issues and saw them as a cause for serious concern. These participants were strong advocates for open access, or had strong opinions about how repositories should be used.

Most participants thought that a library service to collect and make all the research of the university publicly available sounded useful. One associate professor at Virginia Tech said, “That would be awesome! Do we have that?” An assistant professor at UNT thought it sounded helpful from a standpoint of personal organization, saying “It may be preferable, because for example I just had somebody request one of my papers, and honestly, I don't have a copy of it. Part of me would really be in favor of a repository that would have everything that I have published available to me for that kind of situation.” This participant was generally enthusiastic. Under the conditions that he maintains copyright and intellectual property, he said, “you name it.” He was also ok with giving a permanent license to his institution to host his work.

A few participants were highly informed about some of the issues discussed in the interview, but were mostly unaware of others. Regarding the concept of open access, many participants were unaware of institutional repositories and understood open access only in terms of hybrid journals where the author, or the author's institution or funding agency pays a fee to make the article freely accessible without a subscription. One full professor at Virginia Tech

indicated some familiarity with open access and claimed to have been approached by open access publishers, but had not been involved yet. On the other hand, this participant had been a fellow at NSF several years previously and was familiar with some of the data management policy issues. Another full professor was somewhat aware of data management requirements in grants, but was apparently not affected by federal data sharing policies or journal data sharing policies. Much of this respondent's data collection came in surveys and interviews that involved human subjects, so publication outlets might not be as focused on data sharing. This respondent claimed not to have been required to share data before, stating: "I remember writing a proposal for NSF or something like that and they said the data had to be made available, but I don't think that is very common. It's never been a concern for me because no one has asked me to do it. I imagine that there are some funders out there that would ask to do it. NSF is one of them, but it's never been an issue." This is a significant issue because this interview occurred in mid-2012.

Approximately six months after this interview, the White House Office of Science and Technology Policy released the Policy Memorandum (described in Section 2.2.2) requiring that all federal agencies with more than \$100 million in research expenditures devise plans to make the findings of funded research more accessible. More federal agencies have data sharing requirements now and the data collection and analysis period co-occurred with a period of rapid policy change for data sharing in the United States.

Most participants exhibited sophistication about search interfaces, understood differences between full-text search and Boolean search, and how to use each effectively. During the interview, one participant accurately, if partially, described the concept of search-engine optimization when discussing dissemination strategies. This participant however, like nine out of

fourteen of the participants in this study, did not seem to be aware of library-managed repositories for dissemination of scholarly communication or management of research data.

There were varying levels of engagement with and awareness of open access in scholarly journals. This is important to document because open access can be achieved through the publication process, or it can be achieved through archiving a work online with or without peer review or an editorial process. The following examples illustrate how “open access” was a confusing term for many of the participants, and how some participants conflated “open source” or “online” with open access. In one example, a full professor at UNT asked, “Do you mean ‘open access’ in the sense that people can go in and revise what I write, or... in the sense of anybody, anywhere can get them?” In a different example, when I asked a full professor at Virginia Tech if any of his articles were in open access journals the participant asked if open access meant “online.” When I gave a concise definition of open access as “anyone can read them without a subscription,” the participant replied, “besides going to the library? I don't know. I guess, it would be online but I don't know.” I asked if the participant was familiar with open access, and the reply was, “not so much. Most of the journals that I publish with are published through the old traditional publishing houses ... I am pretty suspect of open access. I tell my students, I can put anything I want online. It's not [necessarily] going to be truthful.” Several professors at UNT were unaware that UNT has an Open Access Policy. On the other hand, one of the participants at UNT had been involved in drafting the university’s open access policy, demonstrating that in some cases participants had low levels of engagement with open access, but in other cases participants had high levels of engagement.

Other participants were aware of open access journals, or that one can pay a fee to certain hybrid journals to make their article open access. They seemed to sense that there was some kind of gray area with what they could legally email to colleagues or put on their website. Many participants seemed to think of open access as something that did not require authentication to access, but in many cases they were less confident. For example, more than one participant said they could access a particular work, but that they were “on a university computer,” with the implication that they might have been authenticated through their IP address. One participant opted to make their doctoral dissertation open access, but had not published any other research through open access journals. One humanist in environmental philosophy said that most of the discipline’s journals were open access, unlike in life sciences where there is a lot more money involved. This suggests that participants have varying levels of engagement with the subject of open access journals.

In addition to different levels of awareness about open access channels in published scholarly communication, there were also different levels of awareness about library services that support research and scholarly communication. The following excerpts indicate a difference in understanding between what academic libraries recognize as electronic resources with licensed content and what they recognize as digital libraries and institutional repositories. When asked about digital libraries, an associate professor at Virginia Tech asked, “would that be like when I use Web of Science to get articles, would that count as digital libraries or is that something different?” A full professor at UNT said, “I go in and get stuff all the time, but I don’t really pay attention to what it is called or where it is coming from.” The confusion about the types of library service is probably because the distinction between two access mechanisms for retrieving

relevant scholarly articles is not important to end users in this population.

Regarding institutional repositories, three categories of responses emerged in the interview data. One category represents participants who had never heard of institutional repositories. The second category includes participants who had heard of institutional repositories, but had not used them. The third category includes participants who were already familiar with institutional repositories, and had strongly held opinions about them.

The first category is the largest and includes a diverse set of reactions, though all participants expressed at least general interest. One associate professor at Virginia Tech said, “I don't have any experience with that.” When I asked if it was something the participant wanted to explore, the participant replied, “I'm not sure I even understand exactly what you mean.” After I defined the concept and explained the service the participant said, “I've never given that a whole lot of thought. I have never been approached about doing something like that so it sounds like something that would be helpful but not something I've given a lot of thought to and not something that has been on my radar.” When another participant at Virginia Tech heard of the institutional repository, the participant said “it would be cool if more people knew about it.” This indicates that it sounded useful, but that previous outreach had not been successful. One assistant professor had lots of questions about UNT's Scholarly Works: “How do you go to the Scholarly Works? Is it in yet another database that you need to look through? Is it a database that is now in UNT's database aggregator? Is it something I would direct someone interested in my work to visit? Or is it something that I as a researcher would visit to find more information about a given topic?” These questions seem to indicate a lack of previous experience with repositories, but also a strong interest in learning more about them and how a researcher could

use them to his or her advantage. These examples demonstrate that in spite of previous unfamiliarity with institutional repositories, some participants wanted to learn more and discover how to potentially use and benefit from the service. Other participants thought repositories sounded interesting and useful, but their curiosity about the service was less enthusiastic.

In the second category, an assistant professor at UNT initially sounded unfamiliar with the Scholarly Works Collection there, but remembered it once I started to explain it. The participant said, “oh yeah, I am familiar with that. I gave a talk last year and they were telling me about it and asking me if I would submit my talk and ... I think I just got caught up in other things and didn't respond to their email.” This example indicates that in this participant’s case, outreach had been somewhat successful, but there had not been sufficient follow through.

The third category was the smallest but also the most vocal. This participant had few questions about the service, but had lots of opinions about its scope and execution. A full professor at Virginia Tech had very strong opinions about what should and should not go into a repository due to the huge institutional commitment to maintaining stable access to a resource, demonstrating that he was aware of institutional repositories and had thought about them enough to articulate opinions.

The division of these three categories of participant engagement with institutional repositories, and the respective sizes of each category, suggest that a lack of awareness of institutional repositories is the norm, even at the institution with an institution with a campus-wide Open Access Policy.

3.6. Data Analysis and Interpretation

Using a grounded theory approach, I analyzed data through an iterative process of identifying and labeling common themes in the interviews and observations. My first steps were to file and transcribe each recorded interview. Once the transcriptions were complete, I sent them to the participants for review. Then I used the qualitative research software NVivo to help me annotate transcriptions with terms representing concepts that I discovered while reading the transcriptions. Using word search features, I identified concept terms and applied them throughout the transcribed interviews wherever I believed those concepts occurred. NVivo allowed me to organize occurrences of the concept terms into themes. These dominant themes that emerged from the interviews are the main findings of this study.

3.7. Validation

Validation of findings in this study occurs through several strategies—triangulation, member-checking, and clarification of bias.

Triangulation occurs by collecting data from two sites in order to justify the themes defined in data analysis and interpretation. I engaged in member-checking by sharing the themes with the interview participants to determine whether they felt the themes accurately represent what they expressed in the interviews. I clarified the bias that I brought to the study by explaining my relationship to the participants and to the institutions that were the sites of the research.

3.8. Chapter Summary

This chapter described the strategy of inquiry, the limitations of the study, the setting, the participants, and the methods of data collection, analysis, and validation.

CHAPTER 4

FINDINGS

4.1. Overview

This chapter presents the findings from both research questions, as well as a summary of those findings. The data is grouped into sections based on the research questions. Section 4.2 presents findings related to the research question, “How do faculty want to disseminate their research?” Section 4.3 presents findings from the research question, “What attitudes do faculty have towards data sharing?” Section 4.4 summarizes the findings. All participant names in this section are pseudonyms to protect their identity and privacy.

Not all interview questions yielded useful data for this study. The interview questions about information seeking and about instruction produced useful data about how participants seek information and how their research practices differ from their instructional practices, but they did not produce relevant data for understanding sociotechnical factors in faculty attitudes towards scholarly communication.

The observation component of the on-site interviews was useful for establishing the level of systems and retrieval knowledge for each participant, but the observation data contained evidence more appropriate for a narrow usability study rather than the topic of this dissertation. The observations came from a video recorded screen capture of the eight Virginia Tech participants using the UNT Digital Library’s Environmental Policy Collection, along with an audio recording of what the participant said while interacting with the collection. Participants spoke aloud while exploring the interface and the contents of the collection. A few participants who had research interests in environmental policy were interested in the content. Participants

who were not interested in policy and did not use gray literature had little use for the collection. Participants reporting lack of a specific need for environmental policy materials said they used peer reviewed literature and book chapters for course materials and for their own research. Participants also asked questions about the interface and search and browse features. In summary, the observation component contained feedback about the interface and the content of the collection. The context was in regards to information seeking for teaching and research and was not of value for this study.

4.2. Ways Faculty Want to Disseminate Their Research

This section discusses the findings related to the first research question in section 1.6, which asks, “How do faculty want to disseminate their research?” All participants stated that they disseminate their findings through traditional methods such as peer-reviewed journal articles, conferences, books, and technical reports. When asked about how research findings are disseminated, Professor Cooper, who conducts quantitative research, summarized the choices as, “You present, or you write it and publish it. That's pretty much it.” Dr. Blackburn, an associate professor in the humanities, similarly reported traditional venues, such as books and journals, for disseminating research, adding, “I don't blog or use the Internet too much.”

Dr. Stanley, an associate professor, reported a few different ways of disseminating information to different audiences, and alluded to the importance of open access, and paying a fee to make the most important published work available to the broadest audience. Dr. Stanley said, “I have published in a lot of different journals. _____ is one of the open access journals I had been into. I also do reports, just for sponsors, prior to publishing, and just give them

straight up non-theoretical practical stuff.” Dr. Johnson, an associate professor of biology, discussed disseminating research through reports in addition to the journal articles, manuscripts, and conferences. Some reports are specifically required by federal grants and funding agencies at state and municipal levels.

Professor Packard also expressed that there were different channels for dissemination for different audiences with diverse needs. The intended and perceived impact on each audience is different, and the method of disseminating information for each audience is different. Professor Packard spoke of different target audiences saying, “There are the communities of practice that I want to be able to exchange knowledge with, so I have to disseminate the information that way. In the communities of practice, we're really scrambling to try and provide information in such ways that help people. They're pretty self-determining, but they need information, and if they have that information it can have an impact right away in that same growing season. Communicating back is a complicated thing at different levels in different ways.” Professor Packard works with the global poor, so communities of practice include rural agricultural communities in developing countries. These groups have different information needs than academics or professionals in global development organizations, thus the participant uses different techniques for disseminating findings to each community.

At least five participants mentioned non-traditional or less formal research dissemination through blogs and interviews with news media. For example, Dr. Cole, an assistant professor in the humanities, does some participatory or action based research and has a variety of research channels, including a blog that is “mostly just reposting and very little commentary” while trying to maintain a neutral academic tone. The intent for Dr. Cole is to

“engage with the local community on the problems that they face. Dr. Cole said, “I am not really producing any data or texts in these conversations. A lot of it is just verbal and network building stuff.” Dr. Cole also writes for more popular publications, local papers, and national online magazines, in addition to the traditional academic journals and book chapters.

Professor Bryson is a civil engineer who engages in a variety of formal and informal scholarly communication. Professor Bryson reports disseminating findings through peer-reviewed journals, conferences, and a trade journal, as well as through expert interviews in public interest pieces in news media, and a non-peer-reviewed, but widely read news publication produced by a professional society.

Dr. Hayward, an assistant professor with a humanities background, described dissemination through encyclopedias saying, “This is important work because environmental studies is such an interdisciplinary field. Many times a really good encyclopedia article can be quite helpful for a researcher or a student who wants to step into another discipline a little bit more to get the overall scope of the field. It was pretty exciting for me to be able to do that work to chart what the field means. So I, as a humanist, was really shaping the way that these technical volumes were being put together, which for me is an important way to disseminate information.”

In some cases, dissemination channels clearly correlate with the culture of a department. For example, Professor Horne at Virginia Tech said, “We need to be peer-reviewed because that is one of the criteria. I don't think we get anything for a book. We get dinged for not having peer-reviewed articles. It's a pretty rigorous group.”

In addition to these traditional and non-traditional channels, participants also described what they want to accomplish with their research outputs. Table 1 lists the novel ways participants wanted to change their scholarly communication, including increased transparency, increased impact, and increased intellectual property rights, and it is discussed in the sections that follow.

Table 4.1

Novel Ways Faculty Want to Disseminate Their Research

Participant	Rank	Methods	Novelty
Cooper	Full	Quantitative	1,2,3
Stanley	Associate	Qualitative	2
Palmer	Assistant	Quantitative	(None identified)
Rosenfield	Associate	Quantitative	(None identified)
Desmond	Assistant	Mixed	2
Truman	Associate	Mixed	(None identified)
Bryson	Full	Mixed	(None identified)
Horne	Full	Mixed	3
Martell	Assistant	Mixed	(None identified)
Cole	Assistant	Humanist	2,3
Hayward	Assistant	Mixed	3
Johnson	Associate	Mixed	1,2
Packard	Full	Qualitative	2
Blackburn	Associate	Humanist	2

Note. 1=Transparency; 2=Impact; 3=Intellectual Property

4.2.1. Increased Transparency

As indicated in Table 4.1, Professor Cooper indicated a strong interest in supplementing traditional publications through the availability of supporting research data and algorithms

used to generate the findings, and thereby improving transparency of the scientific process. Professor Cooper also expressed frustration over the peer-review processes of some journals, particularly about how some editors use their position to try to increase the impact factor of the journal. He said, “We routinely get requests from editors to add such-and-such, and such-and-such citations, but it's almost always for their journal, which is of course another little gaming factor. To get extremely egregious, we had an editor ask for his own stuff to be cited, just to increase his own impact factor, which is lame beyond measure. We're like, ‘oh, come on!’”

4.2.2. Increased Impact

While nearly every participant in the study expressed the desire to disseminate research in a way that maximizes impact, participants define impact in differing ways including bibliometric impact, readership, and social change. Bibliometric impact refers to quantitative indicators such as the Institute for Scientific Information (ISI) impact factor. While there were other reasons participants gave for publishing in ISI indexed journals, this particular theme refers to an interest in a high impact score using bibliometric methods. “Readership” in this section refers to instances where participants indicated that their goal was to have their ideas and findings read by as many people as possible. “Social change” in this section refers to instances when participants wanted their scholarship to change the values and priorities of a society.

Five participants indicated a strong interest in increasing access and availability of their published research to broader readership. Dr. Stanley, for example, wanted broader

readership, and simply liked the idea of having published works be more accessible and read by a larger audience. Dr. Stanley said, “to answer your question from a faculty perspective, we want our stuff to be read ... to have an impact in whatever way I can, whether I get cited or not. So first and foremost, we want people to be able to find it and get access to the full text as easily as possible; I know from experience that a lot of folks, if they are looking for something but can't click on it...will find something else that talks about the same thing and use that instead. So anything that enhances the clickability off of Google Scholar or any search engine would be awesome.” To enhance the accessibility of previous research, Dr. Stanley, and other participants, had paid a publisher’s fee to make an article open access in the past.

Reasons for paying to make an article open access varied. While other faculty claim to make articles open access when they have the money, Dr. Stanley waited to make a specific publication—the most important one out of a research project—publicly available by paying for it to be open access. Dr. Stanley said, “we've actually been working on one project for about six years, and there have been a lot of intermediate publications, but we finally got to a point where we have the whiz-bang-here-is-the-golden-nugget-at-the-end-of-six-years, and that one we are paying to make open access. It's a three thousand dollar fee to do it.” As Dr. Stanley said, there were several prior articles, but the final output of the research will be open access because of its significance.

Dr. Blackburn said that within environmental philosophy, “Most journals are open access. I think our journals are usually not that expensive; it's not like a medical journal, which would cost a lot of money to get a subscription, so they prevent open access. I think in

humanities our journals are more available; there's not that much money involved so it's easier to make publications available for open access, but I honestly don't know for sure.”

Professor Cooper had also paid open access fees and said, “We try to do open access if we have money, and this is always the issue because somebody's got to pay. Fifteen hundred to three thousand dollars is about what I have seen journals charge for open access fees, so we have done that several times recently, just because of my own wish to try to retain this kind of thing. You just have to pay to play I guess. If the libraries aren't paying, then basically the person who is doing the scholarship has to pay somehow.” Professor Cooper was referring to programs at some academic libraries that set up a subvention fund to help researchers pay to publish their articles in open access journals. This demonstrates an interest among participants in open access as a way to increase impact through broader readership, while also acknowledging the reality that this method of increasing access costs money, which is a barrier to dissemination.

Professor Cooper is willing to pay open access fees for hybrid journals, and does not correlate a journal's reputation with its access model. In spite of that, Professor Cooper demonstrates concern about journal reputation: “the last thing you want to be seen doing is publishing in a, quote 'easy journal,' or 'lame journal' with no standards or whatever. But it has [caught on], so we've had a few things out there in the last year.” This demonstrates that while Professor Cooper wants to make research as widely available as possible, it is important that there is a rigorous review. Professor Cooper's use of the phrase “the last thing you want to be seen doing...” suggests that the rigor is not only important for the advancement of science, but also for one's reputation as a researcher.

Several participants discussed a few other strategies for increasing impact through readership. Dr. Blackburn said “I don't always go to the same journal; I try to get a broader audience.” Professor Cooper discussed several strategies to ensure greater dissemination and impact. He noted that as long as people have access to typical citation databases, they probably will not have a problem finding his research. One of his strategies is to “make a concerted effort to be in ISI journals. Partly because they are discoverable, but also because of their vetting of the peer-review process, and independence of the editorial board ... there tends to be a little better quality.” This example demonstrates that Professor Cooper makes a conscious choice in the channel for disseminating research based on discoverability, transparency, and editorial rigor. Professor Cooper's other strategy for increasing impact was to ensure that peers in one's field are aware of one's new research developments. Professor Cooper said, “The biggest thing I sense that increases the citability of an article is having influential people in that area know about it. For example we had one published in [*journal title*] recently, which used to be independent, but they let Springer manage it. It's quite a good journal. They asked me at the time of publication for twenty email addresses of people that I thought would be interested in seeing it. In retrospect it was really quite clever, because in the end it increases citability. So I think that impact improves primarily through vehicles like that, which is the first time I'd ever seen anyone doing that, or informal networks, or things like conferences.” These participants' ways of achieving impact demonstrate attention to strategy, and especially in the case of the latter, concern about the quality and rigor of the peer review and editorial process.

Professor Cooper indicated specific frustration targeted at Elsevier as a synecdoche for commercial publishing in general. He said, “I really get peeved by the “Elseviers” of the world. One of the top-rated journals in my field is an Elsevier journal, and I have always dissuaded students from publishing there, partly because [I’ve worked in a developing country without access to an academic library] and I remember how difficult it was for people to have access to information in that context. I’m like, ‘well guys, if you do that, you’re just basically saying that ‘them that has, gets’ for this stuff. They’ll say, ‘well, its highly rated, I’ll get a better job’ or whatever, and I tell them, ‘maybe, but I can tell you that my most cited stuff outside of the big stuff, like *Science*, or whatever is going to be not at all related to the impact factor of the journal,’ which from what I have seen, they’re gaming anyway.”

Only a few participants brought up social impact, but it was incredibly important to them. Professor Packard described several layers of social impact: “Academics can be really helpful in guiding policy, but if you get that out to the community of philosophers it could change the world, but you’re not going to know that for a hundred years or so. It’s much more of a long term process to change the ideologies, particularly in the global north, which is the primary academic context that I am trained to participate in. Changing the ways that people think, that’s a long slow process. It’s not changing someone’s ethical view, it’s changing the very foundational conceptual structures that determine how people understand and set up worldviews. That is long-term change. In different places I expect different levels of change and different levels of response and different kinds of response from the information dissemination. The philosophical context is the longest term, and the policy context is—you don’t change policy overnight—but I hope that the work in the policy context could make a

difference in five or ten years.” This description shows that, quite differently from an impact factor, there is an interest in having an impact on policy on one level, having an impact on the ethical view of a reader at a second level, and changing basic conceptual structures and metacognition on a third level. Professor Packard, who already has tenure, is enthusiastic not just about being read, but about changing the world. Citations and promotion are not even mentioned.

Similarly, Dr. Blackburn exhibited a very high level of altruism regarding impact and research dissemination, to the extent that she reported a preference for social impact over how colleagues view the work, and even reported not caring too much about the potential of being plagiarized, so long as the research could have a strong impact on society. About open access, Dr. Blackburn said, “I am all in favor of completely open access even if it makes it harder to get points in my faculty evaluation; I would sacrifice those points for the accessibility. For me, it's much more important to have an impact than to be acknowledged by some university committee. It's more about the message than about my name. Having said that, there are some concepts that I am working on, and I hope I will get some credit for having come up with those ideas. Ultimately, the most important thing for me is that things percolate and that in my field there will be rising awareness about [these environmental issues], and if it's my work or someone else's work, or someone stealing from my work, I ultimately don't care too much. It's not about my name necessarily or my institution's name but way more important, the themes that get attention.”

Examples in this section illustrate different ways that participants define impact in scholarly communication. Categories of impact include bibliometric impact, readership, and

societal change. While trying to increase exposure through broader readership, three tenured participants state that they have paid to publish research in open access journals. One of these three tenured participants waited until she had tenure before she felt confident enough to publish in an open access journal. One tenured participant was skeptical about open access due to its newness. Several assistant professors were interested in open access, but expressed concern about how a promotion and tenure committee would evaluate open access publications. At least two tenured participants had published in open access journals that do not charge an author fee. This demonstrates an interest in open access publications as a means to increase readership as a form of impact, but some participants remain wary about the reputation and perceived rigor of open access journals. Those who remain concerned specifically cited the promotion and tenure process as the source for their concern. While at least three participants were interested in societal change as a form of impact, it was unclear how this would be evaluated.

4.2.3. Increased Intellectual Property Rights

The desire to increase access to one's research was tied in some ways to the desire to have greater rights over the work. Issues that emerge in the interviews include difficulty in retaining rights to use one's own work, apprehension about how work can be personally or informally shared with colleagues, and in some cases, authors knowingly violating copyright law. Participants want to maintain ownership over their work and decide what they can do with it after it has been published. For example, participants want to feel they can post their

articles on their personal webpage, or email them to colleagues without violating intellectual property laws.

At least five of the participants were concerned about how permissions with various publishers would be negotiated, since journals have different policies. Three participants expressed some frustration with the practice of some publishers owning copyright over published faculty research. The participants found it difficult to obtain the rights to reuse their own work. Dr. Cole had to buy a copy of the publisher's version (with correct pagination) of an article that he wrote, while Professor Horne had previously published five dissertation chapters as separate peer-reviewed articles, and then had to get permission from four or five journal editors to publish them as a book. Permission was granted, but Professor Horne sounded slightly bemused by the process, saying, "it was kind of interesting, but it was another step." Dr. Hayward was unfamiliar with, but intrigued by the idea of repository services in part because "some of the publishing companies for different journals are much more permissive and some have been much less permissive... I had to fight to get a PDF copy of my own article for my own use because I was putting together documents for my last big review, but they didn't even want to give me a copy for myself. The library didn't have that journal so eventually I ordered it through interlibrary loan, and it just seemed absurd that I had to order it through interlibrary loan when it was my own article. Of course I have the text but I wanted at least one copy with the right pagination for personal use."

A few participants said that they email their articles to anyone who asks, and claim that this is a standard dissemination practice in the academy. Some also claimed to post publisher versions of articles on their own websites. Professor Cooper also said it was "probably quite

dreadfully against the copyright in about 1/3 of the cases,” and “I'm just waiting for someone to sue me.” This demonstrates that participants feel strongly about their need to freely disseminate their work and share their results. This remark could be interpreted either that Professor Cooper is consciously breaking the law and does not care, or perhaps he is just being flippant. Additionally or alternatively, Professor Cooper might be saying that this practice is so common in higher education that a publisher is unlikely to enforce their copyright. Professor Cooper, and most other participants, indicated awareness of publisher policies and intellectual property laws, and were concerned about how to submit material to a repository without openly violating copyright. As illustrated in the last example, however, some participants seem to dislike certain copyright laws that inhibit formal scholarly communication, and in some cases knowingly work around them.

Professor Cooper stated a preference for journals owned and operated by professional societies. He said, “Most of the society journals let you [deposit into a repository]. You just can't have their name on the masthead. Some of the non-society journals are less good. If it's IEEE or EGU, which is where I tend to try to publish because I have always been a strong believer in professional societies controlling publishing, they are really good about it.”

Professor Cooper clearly feels that academic and professional societies, rather than commercial publishing houses, should control the dissemination of research, and the reason seems at least partially tied to the extent of control an author maintains over published research.

4.2.4. Status Quo

I did not detect any desire for novel ways of disseminating research in four participants.

Dr. Rosenfield is an associate professor who did not feel constrained by authors' rights determined by publishers. Dr. Palmer and Dr. Martell are assistant professors who focus more on maximizing the number of their publications. Dr. Johnson does seem to desire novel ways of disseminating research, but feels constrained by the promotion and tenure system.

Professor Horne, is a full professor who did not indicate a desire for change in dissemination methods, and did not seem necessarily attached to the status quo either. He spoke skeptically about Open Access as a means of increasing readership and impact saying, "It's just so new, and we're so tradition-bound to these other journals that we pay subscriptions for."

Dr. Rosenfield described publishers in positive terms. He spoke favorably of the submission process and turn-around time in particular at two Elsevier journals and noted that author's rights have never felt constraining. Dr. Rosenfield was unaware of the institutional repository, and said that it "sounds like something that would be helpful but not something I've given a lot of thought to and not something that has been on my radar." He stated, "I've never really felt constrained in publishing in either one of those journals. Honestly, I've been very satisfied with the publication process in both instances. The manuscript submission process that they use for both is pretty painless, and they're pretty responsive in terms of the turn-around time of the reviewers, as well as getting them back to you in a comprehensible way, and then through the revision and resubmission process. I don't ever recall having been constrained [in what the publishers allow]. In fact with one they had a companion publication, which is kind of a professional journal, and they actually encourage the authors of their scientific articles to take those and bring them down a level, to make them a little less 'science'

and a little more practical, to disseminate the findings to a more general audience. In that regard, I think that that particular publication is facilitating access to a broader audience.”

Dr. Johnson spoke favorably of Open Access, but acknowledged that among promotion and tenure committees, there is a stigma attached to paying to submit an article. She said, “They don't see it counting the same, even though the peer review process is the same. They think that there is a bias in the financial side. So, in trying to get tenure, that was not a risk that I was willing to take. I have received tenure now and I am more likely to feel comfortable. I'm a firm believer in open access and I'm one of the few lucky professors where I do have access to indirect monies where I can pay the fees. Some of my colleagues still believe that open access is not peer-reviewed, and that there is a financial bias. Until my colleagues who have been around a long time get more learned about open access journals and how they work, and understand that they are not vanity journals, this is not likely to change.” This demonstrates that while certain participants want to explore open access as a means of disseminating research, they are concerned about negative consequences in the promotion and tenure process. Therefore Dr. Johnson had to work against her own interests in some cases in order to satisfy the promotion and tenure committee. This was particularly interesting because Dr. Johnson's institution, UNT has a university-wide open access policy. The culture of the discipline or department may be a factor for some researchers in choosing research dissemination venues, and it may outweigh university policy in incentivizing behavior.

Examples in this section demonstrate that a large minority of participants does not seek alternative ways of disseminating findings and that three reasons emerged. One participant does not feel constrained by current scholarly communication norms. Two participants focus

achieving impact and professional advancement by maximizing publications rather than. One participant felt constrained by the ignorance of her peers but was not confident about working outside of her departmental and college standards until she had tenure.

4.3. Faculty Attitudes Towards Data Sharing

This section discusses the findings from the second research question in Section 1.6, which asks, “What attitudes do faculty have towards data sharing?” As noted in section 3.5, there was some variance in a number of participant traits, such as academic rank, methodology, and disciplinary culture. These traits in turn correlate with variance in attitude. Pre-tenure faculty for example tend to be more guarded and conservative than senior faculty because they are seemingly more risk averse, particularly in areas represented in their dossier for promotion and tenure.

Many participants in the study appear to have conflicting views about sharing their data. Nearly all participants see value in sharing research data, but there were a few prevalent concerns about providing free global access to one’s own research data. Table 2 demonstrates how different factors make participants concerned about sharing research data, especially in an open repository or similar system where participants do not have control over the data. In some cases, the participants were still interested in the benefits of data sharing, but they had concerns about how it would be done and the conditions under which they would be willing to share data.

Different themes emerge in the interviews. Section 4.3.1 discusses perceived benefits of data sharing including increased transparency of the scientific process and public access to

publicly funded research. Emerging themes that indicate concern about data sharing include research ethics and the protection of human subjects (4.3.2), worries about research data being taken out of context (4.3.3), concern among senior faculty for the ability of junior researchers to publish findings (4.3.4), and concern among junior faculty about their own ability to publish findings (4.3.5).

Some emerging themes correlate more strongly with distinct categories of participants. For example, participants in the non-tenured faculty category expressed concern about other researchers poaching their findings. While tenured faculty in this sample expressed concern about researchers poaching the data of their graduate students and post-doctoral associates, they did not express similar concern about their own ability to publish unique findings from shared research data. Therefore, in this example, the emerging theme of concern about one's own data being poached correlates with junior researchers rather than tenured faculty.

Context is another emerging theme that strongly correlates with one category of participant. In this study, "context" refers to the information necessary to ensure that research data is understood. This area was primarily a concern for tenured researchers. Some of the participants with tenure worry about their research being misunderstood by people without training, and feel that the job of a researcher is to provide quality data and explain its meaning. If research data is publicly shared without an article attached to it to explain nuance, participants see a risk of students, media, and policy makers misinterpreting the data, and lay people in general not being well served. As illustrated in the subsections below, there are issues with the politicization of science, as well as issues with definitions of terms as they are used between different disciplines, and as they are used in popular discourse.

Table 4.2

Attitudes of Faculty Towards Data Sharing

Participant	Rank	Methods	Themes
Cooper	Full	Quantitative	2,3
Stanley	Associate	Qualitative	1,2
Palmer	Assistant	Quantitative	3,4
Rosenfield	Associate	Quantitative	No objections identified
Desmond	Assistant	Mixed	What if I did something wrong?
Truman	Associate	Mixed	4
Bryson	Full	Mixed	4
Horne	Full	Mixed	3
Martell	Assistant	Mixed	3
Cole	Assistant	Humanist	No traditional data
Hayward	Assistant	Mixed	3
Johnson	Associate	Mixed	2,3,4
Packard	Full	Qualitative	1,4
Blackburn	Associate	Humanist	No traditional data

Note. 1=Ethics; 2=Transparency; 3=Career Protection; 4=Context

4.3.1. Transparency

Transparency of scientific research is another significant aspect of disciplinary culture. While few participants really focused on this issue, it was a key theme in the interview with Professor Cooper, who conducts strictly quantitative scientific research. He feels very strongly that research data and software that support published research should be available to the public in order to demonstrate the validity and reproducibility of the findings. He additionally thinks that data should be available because in many cases research is funded by taxpayers either through federal grants, or through state support of public universities. The underlying

belief is that taxpayers should not have to pay for the benefits of publicly funded research through taxes, and again through publishers.

In addition to the benefits that publicly accessible research data can give to taxpayers, the same participant also saw benefits for the scientific method. Professor Cooper said, “in some instances we're proposing a new way to do something and that these are the scientific results from that new way, and other people need to be able to corroborate that in their own systems. That's just how science works. I've had colleagues whose models won't be publicly released, and you're wondering what kind of sleight of hand really went into the publication. I mean, honestly. But if anybody can right click and download, then even if they never actually use it, there is this feeling of transparency. That's really important to us. It's not just the data, it's the models themselves, and other analytical techniques.” This quote demonstrates how Professor Cooper feels that public access to scientific data is valuable to the public and to science.

On the other hand, greater transparency takes extra effort on the part of the researcher. Dr. Stanley saw the benefits of data sharing, but explained the personal cost with making data accessible to peers. He said, “I think it would be great if someone wanted to replicate my study, but again, it would make work for me. I would have to go in and make that data a little more—I know best practice is that I should keep my data all better than I do, and my students keep their data better because I teach them to, but I do as I do rather than what I teach.” Dr. Stanley happens to conduct a lot of qualitative research, and the data is therefore more nuanced and harder to organize using standardized tools. Due to Dr. Stanley's work with

human subjects, there are additional costs, and potentially fewer benefits to the accessibility of shareable data, which will be discussed in greater detail in Section 4.4.2.

4.3.2. Research Ethics and Protecting Human Subjects

Several faculty in the sample conduct human subjects research. Faculty in this group have to protect the anonymity of research participants. During interviews conducted for this study, faculty voiced concerns about the extra effort involved in desensitizing the data enough to protect human subjects, while still maintaining its utility for reviewers and other researchers who might reuse the data. Researchers who want, are encouraged, or are required to share research data, have to strike a delicate balance between sanitizing the data to protect human subjects and providing enough context for the data to be useful to other researchers. For example, Dr. Stanley said, “a lot of my data has personally identifiable information so I am scared to death of archiving some of my data without totally cleaning it first to a point where it may not even be useful anymore.”

Another constraint on data sharing for faculty who conduct human subjects research include concerns about IRB compliance and procedures. Dr. Stanley said, “for most projects, when the project is done I don't reuse the data because of IRB. You have to reopen the file and go through that process again.” This is a major difference that emerges between methodologies. Quantitative research places a premium on transparency in the scientific method, whereas qualitative research emphasizes the need to protect human subjects. Human subjects researchers have to deal with extra layers of bureaucracy and institutional oversight that supports human subjects protection, making data sharing more difficult and risky.

Participants express different concerns about data sharing that may correlate with differences in their methodological culture. Dr. Stanley said, “My doctoral research was about people doing illegal things and that's super sensitive. There is organizational stuff that I do. I look inside [government agencies] at their internal functioning, and people will say stuff that you definitely do not want their name associated with, just because they're talking about their supervisor.” His methods, and his data sources, ethically prevent him from being able to share all of his research data. Furthermore the effort involved in cleaning the data to make it shareable, as this participant stated in the quote near the beginning of this section, comes with a personal cost in labor, and in addition may degrade the usefulness of the data to other researchers.

Similarly, Professor Packard is interested in data sharing but raises concerns about informing human subjects in advance about what will happen with the interviews. She discussed how her research is not politically or socially risky for the participants, but the rules are still the same. She said, “my concern is always if someone in the community is made vulnerable to violence in any way as a consequence of the research. I am not doing [risky] research, and I have the confidentiality protections in place that I think are ok. When you're talking about if there is enough rain and how the crops are going to come in, it's not entangled in conflict the same way and it doesn't pose the same threat.” Initially Professor Packard said that while sharing recordings would not be appropriate, sharing a sanitized interview transcript that did not reveal identities would be acceptable. After further consideration though, Professor Packard's opinion changed. She said, “I would have to go back to the IRB board first and get an ok to make the transcript itself public rather than just the conclusions I have drawn out of it, and I think I would

need to have a consent form before I did the interview. People need to know that is going to happen with it. I would make sure that all the ethics stuff was in place and the person being interviewed knew there was a possibility that I would do that. And if they agreed to it then I would be ok. I don't know... I guess if you put it out there people would use it, so maybe I should have more concern. If the IRB doesn't care... I don't know. I've never really thought about it, and maybe I need to think about it a little more deeply.” Professor Packard voiced concern about human subjects protection and about the time involved with IRB procedures. She was clearly against retroactively releasing interview data from former studies without informed consent. Her responses in the interview suggest that the nature of her research questions, method of inquiry, data sources, and the research ethics associated with her studies all contribute to her concerns about data sharing. Her hesitation and uncertainty as she explores her own opinion on this issue further suggest that this is a new concept for her to consider, and she feels that her opinion could still change.

4.3.3. Context

Context is another theme that emerged in the interviews. As mentioned in Section 4.3, “context” refers to the information necessary to ensure that research data is understood. Participants ultimately feel that research data needs context to be understood, and providing that context requires effort. Some researchers in the sample expressed the concern that if their research data is made public, then someone else could take certain points of data out of context and produce alternative findings for political reasons. For example, Professor Bryson expressed that when data is presented in an article, the author can describe nuance in the data,

but if the raw data was used in political journalism, the producer might gloss over the details, or cherry pick some data point to advance a certain bias. Participants also feel that that without context, making such data public would be useless or even irresponsible. Ultimately, many participants are concerned that their data could be taken out of context and/or politicized.

As in Section 4.3.2, where participants raise concerns about the cost of protecting human subjects in data publishing, there is also cost in providing context for data when making it publicly available. For example, Dr. Palmer stated that for “most of it, in order to make it available for free, you would have to associate it with a paper that you published, because context is everything, and then in order to make it intelligible to a larger world, it would require a lot of work. It wouldn't be something where I could just post my file. It would take weeks to label it so someone could understand it. It's not that I would be adverse to it but I don't know that it would be worth the effort.” Dr. Palmer demonstrates the idea that data sharing is not as simple as just uploading a set of files to a server; rather, there is a great deal of labor involved.

Professor Bryson demonstrated knowledge of federal agency guidelines on data management and sharing, and agreed with the concept of using data and making it accessible for other users, but has two main concerns: data can be very specialized, and data can be easily misinterpreted. Professor Bryson said, “Once you give it away, you're no longer in control. You could get credit for it in Fox news. Or it could appear in *Science*, but you don't get credit.” By getting credit in Fox News, Professor Bryson was referring to the possibility of data being misinterpreted because of a lack of context, or misused for political purposes. The phrase about research appearing in *Science* but not getting credit meant that if she shared her own

data, then another professional researcher could publish from it in a highly regarded journal, but that there might not be proper attribution to the researcher who collected the data.

On the specialized nature of data, Professor Bryson said that if one is just publishing the data without the context of an article or a book, there are no chapters of introduction describing nuances in the data. Context and nuance are especially difficult to convey in interdisciplinary research where different fields have separate definitions for the same term. For example, according to Professor Bryson, “intake” in water research means something different from “intake” in nutrition research. Professor Bryson is also concerned about environmental data being misinterpreted and gave an example of statistical variation in measurement. She noted that a researcher cannot detect 0% arsenic in water, but can only detect a certain percentage below the detection limit. Once a toxin is below a certain level of quantification, the researcher gives a star that says, “below certain limit.” The public, however, wants to hear “no arsenic in water” instead of “insignificant level of arsenic in water.” Professor Bryson highlights the importance of context in presenting data and said, “As a researcher, my role is to provide quality data, and provide enough background to help people understand it.” This last quote illustrates that she feels that her concerns about context and nuance in data sharing are tied to her identity as a researcher, which suggests that her attitude is an aspect of methodological culture.

Dr. Palmer described an additional concern about context and discussed problems with data sharing in a nascent field where data standards are still developing. She said, “One problem is that the data in this field is not very consistent. Two years ago we set up a committee to sort of standardize data collection so we could support data sharing.” She went

on to explain how different researchers in a new field might collect data in slightly different ways, leading to non-standard measurements, which is problematic in science because research needs to be replicable and generalizable. She said, “There is no consistency in how things are measured, so it's difficult to share data sets. In [the traditional field] that hasn't been the case because they have very standard methods, but what happens in [our field] is that a lot of those methods don't make any sense. It's a more complicated environment, so you end up with these problems.” Dr. Palmer was describing how her new field adopts the standards from a larger, more general field. In the specialized nature of her field, however, the standards are inappropriate for the setting. As a result, her field needs new standards. In the absence of those standards, data sharing is somewhat irrelevant because data reuse requires some basis for comparison.

Other participants acknowledged the difficulty of context in data, but had an attitude different from others in the sample. For example, Professor Packard is a philosopher who conducts anthropological research. She said, “People have an incredible capacity to understand things in different ways than you intend. In some ways, their interpretation is just as valid. You know my interpretation is just as privileged as the next person's, so not really privileged at all. I don't have a great fear of misinterpretation, but I would be much more wary of that with different kinds of research.” This indicates that for Professor Packard, different interpretations of data are a possibility, but as a philosopher who conducts research with anthropological methods, there is subjectivity in interpreting all aspects of human experience. Her attitude towards the idea that someone might see things in her data differently than she did aligns with her methods of inquiry.

Dr. Hayward illustrated the difference between organizing data for herself and organizing data for others. She said, "First of all, to make it public, I would want to make sure it was really clean and tidy and the categories were well explained. I'm not saying I wasn't careful about it, because I was. But I wasn't necessarily writing formal definitions for what went in each category for example." She might have had definitions in mind for each of her categories, but since she wasn't planning on disseminating the data itself, she felt that it didn't need to be as clear to others in the dataset since it would be clarified in the publication. Going back and properly labeling the research data would involve effort to ensure it was useful and clear to others.

Dr. Johnson at UNT discussed the importance of context and gave two examples of pitfalls in data interpretation. She said, "If data doesn't have a key, then that's going in the wrong direction. You know, like a key to a map tells you about its size, dimensions, orientation, and that kind of thing. If there is not really good detail with the data, it could be misconstrued. A '1' could be one dollar, it could be one percent, it could be a holding point, it could be translated into an 'A.' That kind of thing scares me for a person who doesn't have that capability in statistics. My own students struggle with this, and they are graduate students. That's one of the reasons why we see so many articles produced by newspapers and magazines, and things are misconstrued." The first pitfall she illustrates here is that data can be easily misconstrued even if it is well organized and clear to the researcher. In order to share it with others, it needs a key to clarify its meaning. She also said, "The other day on the news I was hearing a reporter say that Dallas was 60% minority. Well, 60% is a majority, so they were using "minority" as a classification of people rather than as mathematical representation. That kind of

thing happens even with well-educated people, so it scares me to just have this random data and not know.” The second pitfall illustrated in this quote is that sometimes, statistical terms, such as ‘minority’ are used colloquially, and that can confuse a reader. She continued, “I think that anybody who would read the paper should see the data. I love all of that. I want to see other peoples’ data. I don’t want to just see the little piece that you have given me. I want to see it all. I completely agree with that. But, Sally down the street just having the raw data, I don’t see that as being a good idea. It would need context. So it makes a researcher a little nervous I think.” This last part of Dr. Johnson’s quote articulates how she feels strongly that data should be available, but it needs to be properly explained in its context.

This concern about providing enough context and organization in research so that other people can see the data and come to the same conclusions relates to a similar point that several other participants made. Research data has traditionally been organized and labeled in an environment where only the researcher will see it or use it. A common reason some participants did not want to share their data was the amount of work required to organize the data to make it useful to someone else. Several participants demonstrated concern about the burden of preparing research data for broader use and public dissemination. Some participants acknowledge high social value in sharing research data, but see high personal cost in terms of time and effort to prepare research data for sharing. Those participants that see low social value in sharing research data share the belief that the time and effort required to clean data is too great compared to the perceived benefit. The faculty members who see little value in sharing their data believe that their data is too narrow to be useful in a different study. Many participants feel that there would be a lot of work involved to make the data suitable and

accessible for sharing. Some assumed that they would not get much credit for that work, and could not readily imagine an alternative use for their research data, so they did not feel that it was worth their time. Examples in this section illustrated different issues for faculty who are concerned about context in data sharing. Those issues include differences in definitions of technical terms between specialized fields; organization and labeling of research data; explanation of what the data represents; and ability for readers to look at data and come to different conclusions due to either difference in perspective, lack of training, or misuse of terminology.

4.3.4. Career Protection

Many participants were reluctant to share research data because of their perceived need to maintain exclusive access for themselves or for their research group. Two main categories arose under the career protection theme. In the first main group, many junior faculty members and a few tenured participants expressed concern about their ability to publish the findings from their data before someone else could if it was freely available online. Most of these participants thought that publicly funded data should be freely accessible, but they would not want to share their unfunded research data because of the possibility that someone else could “run with it” before they could fully capitalize on it in their own publications. The other main category in this theme was tenured faculty who want to protect the research programs of pre-tenure faculty, post-doctoral research associates, and graduate students. Tenured faculty who advise, supervise, or otherwise collaborate with students and post-doctoral research associates worry about research data from their projects being broadly

available before their advisees can complete their own publications. The concern is that outside researchers who find and use shared research data may be able to publish on the findings first, hurting the junior researchers' ability to publish and build their own careers.

Dr. Palmer, for example, was ok with data sharing in principle if a few conditions were met. She said, "If I was working on a project and I hadn't published it, and you came up and said, 'Can I have your data? I am working on a paper and I want to use it.' I would probably be like, 'well not just yet, because if you publish the data then I can't publish it.' It might be one thing if I could still publish it. They would have to be using it for a totally different purpose I guess ... It's not that I mind sharing it per se, but you just can't publish things twice, and I need to publish my papers, or I will get fired. So [a significant difference between my work and their work] would be the condition [for sharing] I guess. The basic thing for me is that I need to maximize my number of publications. I have to do that. In order to do that, I need to publish papers, and if someone else were to publish a paper using data that I collected before I published it, then essentially I wouldn't be getting my job done." The important factor for Dr. Palmer is the need to be able to continue to produce publications, and she believes that sharing data prematurely could ultimately inhibit job performance. As an assistant professor, she is primarily evaluated by the amount of research she produces, and anything that inhibits her ability to publish will have a negative impact on her performance.

Dr. Hayward is a humanist who primarily analyzes texts, but has conducted several mixed methods studies where she had amassed data sets. She expressed conflicting thoughts on data sharing saying, "I haven't really thought about this a lot, and I don't know what I think about me thinking this, but I feel a little bit cautious about sharing that information because, on

the one hand I would want people to be able to build upon that work. On the other hand, this sounds kind of... I don't know. There's not a formal recognition process for recognizing that type of work, and my research assistant and I spent months putting that research together, and I don't necessarily like the way this makes me sound, but in the academic world, at least in the humanities fields, you get rewarded for your end publication, not for the data you put together. So if I did all of that work and then put it out there for anybody to use, on one hand that would be awesome and advance the field. On the other hand there is no formal recognition for that. Now, maybe I should get recognized for that and that would be great, but on the other hand, it seems like I would be giving up something pretty easily. And that I think is a real challenge of the academic system that we have today. Why am I thinking that way? Well, I'm a tenure-track faculty member. I have to think about protecting my research, as well as collaborating, and I have to figure out how it makes sense to use my time."

One of the most interesting things about Dr. Hayward's answer is the conflict about the perceived social benefits and the personal costs in sharing research data. This answer simultaneously demonstrates some altruism as well as some ethical egoism. Altruism is apparent in statements like, "I would want people to be able to build upon that work" and "that would be awesome and advance the field." Ethical egoism is meanwhile reflected in statements like, "there is no formal recognition for that," "I have to think about protecting my research," and, "you get rewarded for your end publication." This answer is also interesting because of the participant's self-awareness and discomfort with the cognitive dissonance in the answer, evident in phrases such as, "I don't know what I think about me thinking this," "this sounds kind of...I don't know," and, "I don't necessarily like the way this makes me sound." It is also interesting to note how Dr.

Hayward says, “as a researcher that primarily works with text, I haven't really thought about this a lot,” which suggests that she does not think much about data sharing due to having a different relationship with data than a scientist has. It is clear that she was thinking deeply during the interview, and even articulating her metacognition. It is possible that her view on data sharing could evolve in a number of directions.

Professor Cooper explained why it is important for a student’s career to wait an appropriate period of time before sharing data. He used my own research to illustrate to me the importance of unique data. He said, “In the end, we like to make data available once the period of initial extraction and publication is done. You’re working on something right now. You don't want someone else to suck all the use out of these interviews before you’re done. Usually what we try to do, particularly with students, is protect that space. Once we feel like the dissertation, and the papers out of the dissertation are complete, then we're perfectly happy to make it public and let other people get what they can out of it. People can spend literally years on their research, and you're spending years on something here too. You want to make sure that you're protecting that person's scholarship, if it needs protecting. There are some instances where they're using data anybody can download and it doesn't make any difference anyway—the novelty is in their approach, or something else. But in other instances, we've cobbled together all sorts of interesting collaborations where the datasets are unique, and it's our access to those data that makes the research unique. In those cases, we don't want to open it up the moment we get it. That's poor management.” In this example, and in others, Professor Cooper demonstrates a strong interest in the transparency of scientific research, but in this example he also gives a caveat. When research data is unique, then it is extremely

valuable to the person who collected it. Since Professor Cooper advises and supervises graduate students and other researchers, he wants to protect their ability to publish their findings. If the data is available to everyone, then someone else might publish the main findings before the student or post-doc. Once that happens, journal editors will be less interested in the publication by the person who actually collected the data and put it online.

While few tenured participants displayed any concern about sharing their own research data, Dr. Stanley, an associate professor at Virginia Tech said, “If I am still working with that dataset and publishing from it, I would be pretty protective of it and make sure that whomever requests it is using it for a different purpose. I have actually shared data once with a colleague at another university. I had no personally identifiable information in it, and the reason I was comfortable is that they were just using it to draw a sample. Great, take my data. But if it was somebody who wanted to ask the same types of questions I was asking—unless it was years after and I wasn’t going to re-analyze—I wouldn’t share it.” Dr. Stanley does a lot of qualitative research. Since Dr. Stanley and Professor Cooper are both in the same college, their differences in attitude towards data sharing could stem from differences in methodological cultures.

Dr. Palmer had also shared data in the past. She described an experiment that involved a lot of a trial and error work in developing the design, and then shared data with another researcher who was doing a similar experiment. Apart from that, Dr. Palmer claimed to have not been asked for data in the past, but had a graduate school classmate who, on several occasions, solicited other researchers for their data based on questions he developed while reading their publications. Dr. Palmer said, “You develop this data set; you want to get your publications out of it before you just put it up on the web for anyone to use. But there’s always

more data in some data sets that can be extracted if someone wants to use it. I knew a grad student when I was getting my masters degree who would read a paper and say 'I wonder why they didn't look at this.' Then he would call up the person and ask them to send over their data and he would re-analyze it and write a whole new paper about some aspect the original investigators ignored. He must have published something like ten papers while he was a Ph.D. student doing that. I think you can do that in some cases, and I am sure that the people who gave him the data didn't mind." It is unclear from this anecdote how the graduate student cited the researchers that provided him with data, but since he did it many times he may have had a procedure. The important thing about this quote is that it seems Dr. Palmer admired her classmate's productivity and did not seem to think he was taking unfair advantage of anyone. It appears that she assumes he had permission to use the research data in further publication on different research questions.

Data citation was an interesting topic for several participants. Dr. Johnson said she would give research data to anyone who asked without the expectation of being a co-author, but with the expectation that the recipient would attribute her role in collecting the data. Professor Cooper thinks that the standardization of data attribution could help make data sharing more attractive to researchers. He said, "One of the things you would want, in the end, is to have some recognition for data development, so why not have a section on our vita and annual reports or whatever, where you put your publications in, and you could have a whole separate section for published data sets. You could see the same sort of things: what kind of citation rate is it having? It seems like, for the science side, it's just another piece where people haven't been paying close attention. And then there is going to be data authorship and that's going to

be another issue—once you can cite something, it's going to be like, 'who's is the first author? Who is the second author?'" This last point demonstrates scenario building around a situation where researchers get more recognition for building data sets than they do now. Professor Cooper envisions the standardization of customs and procedures where data is an established formal research output, and credit is divided appropriately among collaborators according to their role in the design, collection, and organization of the research data.

Dr. Martell expressed the strongest concern among participants within the sample about the need to protect research data. Dr. Martell was aware of federal research data policies and indicated that he would share his data if it were a condition of funding. Otherwise, he felt that the person who collects the data should always be credited as a co-author in any publication that comes from the dataset, and that was his own stipulation for sharing research data. He stated that for him, merely being cited in an article wouldn't be a significant enough benefit to incentivize data sharing, and he listed a number of concerns about being a junior faculty member and having research data that was difficult to collect. He said, "I think my hesitation is as a relatively young academic. I think there's a chance for someone else to take that data and run with it when I am in a position where I can't completely capitalize or recognize what I have." He went on to explain how his research data could become a subset of someone else's larger study, and then once that other person releases their findings, his own work would become unpublishable. Dr. Martell further said that this is common in academics. He said, "There's a lot of stealing of information and taking of information, so I am really reluctant to share my ideas, my questions, and even my data with people. It's a reluctance that's a strategic one because it's an unsavory world out there. Some people can't come up with good ideas, but

man they can finish that stuff and run with it. They'll take your data and say, 'oh, I would never do that' and then they go ahead and use it, and then you did all that work for nothing. That's my impression." Co-authorship however was an acceptable incentive to share research data. He said, "If I make an agreement to be an author on anything that they publish with my data—if that's the case and I am getting credit for it then that's a quite different story than being like, 'Hey, I'm working on this project' or, 'Hey, I'd love to see that data that you've got.' If they said that, there would be this conversation: 'well if I share this data with you, then any publications you get out of that, I'm going to be on the publication right?' That happens quite often: you'll see people who have collected a soil core or an ice core thirty or forty years ago. If anyone uses and analyzes the data that comes from that core, they always have to cite the person who collected the sample, because they put in the work to collect that data. It's their data, so their name is associated with it. Until I see the returns on the time, the money, and the effort that I put into collecting the data, analyzing the data, and getting it published into peer-reviewed journals, I am really reluctant to share that with other people, just because it is mine. It's what I have built up; it's my cache. That's pretty standard I think throughout the academy." Dr. Martell clearly feels that there is a strong risk of plagiarism if research data is made available for anyone to use.

One distinct difference between Dr. Martell and Professor Cooper and Dr. Johnson, is that Dr. Martell sees himself as a young academic, whereas Professor Cooper and Dr. Johnson already have tenure and they are more established in their careers. The other key difference is Dr. Martell is a mixed methods researcher, whereas Professor Cooper and Dr. Johnson are both trained as scientists, even though Dr. Johnson currently conducts educational research. As

scientists, their methodological culture may correlate with different attitudes towards data sharing.

Dr. Hayward was also concerned about the added work if the research data was dynamic, longitudinal, or required updating. If the data was private and the researcher moved on to other interests, it would not be a problem, but if the data was live and the researcher moved on to other interests, the researcher would feel obligated to maintain something he or she was no longer interested in. Dr. Hayward said, “part of me thinks it would be interesting if I could make that data available, and somebody else who was interested in similar questions could have access to it. On the other hand, I am a little bit hesitant about that. I did this project but I don't really want to update it every six months with the new research. If I established a website and said, ‘here is all this research,’ I would feel like I would need to keep it updated in a way. I feel that I have moved on in my research and I don't want to have the job of compiling this data forever.”

This section demonstrates faculty concerns about how sharing research data will affect opportunities to publish findings. Participants with a background in the sciences are more open to sharing research data than participants who employ qualitative research methods typical of social sciences. Furthermore, junior faculty are more guarded about sharing research data than tenured faculty, though some tenured faculty also reported the need to protect access to data for graduate students until they could complete their publications.

4.4. Chapter Summary

In addressing the first research question, I found that participants are interested in

changing some aspects of how they disseminate their research. Some participants want greater transparency in the peer review process as well as in the communication of scientific findings through public access to the research data, models, and algorithms that support published findings. Participants additionally want to increase the impact of their work, but they want to accomplish this in a variety of ways including traditional bibliographic and scientometric measures, broader readership, or social change. To expand readership, some participants engage in different forms of dissemination for different audiences with diverse needs. Some participants are satisfied with the current scholarly communication system. Others are dissatisfied, but have not been confident that they were in a position to affect change. Finally, dissemination choices may be linked to the culture of a discipline or department.

In addressing the second research question, I found that participants have a number of ideas and concerns about sharing research data. As found in other studies, (Nelson, 2009) while most faculty value the idea of public access to research data, they are reluctant to share their own research data. This study produces novel findings about faculty reluctance towards data sharing. As discussed above, reasons include the need to protect human subjects, the difficulty in the bureaucratic challenge to re-open former studies through Institutional Review Board and contact former research participants, the difficulty in providing enough context to help users understand research data, the effort to organize and document research data to make it useful, and the need to have the data linked to peer-reviewed publications that explain it. The final reason found in this study is the concern among faculty that if their research data is available online, they will not be able to publish findings from their data before another

researcher uses it in a publication, affecting their own ability to publish research and advance their careers.

CHAPTER 5

ANALYSIS AND SUMMARY

5.1. Overview

This chapter analyzes the broader theories of technology and culture introduced in Chapter 2 to elucidate and contextualize the findings from Chapter 4. Section 5.2 presents the findings from the first research question, “How do faculty want to disseminate their research?” Section 5.3 presents findings from the second research question, “What attitudes do faculty have towards data sharing?” Section 5.4 presents recommendations for repository programs. Based on the findings from Chapter 4, section 5.5 explores how this research advances the study of library and information science (LIS) and discusses pathways for future research.

In addressing the research questions, most participants express an interest in changing the ways that they disseminate research, but they are reluctant to do things differently because they face personal risk and personal cost. Faculty are evaluated based on specific criteria, and they need to engage in research and dissemination that fit within those criteria. If they engage in alternative types of research and dissemination, they perceive the risk of a negative evaluation in the promotion and tenure process. Since participants are evaluated on specific criteria, they see personal cost in the time it takes to share research data, no benefit in the promotion and tenure process and sometimes limited benefit in their field.

5.2. How Faculty Want to Disseminate Research

In addressing the first research question, “How do faculty want to disseminate their research?” I found that ten participants are interested in changing some aspects of how they

disseminate their formal findings. In discussing aspects of research dissemination that they would like to change, emerging themes include increased editorial transparency, increased impact, and increased intellectual property rights over their published work. Within the increased impact theme, there were several categories of impact, including traditional bibliometrics, broader readership, and societal change. Four participants did not specify anything they wanted to change. Within this group of four, one participant was content with current options and methods for disseminating research, and three participants focused on their need to produce more articles.

All participants produce and disseminate traditional forms of scholarship such as peer reviewed papers and monographs. Some participants discussed other forms of research outputs, such as through reports, blogs, trade journals, and popular media. The intended and perceived impact on each audience is different, and the method of disseminating information for each audience is different. Some faculty conduct multiple types of research that result in different types of impact on their field and on the world. In spite of the high levels of engagement that participants have with their alternative research, they feel that the promotion and tenure process does not assess the impact that results from these projects.

In some cases, dissemination channels correlate with disciplinary culture. Some trends are obvious: faculty in STEM fields need to produce peer-reviewed articles, and faculty in the humanities need to produce academic books. One participant mentioned the practice of peer-reviewers or journal editors engaging in coercive citations, and he was the only one who listed editorial transparency as a concern in research dissemination. He was also the only participant engaged in purely quantitative research who was also a full professor. In contrast, the three

participants who identify as philosophers, plus the participant who identifies as a creative writer were all pushing the boundaries of what constitutes research and what constitutes impact. Disciplinary difference is one of the two most significant findings that emerged in the interviews.

Faculty attitudes in scholarly communication also correlate with faculty rank. Participants who focus primarily on increasing output in traditional channels are all pre-tenure and several of these participants specifically stated their need to increase their peer-reviewed publications. Some tenured participants also described strategic scholarly communication choices they made while they were assistant professors in order to minimize risk in the promotion and tenure process. This was the other most significant finding that emerged in the interviews.

5.3. Faculty Attitudes Towards Data Sharing

In addressing the second research question, “What attitudes do faculty have towards data sharing?” I found that participants have a number of ideas and concerns about sharing their research data. As found in other studies, most faculty value the idea of public access to research data but they are reluctant to share their own research data. My study produces novel findings about why faculty reluctant to share data. As discussed in section 4.3, concerns about sharing research data include the need to protect human subjects, the difficulty in the bureaucratic challenge to re-open former studies through IRB and contact former research participants, the difficulty in providing enough context to help users understand research data, and the effort to organize and document research data to make it useful. Additionally,

participants are concerned that if they do not control access to their data, they will not be able to publish findings from their data before another researcher uses it first, affecting their own ability to publish research and advance their careers. Unsurprisingly, research ethics and the protection of human subjects was only a concern among participants who actually engage in human subjects research.

Context in research data was a bigger concern for scientists and social scientists than for humanists, and it was more commonly cited by tenured participants than non-tenured participants. Among participants who discussed context in data sharing, three were tenured and one was not tenured. Since this concern mostly occurred with participants who had tenure, it may correlate with a certain amount of experience conducting and disseminating research and seeing popular media coverage of research.

5.4. Recommendations for Repository Services

Based on the findings and analysis, the recommendations of this study are that repository managers and developers understand faculty needs for advancement as well as offer mediated deposit and copyright assistance. Furthermore, many participants had never heard of the repository, including at UNT, which has a mandatory institutional Open Access policy. This demonstrates a failure to date in outreach and communication. New strategies are required for increasing visibility of services. Embedding repository services in faculty workflow, such as through annual reporting, and in the promotion and tenure review process could improve faculty awareness and use of repository services. It is productive to tailor programs and services to address problems that researchers actually have.

The findings in this study support Harley et al.'s (2010) findings that motives such as advancing one's career and advancing one's field are dominant incentives for faculty when deciding how to disseminate research. Faculty researchers select channels to disseminate their most significant research based on the factors of journal prestige, speed to publication (how long it takes from submission to publication), and visibility (Fowler, 2011; Harley, 2013). Some participants in my study liked open access publications for their most significant research in order to increase the visibility, but accessibility was not the only criteria or even the most important criteria.

Institutional open access policies theoretically make useful knowledge more available to society, but they are controversial in many institutions and they often cost significant political capital. As demonstrated in Chapter 4, people have different ideas of what open access means and have concerns about their ability to publish and get credit for research that is seen as rigorous. Ideally, open access policies are paired with other efforts, such as investing in personnel and systems to gather all of the citable publications generated by an institution, and then ingesting article versions that publishers allow. This method for populating a repository also falls within a range of traditional roles for librarians: identifying, finding, collecting, and managing documents.

Most participants indicated that their interest in submitting their work to an institutional repository would depend on how much of their time it would consume compared to perceived benefit. Some participants indicated that they would prefer if someone else would do the work for them, or that they would need a certain amount of assistance. Dr. Martell described an experience of being approached by a repository librarian. He said, "I think I just

got caught up in other things and didn't respond to their email. I felt like I was waiting for someone to say 'this is what you can do, this is how I can help you do this, and this is what it will entail,' and I never quite got the guidelines, or support, or confidence that it warranted."

Dr. Martell was not disinclined to participate, but he might have been approached at the wrong time, and without sufficient follow-up. Workload was another common problem. Dr. Cole stated, "A lot of this would depend on how much work would be involved for me. Not to sound selfish about it, but if I am super busy, the work would have to be done by somebody else."

These quotes show how researchers might be inclined to use a repository, but get distracted by the daily pressures of their job or feel that they do not have enough information to act.

5.5. Recommendations for Future Research

This study raises questions for future research in a number of areas, both for practitioners, and for LIS theorists. More research is needed to provide a better understanding of faculty attitudes. It would be useful to conduct institutional or inter-institutional surveys of researchers to determine faculty attitudes towards scholarly communication to examine differences between academic ranks, including graduate students, postdoctoral research associates, and adjuncts, and also to determine differences between academic disciplines and research methodologies, and differences between liberal arts colleges and research universities, or other types of institutions.

There is already a significant amount of data and commentary in the study of data management and repository services in academic libraries. However, there has been less examination into root causes and trends behind the findings of previous studies and the use of

sociotechnical theory and phenomenology has been largely absent. This dissertation's use of phenomenology and sociotechnical theory demonstrates the applicability of these theories. While phenomenology has been defined and sporadically applied in LIS research, it is not a widely used approach in this discipline (T. D. Wilson, 2002). Further work is necessary to demonstrate its wider utility.

Howcroft, Mitev, and Wilson (2004) assert that sociotechnical systems theory is a natural fit for the study of LIS, yet it is not widely adopted in the study of scholarly communication. The Rob Kling Center for Social Informatics at Indiana University is one exception, and the Special Interest Group for Social Informatics of the Association for Information Science and Technology has held annual symposia since 2004. The use of sociotechnical systems theory has been expanding in LIS in the last ten years, but there have been few applications of this theory in the study of scholarly communication.

This dissertation builds on the work of Harley, and Foster and Gibbons to help develop a deeper understanding of the needs of faculty and improve institutional repository programs in large academic libraries. Section 5.2 illustrates how the findings discussed in chapter 4 inform the future of institutional repositories. As discussed in chapters 2 and 3, this study employs methods of inquiry that have been underutilized in the study of institutional repositories. The findings of my study suggest that my approach was productive and potentially useful to the advancement of LIS.

This study provides a greater understanding of faculty attitudes toward institutional repositories. The findings of the study can assist academic libraries in understanding how to create institutional repositories to support data managements and scholarly communication in

their parent organizations. It further provides an opportunity to learn more about the attitudes of faculty toward data sharing and how they want to disseminate their research. Institutional repositories continue to hold great promise for the information professions. It is essential to understand issues that impact the long-term success of such repositories.

APPENDIX A
IRB APPROVAL LETTER



A green light to greatness.™

Office of the Vice President of Research and Economic Development
OFFICE OF RESEARCH SERVICES

January 2, 2013

Supervising Investigator: Dr. Brian O'Connor
Student Investigator: Nathan Hall
Department of Library and Information Sciences
University of North Texas

RE: Human Subjects Application No. 11583

Dear Dr. O'Connor:

The UNT Institutional Review Board has reviewed and approved the extension you requested to your project titled "Information Seeking Behavior of Environmental Studies Researchers and Instructors." Your extension period is for one year, **January 10, 2013 through January 9, 2014. Federal policy 45 CFR 46.109(e) stipulates that IRB approval is for one year only.**

It is your responsibility according to U.S. Department of Health and Human Services regulations to submit annual and terminal progress reports to the IRB for this project. The IRB must also review this project prior to any modifications. **If continuing review is not granted before January 9, 2014, IRB approval of this research study expires on that date.**

Please contact Shelia Bourns, Research Compliance Analyst, or Boyd Herndon, Director of Research Compliance at extension 3940, if you wish to make changes or need additional information.

Sincerely,

A handwritten signature in cursive script, appearing to read "Patricia L. Kaminski for PLK".

Patricia L. Kaminski, Ph.D.
Associate Professor
Chair, Institutional Review Board

PK:sb

UNIVERSITY OF NORTH TEXAS

1155 Union Circle #305250 Denton, Texas 76203-5017
940.565.3940 940.565.4277 fax <http://research.unt.edu>

APPENDIX B
RECRUITMENT LETTER

Dear _____:

I am a doctoral candidate in Information Science at the University of North Texas (UNT), and I am a librarian at Virginia Tech. I am writing to request an appointment to interview you as a part of my dissertation research.

I contacted you because I found your email address and your research and/or teaching interests on your web page at _____.

I am conducting interviews with faculty members and graduate students at Virginia Tech and at UNT whose teaching and research deals with scientific, social, or ethical aspects of environmental issues such as climate change, land use, biodiversity, natural resources, etc. The interview involves a set of open-ended questions about your research habits and about your opinions of some specific library web-based services.

The data gathered from interviews will be used in my dissertation, which will be made available online through the UNT Libraries, as well as through Virginia Tech University Libraries. Additionally, I hope to publish the results in an article through a peer-reviewed journal. The benefit of participating in this study is to improve services for yourself and all library users at UNT and at Virginia Tech.

For more information, please contact me at the email address below.
I hope to hear from you.

Sincerely,

Nathan Hall
Virginia Tech University Libraries

APPENDIX C

INTERVIEW PROTOCOL

My name is Nathan Hall and I am a librarian at Virginia Tech University Libraries. The library is adding some new services and I am conducting interviews to get a better understanding of our target audience. Furthermore this is part of my research for my doctoral dissertation. I plan to present and publish on the results I obtain.

I really appreciate you sharing your time with me. Please remember that this interview is voluntary. We can stop at any time you choose, and you can decide not to answer any question.

Please tell me about your research. Domain?

What kinds of collaborations are you involved in? (if necessary, just focus on a few projects)

Methodology?

Describe how you find information?

Library resources? Internet? Subscriptions? Professional organizations? Social network?

SCOPUS? Web of Science?

Tell me about the data you collect.

Format? Size?

How do you typically store your raw data?

Storage? Preservation? Accessibility?

Is your research funded?

What are the primary agencies?

Does the funding institution require a data management plan?

How do you typically disseminate your findings?

Journals? Conferences? Social media? Open access?

Could your raw data be re-used in other research?

Are there restrictions?

How would you feel if someone expressed an interest in using your data for additional research?

How would you feel about making your raw data publicly accessible online?

Would you require any tracking?

How would you feel about making your research publicly accessible online?

If journal publisher's IP is mentioned? What if the journal allowed a preprint version?

Would you do this upon publication? Would you require an embargo for a period of time?

What are your basic needs for disseminating research?

Citation analysis? Usage?

TEACHING FACULTY

Please tell me a bit about the classes you teach

Do you work with graduate students?

How do you select information for instructional use?

How do you typically store your teaching materials?

Are your teaching materials publicly accessible online?

What are your basic needs for storing your academic work?

Server space? Encryption? Security?

If I say the phrase "digital library," what comes to your mind?

If I say the phrase “institutional repository” what comes to your mind?

What are the most important resources you would use from a digital library?

Have you ever used a digital library to find information for research or teaching?

If yes, describe that experience.

If no, why not?

Have you ever used digital libraries for hosting your own work?

If yes, describe that experience.

If no, why not?

What would you suggest for improving findability of your own work?

If the university started actively collecting faculty or student work, what would encourage you to submit your work for inclusion? (Policies? Usability? Services? Embargoes?)

What would discourage you?

Is there anyone else you think I should talk to? Colleagues? Graduate students?

REFERENCES

- Akers, K. G., & Doty, J. (2013). Disciplinary differences in faculty research data management practices and perspectives. *International Journal of Data Curation*, 8(2), 5-26.
[doi:10.2218/ijdc.v8i2.263](https://doi.org/10.2218/ijdc.v8i2.263)
- Allen, B. L. (1996). *Information tasks: Toward a user-centered approach to information systems*. San Diego, CA: Academic Press.
- Allen, J. (2005). *Interdisciplinary differences in attitudes towards deposit in institutional repositories*. (Masters thesis). Manchester Metropolitan University, UK.
<http://eprints.rclis.org/6957/>
- Association of Research Libraries. (2014). *Scholarly communication*. Retrieved from
<http://www.arl.org/focus-areas/scholarly-communication>
- Association of Research Libraries. (2006). *To stand the test of time: Long-term stewardship of digital data sets in science and engineering*. Retrieved from
<http://files.eric.ed.gov/fulltext/ED528649.pdf>
- Bethesda statement on open access publishing. (2003). Retrieved from
<http://dash.harvard.edu/handle/1/4725199>
- Bijker, W. E. (1995). *Of bicycles, bakelites, and bulbs: Toward a theory of sociotechnical change*. Cambridge, MA: MIT Press.
- Bohannon, L. (1966). Shakespeare in the bush: An American anthropologist set out to study the Tiv of West Africa and was taught the true meaning of Hamlet. *Natural History Magazine*, 75(7), 28-33.

- Briet, S. (1951). *Qu'est-ce que la documentation*. Paris: Éditions documentaires, industrielles et techniques.
- Brown, J. S., & Duguid, P. (2000). *The social life of information*. Boston, MA: Harvard Business School Press.
- Buckland, M. (1991). *Information and information systems*. Westport, CT: Greenwood Press.
- Carlson, J. (2013). Opportunities and barriers for librarians in exploring data: Observations from the Data Curation Profile workshops. *Journal of eScience Librarianship*, 2(2), Article 2. doi:[10.7191/jeslib.2013.1042](https://doi.org/10.7191/jeslib.2013.1042)
- Case, D. O. (2006). Information behavior. *Annual Review of Information Science and Technology*, 40, 293-327. doi:[10.1002/aris.1440400114](https://doi.org/10.1002/aris.1440400114)
- Correia, A. M. R., & Teixeira, J. C. (2005). Reforming scholarly publishing and knowledge communication: From the advent of the scholarly journal to the challenges of open access. *Online Information Review*, 29(4), 349-364. doi:[10.1108/14684520510617802](https://doi.org/10.1108/14684520510617802)
- Courtright, C. (2007). Context in information behavior research. *Annual Review of Information Science and Technology*, 41, 273-306.
- Cresswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Cronin, B. (2005). *The hand of science: Academic writing and its rewards*. Lanham, MD: Scarecrow Press.
- Davis, P. M., & Conolly, M. J. L. (2007). Institutional repositories: Evaluating the reasons for non-use of Cornell University's installation of DSpace. *D-Lib Magazine*, 13(3/4). Retrieved from <http://www.dlib.org/dlib/march07/davis/03davis.html>

Dryad. (2011). *Joint data archiving policy (JDAP)*. Retrieved from

<http://datadryad.org/pages/jdap>

Editorial: "Data's shameful neglect. [Editorial]. (2009). *Nature*, 461(145). doi:[10.1038/461145a](https://doi.org/10.1038/461145a)

Editorial: The impact factor game. [Editorial]. (2006). *PLoS Med*, 3(6), e291.

doi:[10.1371/journal.pmed.0030291](https://doi.org/10.1371/journal.pmed.0030291)

Ellis, D., Cox, D., & Hall, K. (1993). A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation*, 49(4), 356-369.

Fisher, K. E., & Julien, H. (2009). Information behavior. *Annual Review of Information Science and Technology*, 43, 1-73. doi:[10.1002/aris.2009.1440430114](https://doi.org/10.1002/aris.2009.1440430114)

Foster, N. F., & Gibbons, S. (2005). Understanding faculty to improve content recruitment for institutional repositories. *D-Lib Magazine* 11(1). Retrieved from
<http://www.dlib.org/dlib/january05/foster/01foster.html>

Fowler, K. K. (2011). Mathematicians' views on current publishing issues: A survey of researchers. *Issues in Science and Technology Librarianship*, 67.
doi:[10.5062/F4QN64NM](https://doi.org/10.5062/F4QN64NM)

Fry, J., Proberts, S., Creaser, C., Greenwood, H., Spezi, V., White, S. (2011). *PEER behavioural research: Authors and users vis-à-vis journals and repositories. D4.2 final report*.
Loughborough: LISU and Loughborough University. Retrieved from
<https://dspace.lboro.ac.uk/2134/12967>

Fry, J., & Talja, S. (2007). The intellectual and social organization of academic fields and the shaping of digital resources. *Journal of Information Science*, 33(2), 115-133.

- Garcia, A. C., Dawes, M. E., Kohne, M. L., Miller, F. M., & Groschwitz, S. F. (2006). Workplace studies and technological change. *Annual Review of Information Science and Technology*, 40, 393-437.
- Gargouri, Y., Hajjem, C., Larivière, V., Gingras, Y., Carr L., Brody, T., Harnad, S. (2010). Self-selected or mandated, open access increases citation impact for higher quality research. *PLoS ONE*, 5(10), e13636. doi:[10.1371/journal.pone.0013636](https://doi.org/10.1371/journal.pone.0013636)
- Harley, D., Earl-Novell, S., Arter, J., Lawrence, S., & King, C. J. (2007). The influence of academic values on scholarly publication and communication practices. *Journal of Electronic Publishing*, 10(2). doi:[10.3998/3336451.0010.204](https://doi.org/10.3998/3336451.0010.204)
- Harley, D., Acord, S. K., Earl-Novell, S., Lawrence, S., & King, C. J. (2010). *Assessing the future landscape of scholarly communication: An exploration of faculty values and needs in seven disciplines*. UC Berkeley: Center for Studies in Higher Education. Retrieved from http://escholarship.org/uc/cshe_fsc
- Harley, D. (2013). Scholarly communication: Cultural contexts, evolving models. *Science*, 342(6254), 80-82. doi:10.1126/SCIENCE.1243622.
- Harnad, S., Brody, T., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y.,...Hilf, E. R. (2004). The access/impact problem and the green and gold roads to open access. *Serials Review*, 30(4), 36-40. doi:[10.1016/j.serrev.2004.09.013](https://doi.org/10.1016/j.serrev.2004.09.013)
- Haythornwaite, C. (2006). Learning and knowledge networks in interdisciplinary collaborations. *Journal of the American Society for Information Science and Technology*, 57, 1079–1092.
- Heidegger, M. ([1927] 1972). *On time and being* (J. Stambaugh, Trans.). New York: Harper & Row.

- Holdren, J. P. (2013, February 22). *Memorandum for the heads of executive departments and agencies: Increasing access to the results of federally funded scientific research*. Office of Science and Technology Policy, Executive Office of the [US] President, The White House. Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access memo_2013.pdf
- Howcroft, D., Mitev, N., & Wilson, M. (2004). What we may learn from the social shaping of technology approach. In J. Minger, & L. Willcocks (Eds.), *Social theory and philosophy for information systems*. Hoboken, NJ: John Wiley & Sons, Ltd.
- Jamali, H. R., & Nicholas, D. (2006). *Communicating and information-seeking behavior of research students in physics and astronomy*. *Proceedings of the American Society for Information Science and Technology*, 43(1), 1-18. doi:10.1002/meet.14504301106
- Khoo, M. (2006). Evaluating digital libraries: A sociotechnical approach. *Proceedings of the American Society for Information Science and Technology*, 43(1), 1-9. doi:[10.1002/meet.14504301236](https://doi.org/10.1002/meet.14504301236)
- Khoo, M., & Hall, C. (2010). Merging metadata: A sociotechnical study of crosswalking and interoperability. In *Proceedings of the 10th Annual Joint Conference on Digital Libraries* (pp. 361-364). New York, NY: Association for Computing Machinery.
- Kim, J. (2011). Motivations of faculty self-archiving in institutional repositories. *Journal of Academic Librarianship*, 37(3), 246-254. [doi:10.1016/j.acalib.2011.02.017](https://doi.org/10.1016/j.acalib.2011.02.017)
- Kling, R. (1999). What is social informatics and why does it matter? *D-Lib Magazine*, 5(1). Retrieved from <http://www.dlib.org/dlib/january99/kling/01kling.html>

- Kling, R. (2000). Learning about information technologies and social change: The contribution of social informatics. *The Information Society*, 16(3), 217-232.
- Kling, R., & McKim, G. (1999). Scholarly communication and the continuum of electronic publishing. *Journal of the American Society for Information Science and Technology*, 5(7), 890-906.
- Kling, R., & McKim, G. (2000). Not just a matter of time: Field differences and the shaping of electronic media in supporting scientific communication. *Journal of the American Society for Information Science and Technology*, 51(14), 1306-1320.
- Kling, R., McKim, G. & King, A. (2003). A bit more to it: Scholarly communication forums as socio-technical interaction networks. *Journal of the American Society for Information Science and Technology*, 54(1), 47-67.
- Lage, K., Losoff, B., & Maness, J. (2011). Receptivity to library involvement in scientific data curation: A case study at the University of Colorado Boulder. *Portal: Libraries and the Academy*, 11(4), 915-937. doi:[10.1353/pla.2011.0049](https://doi.org/10.1353/pla.2011.0049)
- Levy, D. (2003). Documents and libraries: A sociotechnical perspective. In A. Bishop, N. van House, & B. Battenfield (Eds.), *Digital library use: Social practice in design and evaluation* (pp. 25-42). Cambridge, MA: MIT Press.
- Lynch, C. A. (2003a, February). *Institutional repositories: Essential infrastructure for scholarship in the digital age* (ARL Bimonthly Report 226).
- Lynch, C. A. (2003b). Questions about audience, economics, and control of digital libraries. In A. Bishop, N. van House, & B. Battenfield (Eds.), *Digital library use: Social practice in design and evaluation* (pp. 191-216). Cambridge, MA: MIT Press.

Marchionini, G., Plaisant, C., & Komlodi, A. (2003). The people in digital libraries: Multifaceted approaches to assessing needs and impact. In A. Bishop, N. van House, & B. Battenfield (Eds.), *Digital library use: Social practice in design and evaluation* (pp. 119-160).

Cambridge, MA: MIT Press.

Max Planck Society. (2003). *Berlin declaration on open access to knowledge in the science and humanities*. Retrieved from <http://openaccess.mpg.de/286432/Berlin-Declaration>

McDowell, C. S. (2007). Evaluating institutional repository deployment in American academe since early 2005: Repositories by the numbers, part 2. *D-Lib Magazine*, 13(9/10).

Retrieved from <http://www.dlib.org/dlib/september07/mcdowell/09mcdowell.html>

Meyer, E. T. (2006). Socio-technical interaction networks: A discussion of the strengths, weaknesses and future of Kling's STIN model. In J. Berleur, M. I. Numinen, & J. Impagliazzo (Eds.), *International Federation for Information Processing: Vol. 223. Social Informatics: An Information Society for All? In Remembrance of Rob Kling* (pp. 37-48).

Boston, MA: Springer.

National Endowment for the Humanities. (2013). *Data management plans for NEH Office of Digital Humanities proposals and awards*. Retrieved from

http://www.neh.gov/files/grants/data_management_plans_2013.pdf

National Science Foundation, Office of Budget, Finance, and Award Management. (2010). *Data management & sharing frequently asked questions (FAQs)*. Retrieved

from <http://www.nsf.gov/bfa/dias/policy/dmpfaqs.jsp#1>

- National Science Foundation. (2005). *Long-lived digital data collections: Enabling research and education in the 21st century*. Retrieved from www.nsf.gov/nsb/documents/2005/LLDDC_report.pdf
- National Science Foundation, Office of Budget, Finance, and Award Management. (2010). *Dissemination and sharing of research results*. Retrieved from <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>
- Nelson, B. (2009). Data sharing: Empty archives. *Nature*, 461, 160-163. doi:[10.1038/461160a](https://doi.org/10.1038/461160a)
- Office of Management and Budget. (1999). *Uniform administrative requirements for grants and agreements with institutions of higher education, hospitals, and other non-profit organizations* (Circular A 110). Retrieved from http://www.whitehouse.gov/omb/circulars_a110
- Open Society Institute. (2002). *Budapest open access initiative*. Retrieved from <http://www.budapestopenaccessinitiative.org/read>
- Otlet, P. (1934). *Traité de documentation: Le livre sur le livre, théorie et pratique*. Bruxelles: Palais Mondial.
- Palmer, J. (1991). Scientists and information: II. Personal factors in information behavior. *Journal of Documentation*, 47(3), 254-275.
- Pool, R. (1997). *Beyond engineering: How society shapes technology*. Oxford: Oxford University Press.
- Reilly, M. & Dryden, A. R. (2013). Building an online data management plan tool. *Journal of Librarianship and Scholarly Communication*, 1(3), eP1066. doi:[10.7710/2162-3309.1066](https://doi.org/10.7710/2162-3309.1066)

- Reinsfelder, T. (2012). Open access publishing practices in a complex environment: Conditions, barriers, and bases of power. *Journal of Librarianship and Scholarly Communication*, 1(1), eP1029. doi:[10.7710/2162-3309.1029](https://doi.org/10.7710/2162-3309.1029)
- Ricœur, P. (1967). *Husserl: An analysis of his phenomenology*. Evanston, IL: Northwestern University Press.
- Rogers, E. (1995). *Diffusion of innovations*. New York, NY: The Free Press.
- Sallans, A., & Donnelly, M. (2012). DMP Online and DMPTool: Different strategies towards a shared goal. *International Journal of Digital Curation*, 7(2), 123-129. doi:[10.2218/ijdc.v7i2.235](https://doi.org/10.2218/ijdc.v7i2.235)
- Salo, D. (2008). Innkeeper at the roach motel. *Library Trends*, 57(2), 98-123.
- Savolainen, R. (2006). Time as a context of information seeking. *Library & Information Science Research*, 28, 110-127.
- Sawyer, S., & Rosenbaum, H. (2000). Social informatics in the information sciences: Current activities and emerging directions. *Informing Science*, 3(2), 89-96.
- Schensul, S. L., Schensul, J. J., & LeCompte, M. D. (1999). *Essential ethnographic methods: Observations, interviews, and questionnaires*, Vol. 2. Walnut Creek, CA: Altamira.
- Schneiderman, B. & Plaisant, C. (2005). *Designing the user interface* (4th ed.). Boston, MA: Pearson/Addison Wesley.
- Seaman, D. (2011). Discovering the information needs of humanists when planning an institutional repository. *D-Lib Magazine*, 17(3/4). doi:[10.1045/march2011-seaman](https://doi.org/10.1045/march2011-seaman)
- Shannon C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana-Champaign, IL: University of Illinois Press.

- Smith, J. S. (2007). The impact of electronic communications on the science communication process—investigating crystallographers in South Africa. *IFLA Journal*, 33(2), 145-159. Retrieved from <http://archive.ifla.org/V/iflaj/IFLA-Journal-2-2007.pdf>
- Sonnenwald, D. H. (2007). Scientific collaboration. *Annual Review of Information Science and Technology*, 41, 643-727.
- Smith, D. W. (2013). Phenomenology. Stanford Encyclopedia of Philosophy. Retrieved from <http://plato.stanford.edu/entries/phenomenology/>
- Swan, A. (2006a). The culture of open access: researchers' views and responses. In N. Jacobs (Ed.) *Open access: Key strategic, technical and economic aspects* (pp. 65-72). Oxford: Chandos Publishing.
- Swan, A. (2006b). Overview of scholarly communication. In N. Jacobs (Ed.) *Open access: Key strategic, technical and economic aspects* (pp. 3-12). Oxford: Chandos Publishing.
- Thomas, C., & McDonald, R. H. (2007). Measuring and comparing participation patterns in digital repositories: Repositories by the numbers, part 1. *D-Lib Magazine* 13(9/10). Retrieved from <http://www.dlib.org/dlib/september07/mcdonald/09mcdonald.html>
- Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of coal getting: An examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system. *Human Relations*, 4(3), 3-38.
- U.S. Department of Health and Human Services, National Institutes of Health, Office of Extramural Research. (2004). *Frequently asked questions: Data sharing*. Retrieved from http://grants.nih.gov/grants/policy/data_sharing/data_sharing_fags.htm#900

- U.S. Department of Health and Human Services, National Institutes of Health, Office of Extramural Research. (2003). *Final NIH statement on sharing research data*. Retrieved from <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>
- van House, N., Bishop, A., & Battenfield, B. (2003). Introduction: Digital libraries as sociotechnical systems. In A. Bishop, N. van House, & B. Battenfield (Eds.), *Digital library use: Social practice in design and evaluation* (pp. 1-21). Cambridge, MA: MIT Press.
- Velarde-Mayol, V. (2000). *On Husserl*. Belmont, CA: Wadsworth.
- Wallis, J. C., & Borgman, C. L. (2011). Who is responsible for data? An exploratory study of data authorship, ownership, and responsibility. *Proceedings of the American Society for Information Science and Technology*, 48, 1-10. doi:[10.1002/meet.2011.14504801188](https://doi.org/10.1002/meet.2011.14504801188)
- Wilhite, A. W., & Fong, Eric A. (2012). Coercive citation in academic publishing. *Science*, 335(6068), 542-543. doi:10.1126/science.1212540
- Wilson, P. (1977). *Public knowledge, private ignorance: Toward a library and information policy*. Westport, CT: Greenwood Press.
- Wilson, P. (1996). The future of research in our field. In J. Olaisen, E. Munch-Petersen, & P. Wilson (Eds.), *Information science: From the development of the discipline to social interaction*. Boston, MA: Scandinavian University Press.
- Wilson, T. D. (2000). Human information behavior. *Informing Science*, 3(2), 49-55.
- Wilson, T. D. (2002). Alfred Schütz, phenomenology and research methodology for information behaviour research. *The New Review of Information Behaviour Research*, 3(71), 1-15.
- Witt, M., Carlson, J., Brandt, D. S., & Cragin, M. H. (2009). Constructing data curation profiles. *International Journal of Digital Curation*, 4(3), 93-103. doi:[10.2218/ijdc.v4i3.117](https://doi.org/10.2218/ijdc.v4i3.117)

Xia, J. (2011). An anthropological emic-etic perspective on open access practices. *Journal of Documentation*, 67(1), 75-94. doi:[10.1108/00220411111105461](https://doi.org/10.1108/00220411111105461)

Xia, J., Gilchrist, S. B., Smith, N. X., Kingery, J. A., Radecki, J. R., Wilhelm, M. L., Harrison, K. C., Ashby, M. L., & Mahn, A. J. (2012). A review of open access self-archiving mandate policies. *portal: Libraries and the Academy*, 12(1), 85-102. Retrieved from: http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v012/12.1.xia.html