Hannah Wang. “Where the Rubber Meets the Road”: Self-Audit in University Digital Repositories. A Master's paper for the M.S. in I.S. degree. April, 2017. 52 pages. Advisor: Helen Tibbo

This study examines the use of trusted digital repository tools and standards for self-audit in digital repositories. While there has been increasing attention given to external audit and certification of trusted digital repositories, there has not yet been a cross-institutional study of repository self-audits.

Describing a series of semi-structured interviews with six information professionals employed at six university digital repositories, this study examines these repositories’ experiences with self-audit. The study explores the tools that are being used for self-audit, how self-audits are conducted, and the value of self-audit to repositories and their stakeholders. Findings from this study provide some insight into the current state of self-audit in digital repositories, and the paper also suggests areas for improvement and future research in this field.

Headings:

Data libraries

Digital libraries

Digital preservation

Institutional repositories

“WHERE THE RUBBER MEETS THE ROAD”: SELF-AUDIT IN UNIVERSITY DIGITAL REPOSITORIES

by

Hannah Wang

A Master’s paper submitted to the faculty of the School of Information and Library Science of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in Information Science.

Chapel Hill, North Carolina

April 2017

Approved by

Helen Tibbo

TABLE OF CONTENTS

INTRODUCTION

Digital Preservation

Digital preservation in libraries and other institutions has received increasing attention in recent years as the volume and variety of digital content has rapidly grown. The Research Libraries Group (RLG) and Online Computer Library Center (OCLC) define digital preservation as "the managed activities necessary for ensuring both the long-term maintenance of a bytestream and continued accessibility of its contents" (RLG-OCLC Working Group on Digital Archive Attributes, 2002, p. 3). The preservation of digital content has become an especially pressing issue for libraries and archives. These institutions are tasked with the preservation of cultural heritage, and the corpus of these heritage artifacts has necessarily expanded to include digital material. Many of these organizations have taken on the dual mission of maintenance and providing access to this digital content as their guiding principles for digital preservation.

The OAIS Reference Model and Trusted Digital Repositories

Over the past two decades, several tools and standards have emerged to guide the digital preservation activities of libraries and other organizations. Most notably, the *Reference Model for an Open Archival Information System* (OAIS) provides guidance for the building of archival systems. Formalized as ISO 14721 in 2003, the OAIS reference model defines the key concepts and required attributes for long-term preservation and access of digital content (Consultative Committee for Space Data Systems, 2012). Within these concepts, the "Designated Community" of the OAIS is defined as the identified group of users who should be able to access and understand the information held in the archival system. It is the job of the OAIS to make its information (contained in "digital objects") accessible to these users. Although originally developed by the Consultative Committee for Space Data Systems, the OAIS reference model has found wider application in the digital preservation community, and it has become the foundation of planning for and implementing digital repositories.

Many libraries and other institutions have built digital repositories in order to fulfill their missions of digital preservation. These systems store, maintain, and provide access to the intellectual output of the institution (including scholarly papers and articles), electronic journals and books, digital special collections, and other types of digital content. This broad category of systems can therefore include institutional repositories, data repositories, audiovisual archives, museum image collections, and other types of digital libraries.

The concept of "trustworthy digital repositories" evolved concurrently with and is closely related to the OAIS reference model, and it has been used to refine and expand the model further. In the RLG-OCLC (2002) report, *Trusted Digital Repositories: Attributes and Responsibilities*, a trusted digital repository is defined as a digital repository "whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future" (p. 5). In addition to the features required by the OAIS reference model, a trusted digital repository must provide evidence of its "trustworthiness," proving the long-term viability and sustainability of its systems. The attributes of a trustworthy digital repository include:

- Compliance with the Reference Model for an Open Archival Information System (OAIS)
- Administrative responsibility
- Organizational viability
- Financial sustainability
- Technological and procedural suitability
- System security
- Procedural accountability (RLG-OCLC Working Group on Digital Archive Attributes, 2002, p. 13)

Proof of these attributes in a digital repository is meant to signal to the designated community that this digital repository can be trusted. Ross and McHugh (2006) rationalize this emphasis on trustworthiness as "an approach to handling [the] uncertainty" (Introduction, para. 1) and anxiety that surrounds the preservation of digital materials. If a digital repository can meet these standards of trustworthiness, it should be able to allay these fears among the patrons, contributors, and institutions that use and support the repository.

Audit and Certification

Several auditing tools have emerged based on the OAIS reference model and the RLG-OCLC report, in order to operationalize these definitions and provide criteria for assessment. The RLG-OCLC (2002) report recommends a certification process, positing "a program for certification could provide a basis for trustworthiness" (p. 10). In other words, a certified trusted digital repository should signal to its designated community that it is worthy of their trust to steward their digital materials and information.

Each auditing tool serves a slightly different purpose, although they may have significant overlap. The many tools created for trustworthy digital repository audit and certification include the Data Seal of Approval (DSA), the Digital Repository Audit Method Based on Risk Assessment (DRAMBORA), the Trustworthy Repository Audit and Certification (TRAC) Checklist, and ISO 16363 ("Data Seal of Approval," n.d.; "DRAMBORA: About," 2015; RLG-NARA Digital Repository Certification Task Force, 2007; Consultative Committee for Space Data Systems, 2011) . These tools all provide checklists and guidelines for repository architectural and organizational features that help to ensure the long-term preservation and access of digital content. While DRAMBORA is a risk assessment tool meant for solely for internal audit, DSA, TRAC, and ISO 16363 are ultimately intended for external audit and certification, helping producers of digital content decide where to deposit and guiding consumers of digital content to trustworthy repositories of information. All of these tools, however, can be used for the purposes of self-audit. Indeed, a repository must always conduct a self-audit before it can commission an external one. We will see that more and more repositories are using these tools to conduct self-audits, though commissioning an external audit is not always their ultimate goal.

Repository Self-Audit

While there has been increasing attention given to audit and certification of digital repositories (Dillo & de Leeuw, 2015; Lee & Tibbo, 2007; RLG-NARA Digital Repository Certification Task Force, 2007; RLG-OCLC Working Group on Digital Archive Attributes, 2002; Ross & McHugh, 2006), at the same time, many repositories are finding these standards most useful for conducting self-audits (Downs & Chen, 2010;

Innocenti & Vullo, 2009; Krahmer and Phillips, 2016; Pejsova & Vaska, 2012; Schmidt, 2011; Steinhart, Detrich, & Green, 2009). Despite the hope that many repositories will commission external audits and, when the system of credentialing is established, seek certification, many trusted digital repository tools are extremely useful as instruments to guide self-audit. According to the literature, which will be discussed in the next chapter, there are a variety of reasons that repositories choose to conduct self-audits, but to date, there has not been a cross-institutional study of these reasons, nor of the effectiveness of existing tools in guiding the self-audit process. This paper will investigate the current use of trusted digital repository tools to conduct self-audit in North American university digital repositories. In particular, it will ask these research questions:

- What tools are digital repositories using to conduct self-audits?
- How is self-auditing being conducted?
- What is the value of this work to the repository and its stakeholders?

LITERATURE REVIEW

Digital Repositories

For the last decade, researchers have noted the progress and impact of digital repositories. However, one point of inconsistency within the literature is the very definition of digital repositories. The RLG-OCLC report, *Trusted Digital Repositories: Attributes and Responsibilities*, offers a multifaceted definition based on different scenarios, wherein a digital repository may be:

- The repository of digital content for a national library
- An institutional repository maintained by a university library, aimed at collecting and disseminating the scholarly output of the university
- A preservation system for digitized and born-digital content in a museum
- A repository of e-journals
- The digital library of a small cultural institution (RLG-OCLC Working Group on Digital Archive Attributes, 2002, p. 5).

The RLG-OCLC report notes that this list is not exhaustive. In essence, a digital repository is an organization aimed at the long-term preservation and access of digital materials. This definition emphasizes the wide variety of cultural institutions that may host digital repositories, and the many different purposes that those repositories may share.

In their report, "Digital repositories ten years on," Nicholas, Rowlands, Watkinson, Brown, and Jamali offer a more specific classification of digital repositories. In a survey to over 1,600 international researchers, Nicholas et al. supply an operational definition of digital repositories, meant to clarify the terms used in the survey. This definition includes:

- Institutional repositories which aim to collect widely across a particular university or similar institution, possibly covering a wide range of formats.
- Subject repositories based on collecting only within a certain discipline, usually across more than one institution and often international in coverage.
- Format repositories whose scope is limited by collecting in a particular format, e.g. student dissertations and e-theses, research data, digital images (Nicholas, Rowlands, Watkinson, Brown, & Jamali, 2012, p. 196).

Nicholas et al. use a similar approach to the RLG-OCLC report for their definition, using familiar examples of formats and institutions to supply meaning. However, their definition differs in several important ways. There is no mention of the preservation functions of repositories. These three types of repositories also collect the same broad category of content: scholarly output. Since the purpose of this survey was to study the usage and perceptions of digital repositories by scientists and other academic researchers, it only focuses on the repositories in which they may be depositing and accessing content.

Institutional repositories, especially, have been the primary of focus for work in the trusted digital repository (TDR) community (Hank, Tibbo, & Barnes, 2007; Johnston, 2012; Lee & Tibbo, 2007; Li & Banach, 2011), and these were among the first repositories to use TDR audit tools. Institutional repositories are also steadily gaining in number and prominence. As early as 2003, Clifford Lynch declared that institutional repositories were an "essential infrastructure for scholarship in the digital age" (Lynch, 2003, p. 1). Nicholas et al.'s 2012 international survey of researchers at academic institutions showed over 60% of respondents had deposited work in a digital repository, and that around 50% of respondents anticipated institutional repositories becoming more or much more important in the next three years (Nicholas et al., 2012). Indeed, OpenDOAR, an international listing of open access repositories (including institutional, subject, and format repositories), reports that its database has grown steadily since 2007, expanding from 866 repositories in January 2007 to over 3201 repositories in 2016 (OpenDOAR, 2016).

As digital object management and preservation is a major aspect of many trusted digital repository standards, this paper will limit the definition given by Nicholas et al. to include only digital repositories that perform digital preservation activities. This paper will expand the definition used by Nicholas et al. to include not only repositories of scholarly output, but also repositories of digital special collections and archives in libraries, as referenced in the 2002 RLG-OCLC report. For the sake of limiting scope, this paper will only study digital repositories hosted within North American universities.

Trusted Digital Repositories

Implementation

Since the publication of the 2002 RLG-OCLC report, there has been a large volume of literature devoted to trusted digital repository standards and tools. Much of this literature deals with practical implementation: how repositories may use trusted digital repository tools to improve their work and how they might demonstrate their trustworthiness.

Soon after the RLG-OCLC report, there was a large amount of development around technical infrastructures and tools that would support the work of trustworthy digital repositories. Jantz and Giarlo (2005) present an optimistic view of digital preservation, positing that not only is the long-term maintenance of digital objects possible, but there are many readily available technologies that might be integrated into a repository framework. Using the Fedora repository framework, they enumerate the ways in which the technical features of a repository are equipped to bolster its trustworthiness. These features include digital signatures, persistent identifiers, and audit trails. Meanwhile, Lawson and Spies (2004) wrote about a set of tools being developed by the OCLC to catalogue and ingest objects to repositories, based on the recommendations of the OAIS Reference Model and the RLG-OCLC report. These tools supported preservation metadata record creation, metadata harvesting, digital object ingest and administration, and dissemination.

By 2006, there had been a lot of work put into developing criteria for trustworthy digital repositories and developing tools to support them, but very little guidance on how repositories might demonstrate fulfillment of these criteria. Ross and McHugh (2006) focus on the role of evidence in establishing trust in repositories. Indeed, the audit tools being developed at the time, including the Trustworthy Repositories Audit & Certification (TRAC) Checklist, did not include documentation requirements for audit. Ross and McHugh focus on types of evidence that might be useful for auditors when reviewing a digital repository. They propose three different types of evidence that might demonstrate trustworthiness: documentary evidence, observation of practical evidence, and testimonial evidence (Ross & McHugh, 2006). Ultimately, the TRAC document included a note in each section about the types of evidence that might be used to demonstrate that a repository has met that standard (RLG-NARA Digital Repository Certification Task Force, 2007).

In 2006, Cal Lee and Helen Tibbo organized a workshop at the Joint Conference on Digital Libraries entitled, "Digital Curation and Trusted Repositories: Seeking Success." Like Ross and McHugh, the organizers were interested in the practical implementation of trusted digital repository guidelines, as well as how trusted digital repositories might be approached from the context of digital curation. After the conference, the organizers, Lee and Tibbo (2007), assembled a summary of the most promising approaches for implementing the attributes of the first draft of the RLG-NARA Trustworthy Repository Audit and Certification (TRAC) Checklist. Ross and McHugh's call for providing evidence was echoed by several workshop participants, who agreed that "one should not only write or follow rules but should also provide evidence for compliance with the rules" (Promising Approaches section, para. 3). Participants also hoped for more generalizable and shared products across communities to facilitate the development of trustworthy digital repositories. They hoped that this type of collaboration would lead to, among other things, guidance documents built on other existing documents (Lee & Tibbo, 2007).

Around the same time, Hank, Tibbo, & Barnes (2007) conducted a survey of members of the Association of Research Libraries (ARL). The purpose of the survey was to learn about the applications of the current draft of the TRAC Checklist. The authors were particularly interested in the use of the checklist as a planning tool for repositories, rather than in its primary purpose as an auditing tool. The majority of repositories that were aware of the checklist affirmed its usefulness as a planning tool. Dearborn, Barton, and Harmeyer (2013) confirm this usage, discussing the tool's use for repository planning for the Purdue University Research Repository. However, of the 33 respondents to the 2007 survey with repositories in development or production, only 12 reported actually using the checklist. At the time, despite the prominence and perceived usefulness of the checklist, it had not found much practical application. Hank et al. also noted that certification should not be an assumed goal for all institutional repositories. Seeking certification can be an arduous task, and many repositories lack the financial and personnel resources to become certified as a trustworthy digital repository. Hank et al. recommend improvements to the checklist draft to increase usage.

"Trust" and its Discontents

The concept of trustworthy digital repositories is predicated on a notion of "trust." The 2002 RLG-OCLC report offers the standard *Merriam Webster Dictionary* definition of trust:

> assured reliance on the character, ability, strength, or truth of someone or something . . . one in which confidence is placed . . . a charge or duty imposed in faith or confidence or as a condition of some relationship . . . something committed or entrusted to one to be used or cared for in the interest of another (RLG-OCLC Working Group on Digital Archive Attributes, 2002, p. 8).

The report goes on to state that, according to this definition, most libraries, archives, and museums are trusted institutions. The report recognizes that trust is a concept rooted in relationships, and it applies this definition to three relationships inherent in digital repositories: the relationship between the institution and the Designated Community, the relationship between the institution and third-party service providers, and the relationship between the users in the Designated Community and the documents themselves. The report assumes that the first trust relationship is a given, the second one can be established through best practices and certification, and the third one can be cultivated through proof of document integrity (RLG-OCLC Working Group on Digital Archive Attributes, 2002).

Kelton, Fleischmann, and Wallace (2007) propose a much more complex definition for trust in digital information, the Integrated Model of Trust. They extend models of trust from the social sciences to incorporate concepts from information science and human-computer interaction, creating a new model of how trust can be established between an information source and the user. This definition goes beyond the standard dictionary definition used by the RLG-OCLC report, taking into account preconditions for trust, trust development processes, aspects of trustworthiness (including accuracy, objectivity, validity, and stability), influences on judgments of trustworthiness, and elements of the trust itself (Kelton, Fleischmann, & Wallace, 2008).

Several authors challenge these definitions of trust and the very concept of trust within digital repositories. These arguments all take issue with the lack of weight given to the user in judgments of trust. Prieto (2009) calls for a reframing of "trust" in digital repositories as an extrinsic quality:

> While digital repositories may be trustworthy because of adherence to technological standards, accepted practices, and mechanisms for authenticating the authorship and accuracy of their content, it is ultimately their respective stakeholders – both those who deposit and use content – whose perceptions play a central role in ensuring a digital repository's trustworthiness (p. 593).

This understanding of trust relies on both the depositors and end-users of a repository to make the final judgment of trustworthiness, rather than a member of the institution or an external auditor. Prieto shifts the focus from the internal standards and practices of a repository to the Designated Community, which has always been at the heart of the OAIS Reference Model. Prieto's article does not directly criticize the definition of trust used by the RLG-OCLC report or any other – in fact, he asserts that, by acknowledging the role that the Designated Community plays in the ecosystem of a trusted digital repository, they emphasize the importance of the community to the process of establishing trust (Prieto, 2009).

Other authors do not look on the RLG-OCLC report so favorably. Bak (2015) argues that the definition of "trust" implemented in the report does not align with the one quoted from *Merriam Webster*, and that trusted digital repository "standards culture" has actually created a new definition of trust that does not align with any definitions based in current societal norms. According to Bak, the concept of trust is an attribute of a relationship – a digital repository must be trusted by someone in order to be considered trustworthy. However, the RLG-OCLC report "positions trust as a quality that can be unilaterally created, audited and certified" (Bak, 2015, p. 9). In Bak's view, the trust relationship that the RLG-OCLC report takes for granted, between the user and the repository, is symptomatic of this problem. It assumes that cultural institutions are themselves trustworthy, when there is history of these institutions mishandling the records of minority groups and creating a homogenized narrative.

In addition to these criticisms of ill-defined “trust” in digital repositories, there have been several studies looking at user perception of trustworthiness in digital repositories. Yakel, Faniel, Kriesberg, and Yoon (2013) build upon the work done by Kelton and Prieto to define trust from the perspective of the end user, studying the perceptions of archeologists and social scientists of digital repositories. Yakel et al. found in their research that these users associate trust in repositories with transparency, both in terms of mission and willingness to be audited. In addition to valuing preservation and sustainability in repositories, these end users also relied on the recommendations of colleagues and institutional reputation when deciding in which repositories to place their trust. These latter two qualities represent a definition of trust that is clearly distinct from how it is defined by the RLG-OCLC report, and Yakel et al. advocate for a thorough understanding of how repository stakeholders conceive of trust in repositories in order for repositories to meet those needs.

Donaldson and Conway (2015) also investigate user conceptions of trustworthiness, but rather than approaching trustworthiness from the larger repository level, they look at it from the level of the archival document. They find that, at the document level, expert archival users conceptualize trustworthiness in terms of authenticity, tying these concepts in digitally preserved and delivered documents to more familiar concepts in analog archives. They also find that the Kelton Integrated Model of Trust, although accurate in the importance it places on trustworthiness in the larger concept of trust, fails to address the full complexity of how users conceptualize trustworthiness in archival documents. In addition to considering factors such as accuracy and stability, archives users are also concerned about factors “having to do with a document’s first-hand nature, legibility, and form” (Donaldson & Conway, 2015, p. 2439). These factors, tied to archival notions of authenticity, are not addressed in the

Kelton model. Yoon (2015) also looks inside the contents of the repository to study notions of trust and trustworthiness, focusing on the field of data curation and reuse. She studies user trust in data generated by other researchers, focusing on the user-defined trust attributes that formulate their trust judgments.

Repository Self-Audit

Despite the recent literature questioning the nature of repository trustworthiness, many repositories have still found it useful to measure their own preservation activities and trustworthiness against existing standards and guidelines. The literature concerning self-audit in repositories is dominated by case studies. Several repositories have reported their individual experiences with implementing self-audit tools through reports and papers, a sampling of which will be addressed here.

Data Seal of Approval (DSA)

Dutch science organizations Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) and Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) developed the Data Seal of Approval (DSA) in 2008 with the objective of certifying data repositories that can assure depositors that their data will be stored in a reliable manner where it can be accessed and used ("About | Data Seal of Approval," n.d.). Sixteen guidelines comprise the DSA assessment tool. Data repositories must complete a self-assessment using the 16 guidelines and submit the assessment for peer review. To date, 62 repositories have been awarded the DSA.

The Data Seal of Approval has a self-audit process built into its certification criteria. The starting point of the process is conducting a self-assessment with the DSA online tool ("About | Data Seal of Approval," n.d.). After completing this assessment, it can be submitted for peer review. Dillo and de Leeuw (2015) note that preparing and conducting this self-assessment can be a benefit in itself, because it can help the repository improve its communication processes and management procedures. These improvements may result in an overall higher level of professionalism in a repository, enhancing its operations and reputation.

Digital Repository Audit Method Based on Risk Assessment (DRAMBORA)

Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) is an assessment tool provided by the Digital Curation Centre and DigitalPreservationEurope for digital repositories to assess the risks to their digital archiving systems. It was developed between 2006 and 2007. DRAMBORA's approach to digital preservation is one of risk-management: digital preservation specialists can use the tool to identify their objectives and assets in order to assess the risks to digital objects in the repository (“DRAMBORA: About,” 2015).

DRAMBORA differs from other auditing tools in its purpose. The end goal for DRAMBORA is self-audit, rather than certification. Repositories can use DRAMBORA to identify their assets and assess the risks to their systems. Innocenti and Vullo (2009) and Pejsova and Vaska (2012) found DRAMBORA useful for assessing gaps in digital curation activities. The tool has been implemented at a range of digital repositories, from institutional repositories and closed digital collections (Innocenti & Vullo, 2009) to the Czech National Repository of Grey Literature (Pejsova & Vaska, 2012) to the CERN Document Server in Switzerland to the US Geological Survey (“DRAMBORA Interactive: Users,” 2015). The online tool, “Drambora Interactive,” has also been noted for its usefulness – it walks users through the assessment steps, and it can create output reports on the risks identified during the assessment. This can be especially useful for iterative audits (which DRAMBORA is meant to facilitate), as reports from different audits can be compared and used to track the progress of the repository (Pejsova & Vaska, 2012).

Trustworthy Repository Audit and Certification (TRAC) Checklist and ISO 16363

The RLG-OCLC's recommendation for a trusted digital repository certification process resulted in the creation of an RLG-NARA Digital Repository Certification Task Force, which published *Trustworthy repositories audit & certification: Criteria and checklist* in 2007 (Steinhart, Detrich, & Green, 2009; RLG-NARA Digital Repository Certification Task Force, 2007). After a period of pilot testing, TRAC evolved into the *Recommendation for space data system practices: audit and certification of trustworthy digital repositories*, published in 2011 by the Consultative Committee for Space Data Systems, and formalized as ISO 16363 in 2012 (Consultative Committee for Space Data Systems, 2011). Most of the revisions made to TRAC in the development of ISO 16363 were structural, and the core concepts and requirements remained the same (“Digital Preservation Metrics | CRL,” n.d.). Both checklists group their assessment metrics into three categories: Organizational Infrastructure, Digital Object Management, and Infrastructure and Security Risk Management (or, "Technologies, Technical Infrastructure, and Security" in TRAC) (Consultative Committee for Space Data Systems, 2011; RLG-NARA Digital Repository Certification Task Force, 2007). These sections evolved from the "Attributes of a Trusted Digital Repository" named in the 2002 RLG-OCLC report (p. 13). Currently, the Center for Research Libraries (CRL) is using TRAC and ISO 16363 to conduct external audits of repositories and to date they have certified six repositories. However, the CRL has no official standing to conduct these audits and certify repositories, as no body yet exists to certify auditors for ISO 16363, at the time of writing.

Although their primary purpose is to facilitate certification processes, TRAC and ISO 16363 have also been used for self-audits. Ambacher (2007) acknowledges the role that self-audit can have in the certification-seeking process: “A repository should only seek digital certification once it has undergone self evaluation and has addressed all issues which arose from that self evaluation” (Precedents for External Evaluation, para. 2). Some repositories, however, view self-audit with TRAC as an auxiliary use of the checklist that may not result in certification. This is an extension of the observations made by Hank et al. (2007), that while certification is not practical for all repositories, the checklist can still be a useful document for other purposes such as planning. Several repositories have taken these applications beyond planning to self-auditing.

The purposes of self-auditing with TRAC and now ISO 16363 vary. Steinhart, Detrich, and Green (2009) write that the checklist was used to conduct a self-audit at a data-staging repository at Cornell, in order to improve their management practices and establish trustworthiness. This usage of the checklist is notable because the repository in question was not meant for long-term preservation – it is a repository where data is staged in order to be shared or published elsewhere. The self-audit was conducted in order to assess whether this repository was fulfilling its duties in the overall preservation lifecycle of data. Downs and Chen (2010) report on the use of the checklist to conduct a self-audit of a scientific data archive. The ultimate goal for this assessment was to become certified as a trusted digital repository. The self-audit would serve to identify gaps in service that needed to be filled prior to certification. Schmidt (2011) recounts the self-auditing process for an email list archives, which was an iterative process to assess preservation practices and make recommendations for their improvement. A secondary goal of this self-audit was to use TRAC as a research tool to investigate preservation practices for email, thereby contributing to the email archiving field. Krahmer and Phillips (2016) report that TRAC was used to conduct a self-audit of University of North Texas Libraries Digital Collections in order to affirm their commitment to digital preservation, and to increase communication around digital repository activities at UNT Libraries.

Many of these case studies contain useful recommendations for practical implementation of TRAC for self-audit in digital repositories. Steinhart et al. (2009) report that two kinds of documentation were generated during the process of self-audit – external and internal documentation. Steinhart et al. also make a case for the use of TRAC in repositories where long-term preservation is not a goal – there are still large portions of the document that are useful for assessment. Schmidt (2011) and Krahmer and Phillips (2016) emphasize the collaborative nature of completing the checklist for self-audit: the process should involve interviews with and documentation from staff in many different parts of the organization. Krahmer and Phillips even note that this is one of the major benefits of repository self-audit – it can start a larger conversation about digital preservation in the library. Krahmer and Phillips also recommend using the self-audit documents released by other repositories as guidelines and starting points for a self-audit. All authors note that TRAC is particularly useful for identifying and addressing shortcomings in repositories, especially when self-audit is conducted iteratively (Downs & Chen, 2010; Krahmer and Phillips, 2016; Schmidt, 2011; Steinhart et al., 2009).

Using this literature on self-audit as my starting point, I will study the effect of trusted digital repository certification on the internal assessment processes of university digital repositories in North America.

METHODS

The data for this study was collected through semi-structured interviews with repository managers. These participants were identified through snowball sampling and recruited through email invitations. Questions in the interviews followed an interview guide, focusing on the function of the repository, the repository's background in self-auditing, self-audit tools and methods, and the costs and value of self-audit. These interviews were transcribed and coded using open coding, allowing codes to emerge from the data collected. The results of this study demonstrate some of the tools and methods being used for self-audit in repositories and the value of this work.

Participants

Six participants were interviewed for this study. They were staff members from six different institutions with managerial roles in their respective repositories. This included librarians, technical directors, and other types of managers. A snowball sampling technique was used. It proved difficult to recruit participants for this study: the area of trusted digital repositories is relatively new, few repositories have the capacity and resources to conduct a self-audit, and there is no published list of repositories that have conducted self-audits. When it is particularly difficult to identify participants for a study, snowball sampling can be an effective technique for finding and recruiting participants (Wildemuth, 2009). In this study, repositories that were known to be conducting or had conducted self-audits were contacted first, and they were asked to identify other repositories that have experience with the self-audit process. Recruitment notices were also posted in the Archivematica and ALA Digital Curation Interest Group Google Groups; however, there was no response to these notices, so no study participants were recruited using this method.

Each prospective participant was sent an invitation to participate in the study via email (Appendix A). This email explained the purpose of the study and gave an overview of the general procedures, including the expected length of the interview. Individuals that agreed to participate were sent an additional information sheet about the study, with contact information for the interviewer and faculty advisor (Appendix B).

Data Collection

Data was collected using semi-structured interviews. The interviewer used an interview guide (Appendix C), which contained an outline of questions and prompts that could be adjusted depending on each interviewee's responses. Questions in the interview guide fell into six broad categories: general introductions, self-audit introductions, self-audit tools, method of self-audit, risks and costs, and value of self-audit. Each interview concluded with two "wrap-up" questions, inviting the participant to ask questions or add comments not addressed in the interview, and inviting the participant to identify other potential participants.

Interviews were conducted in three different ways. For participants within driving distance of the interviewer's institution, interviews were conducted in person and recorded using an iPhone. For most other participants, interviews were conducted through a video chat service, GoToMeeting, and recorded using the service. For one interview that could not be conducted using GoToMeeting due to scheduling conflicts, the interview was conducted and recorded using Google Voice.

Each interview began with a brief introduction to the study, then proceeded to the main body of the interview guide. Incomplete notes about participants' responses were taken in order to guide the interview. Some participants also utilized the text chat function in GoToMeeting or email to send links to relevant materials on the Web. After each interview concluded, the interviewer took more thorough notes, capturing her initial reactions.

Data Analysis

Each interview was fully transcribed and annotated by the interviewer. Significant responses were categorized based on theme and entered into a spreadsheet. During the process of collecting and recording responses, the interviewer noticed patterns and was able to develop codes accordingly.

Implications

As Wildemuth (2009) notes, snowball sampling may not result in a representative sample of the population. However, the purpose of this study is not to produce generalizable results, but to produce findings that are indicative of certain trends and may be transferred to different contexts.

The results of this study will provide an overview of some of the current self-auditing practices in digital repositories. These findings may provide some guidance for repositories that are considering undertaking a self-audit. The results of this study may also provide some information about what kinds of resources are most useful in the self-auditing process, as well as the types of resources that would help the process, but might not currently exist.

FINDINGS

Six participants were interviewed for this study, each representing a different repository. All six repositories are housed at public universities in North America, though not all repositories are organizationally located within an academic library. Two participants interviewed in this study are employed at data repositories, both of which focus on social science data. These two repositories are organizationally independent of the academic libraries at their universities. Most of these repositories house content for outside institutional partners, such as historical societies, university consortia, and state digital collections. All but one repository had completed at least one self-audit – this repository started the self-audit process but had to stop for personnel-related reasons. It is currently in the planning stages for conducting another self-audit. The type and scope of these repositories are described in Appendix D.

Background

What is a repository?

For some institutions, defining what a repository is – what the institution or service is that is actually being assessed – is itself a sticking point in the self-audit process. For three of the participants, the repository that they examined in the self-audit process was a single discrete entity. The repository itself may have multiple collections contained within it, but all collections share the same preservation infrastructure and access interface.

On the other hand, three of the participants interviewed work in institutions that house more than one repository. In some cases, these multiple repositories share a common infrastructure with multiple access interfaces; in other cases, the same library unit manages the multiple repositories. These different types of repositories include special collections and archives repositories, which might hold image, audio, video, or web materials from a number of sources, and institutional repositories, which are generally responsible for housing scholarly output from the university.

All three of these participants discussed grouping their multiple repository services together under one umbrella in the self-audit. In one case, the library ended up defining a "digital collections" grouping for the express purpose of conducting a self-audit of their repository services.

Why do a self-audit?

The reasons for conducting a self-audit varied among the repositories. Some reasons were external. All five repositories that house materials as a part of partnerships with outside institutions cited accountability to partners as a reason for self-audit. However, a participant from one of these repositories noted, "Zero of our partners asked [if we had done a self-audit]. It was our peers." Indeed, three of these repositories cited respect among peers in the digital preservation and repository field as a major reason to conduct a self-audit. In these repositories, a previously established reputation as a "standard-bearer" for digital preservation practices was itself an impetus to use a trusted digital repository standard to conduct a self-audit.

Other reasons for conducting self-audit were internal. The most commonly cited reason for conducting a self-audit was to identify the gaps in the repository's services or policies. All six participants stated that this gap assessment and the resulting improvements were major reasons for conducting self-audit. Additionally, two participants stated that the desire to increase institutional awareness and understanding of digital preservation work was a reason to conduct self-audit.

What about external audit?

None of the repositories in this study have been certified with an external audit by the CRL using the TRAC Checklist. Two participants expressed a desire to undertake an external audit using TRAC or ISO 16363, but both cited the expense of an external audit as a major barrier. One participant noted that an external audit seems more important for fee-based repository services, which might need to demonstrate external trustworthiness to their memberships through formal certification. Another participant at a repository not interested in external audit stated, “we recognize the process of actual certification doesn’t clearly benefit our libraries in any certain way, but we recognize that a self-assessment should always be part of any preservation program.”

Self-Audit Tools and Methods

Standards used

All six of the repositories in this study had experience using TRAC for self-audit. Most participants stated that the (then upcoming) standardization of TRAC as ISO 16363 was a major reason for using TRAC at the time that they conducted their self-audits. Two participants also stated that the popularity of the standard for external and self-audit was a reason to use it: more resources and community support existed for conducting a self-audit with TRAC. One participant also noted that the structure of TRAC itself was appealing to their repository: “in the introduction to the TRAC standard, it says that TRAC is built on the two pillars of preservation and transparency.”

Three participants stated that their repositories were either preparing to use or currently using ISO 16363 to conduct a self-audit. None of the repositories in this study have yet completed a self-audit using ISO 16363.

Two other standards were mentioned in interviews with participants from the two data repositories. Both of these repositories have used DSA to conduct a self-audit and achieve external certification. One of these repositories also used DRAMBORA to conduct its first self-audit ten years ago.

Not all of these repositories adhered strictly to a standard when conducting a self-audit. One participant stated that their repository had based their work and documentation off both TRAC and ISO 16363, but they also produced supplemental documentation as part of the self-audit process that was not suggested by any tool or standard. Another participant stated that, when another self-audit is conducted in the future, the repository will likely “tweak portions of the self-audit process to better suit our size and our needs.”

Three participants also noted that the scope and scale of the different auditing standards lend themselves to sequential use. DSA is considered more “lightweight” for conducting a self-audit, whereas TRAC is a larger undertaking. One participant noted that, if they were able to do the whole self-audit process over again, they would have done DSA, then the nestor *Catalogue of Criteria* (*Catalogue of Criteria for Trusted Digital Repositories*, 2009), then TRAC or ISO 16363, because that would result in a “nice flow from less complicated to more complicated and more involved.”

Other resources

Participants were also asked about other resources that they found useful in their self-audit processes. Four of the participants stated that looking at past examples of self-audits completed by other institutions was helpful in completing their own self-audits. The Trusted Digital Repository documentation created by Scholars Portal for their own audit was cited by all of these participants as being a useful resource in writing their own documentation (Scholars Portal, 2012). Documentation created by University of North Texas Libraries, CLOCKSS, and MetaArchive for their self-audits were also cited in these interviews as being particularly useful (Phillips, Tarver, Krahmer, Alemneh, & Waugh, 2015; Rosenthal, 2014; Schultz, 2010).

All participants also noted that the digital preservation and repository communities are invaluable resources for conducting self-audits. Two participants noted that they had partnered with another institution for at least a portion of the self-audit process. Four participants said that discussions with professional groups such as the Digital Curation Interest Group of the American Library Association, the Preservation and Archiving Special Interest Group (PASIG), the International Association for Social Science Information Services and Technology (IASSIST), and the International Federation of Data Organizations (IFDO) had been helpful.

Who does the work?

A self-audit using any tool or standard requires a substantial amount of research and writing. The repositories in this study delegated that work in different ways. Three of the repositories were able to form committees to divide up the work of the self-audit, even if one individual did a large portion of the writing. The other three repositories had only one or two people doing all the work of the self-audit, occasionally relying on other staff for consultation or oversight. In two of these cases, a temporary student or intern was responsible for a large portion of the self-audit, either through managing the project or writing documentation.

Challenges of Self-Audit

Time

All participants noted that time was the biggest cost and challenge associated with the self-audit process. For all repositories in the study, staff members had to do the work of a self-audit alongside the day-to-day operations of the repository. In an effort to address the time-consuming nature of a self-audit, two of the repositories in the study are actively pursuing methods to automate their self-audit processes, especially in auditing technical systems requirements.

All participants also stated that, because of its time-consuming nature, a self-audit requires organizational commitment. One participant acknowledged that there are many other priorities to consider in a repository: “there’s lots of priorities – go and get funds, write proposals to keep the lights on, and the doors open.” This participant said that self-audit, including documentation of preservation practices and policies, should be one of these priorities, and it should preferably be included in the job description of a staff member. Another participant noted that having an advocate for digital preservation in library administration was extremely helpful in getting organizational commitment for self-audit.

Several participants noted that a second self-audit had been or would be much easier, because much of the documentation would have already been written and would only need minor updates. One participant noted that writing the policies for digital preservation is the major stumbling block for conducting a self-audit; while many institutions may have the systems and technology to support digital preservation, the policy documentation has not been developed with the same speed and focus. Two participants noted that the self-audit process consisted of "stopping and starting" every time a new policy needed to be written in order to conform to audit criteria.

Mapping language and concepts

Another challenge in the self-audit process, especially when using TRAC and ISO 16363, was in the language used by the tool. One participant noted, "the biggest challenge we ran into was trying to map the TRAC language and TRAC notions over into what our institutional structure looks like." Other participants echoed this sentiment, and it was cited as one of the reasons to look at other institutions' self-audit documentation for guidance. One participant also noted that mapping the trusted digital repository language and concepts described in TRAC could be especially difficult for smaller repositories, which may not have the technical infrastructure or expertise to create and host a trustworthy digital repository, but may be interested in assessing how they handle digital materials. This participant additionally said that resources should be created for such institutions to help them understand the self-audit process and tools, so that they can see how these principles might apply to their own digital preservation work.

Value of Self-Audit

Identifying and fixing shortcomings

As previously noted, all six participants stated that identifying shortcomings or problems in their repositories was a major reason to conduct a self-audit. All five repositories that finished conducting a self-audit cited specific gaps in their policies and processes that they had fixed because of the self-audit, pointing to these improvements as direct benefits of the process. For some repositories, this meant writing a digital preservation policy or a collection development policy for digital materials, which can help fulfill requirements in a TRAC audit. For one smaller repository, the gaps illuminated by the TRAC self-audit led to a complete self-audit of all their policies and donor files. Another participant said that the self-audit process helped them identify problems in their digital preservation systems and refine their processes.

Institution-wide awareness

Three participants, all of whom worked at repositories located within academic libraries, emphasized the value of a self-audit in fostering institution-wide awareness of digital preservation standards and practices. All these participants noted that conducting a self-audit allowed them to have conversations in their libraries about digital preservation and educate other units about the work that the repository was doing. One participant stated that conducting a self-audit was a move by their repository to get the library to take digital preservation more seriously, saying that “doing the self-audit, that’s where the rubber meets the road.”

Two participants also observed that conducting a self-audit increased the vocabulary for digital preservation, both within their units and in the larger institution.

They noted that having the vocabulary and concepts associated with trustworthy digital repositories made it possible to have more conversations about their services and policies. One of these participants also asserted that it had the effect of professionalizing their digital preservation work.

Relationship with stakeholders

The other three participants, who did not directly mention institutional awareness as a benefit of self-audit, all emphasized the benefits related to relationships with stakeholders. Two of these participants work at data repositories, and the other participant works at a repository that houses a large amount of content for outside partners. For these participants, accountability to partners and other stakeholders was a major benefit of conducting a self-audit. One participant saw it as a way "to be able to assure them that any digital collections that they share with us will be cared for and made accessible to future generations." The two participants from data repositories, both of which had DSA certification, also noted that such a certification could be used as self-promotion to potential partners and granting agencies.

Relationship to repository community

Each participant cited contributions to the larger repository community as a benefit of self-audit. One of the participants characterized the digital preservation community as one that is willing to share a lot, in terms of experience and tools. Several participants have given presentations on their self-audit experiences at conferences, and they have found gratification in helping other repositories that may want to pursue a self-audit. One participant noted "glimmers of interest" in representatives from smaller repositories and archives surrounding the idea of self-audit, and believes that it may become a more common practice for such institutions, as long as there are adequate resources.

Two participants noted the value of making self-audit documentation, including policies, available online for others in the community. One participant said that making their documentation available under a Creative Commons license enabled it to become a "conversation piece" in the repository community, as it incited questions and feedback from other repositories. The other participant went even further, saying, "Everyone should have a published self-audit – why not?" This participant argued that the resource barriers to conducting a self-audit are small compared to undergoing an external audit, and that making self-audits available to the larger community increases accountability and fosters communication about digital preservation.

CONCLUSION

Digital repositories are entering a transitional time, where trustworthy digital repository principles are becoming increasingly widespread and standardized. From the findings of this study, we can draw several interesting conclusions about the state of repository self-audit during this time. These conclusions may help guide future research or work in the field.

Quality of Tools and Standards

When asked about the efficacy of the auditing tools and standards, all participants responded that they were satisfied. TRAC, ISO 16363, and DSA were the most popular tools in the study, and they seem to be helping repositories achieve their goals in the self-audit process. These participants and their repositories also place a great deal of trust in the accuracy and breadth of the tools. All participants surveyed were using these tools to conduct gap assessments, to some extent, implying a trust that the tools outline all the necessary requirements for a trustworthy digital repository. This could be due to the extensive work done to create standards such as ISO 16363, and the community input that was included in its development.

Although considered generally trustworthy by their users, these tools are sometimes being used in flexible ways for self-audit. It is not as important to repository managers to adhere to a particular standard or tool when they are conducting an audit for internal purposes only. Some repositories may use an amalgam of several tools, or they may simply use a standard to guide the kind of documentation they produce.

In some cases, the tools themselves are not enough. Several participants called for supplemental resources to guide repositories conducting self-audits. Several repositories relied on examples of documentation from other repositories to structure their own documentation, or to interpret certain language.

Value of Trust

Another conclusion that we can draw from these findings is that trust from the Designated Community is important to repositories. All repositories that host content for outside partners cited accountability to these partners as a major reason to conduct a self-audit. The definition of "trust" in digital repositories and its relationship to users and standards culture continues to be an important and valid criticism of trustworthy digital repositories (Bak, 2015; Yakel et al., 2013; Donaldson and Conway, 2015); however, the repositories in this study are not approaching self-audit with a one-sided systems-centric view. Rather, self-audit is seen as a way to refocus the work of the repository with the Designated Community in mind.

Value of Institutional Support

A recurring theme in the participants' responses was the importance of institutional buy-in for digital preservation. This was seen as both a prerequisite for and a benefit of conducting a self-audit. It seems that committing to an assessment of digital preservation practices and policies can help increase institutional awareness around the importance of digital preservation in repositories. However, many participants also noted that they would not have been able to conduct self-audits without institutional buy-in and support for the process, because it required an investment of staff time and resources. According to participants, this often requires cultivating awareness about digital preservation and having an advocate within the administration of their institution. Future research could be done on how exactly repositories and other units go about soliciting institutional support for digital preservation. It may also be interesting to investigate perceptions of digital repositories in other areas of the academic library or university, in order to understand how and why repositories must work to gain institutional support.

Repository Community

Although digital repositories may have to work to gain the support and awareness of their colleagues and institutions, the findings of this study indicate that repositories have strong support from the repository community. In addition to the several participants that referenced documentation created by other repositories, there was a widespread sentiment that the digital preservation community, specifically those interested in self-audit and trusted digital repository standards, is growing. Groups like the Digital Curation Interest Group of ALA, IASSIST, IFDO, NDSA, and PASIG, as well as user groups for tools like Archivematica and Islandora all provide spaces for information professionals to come together and discuss their experiences with repository self-audit. Many participants in this study expressed optimism due to the strength and support of these groups, as well as the interest exhibited by their memberships.

Limitations

There were several limitations to this study. It proved difficult to recruit participants through both recruitment notices and snowball sampling. As a result, this study could be characterized as exploratory, providing insight into how some repositories are conducting self-audit, but the results do not necessarily indicate the practices of the entire repository community. A study that included more repositories would perhaps produce results that would better speak to the current status of the greater community.

This study also only included repositories located within universities or academic libraries. This excludes government repositories, museum repositories, and data repositories not affiliated with universities. Other studies might try to include a more diverse sample, drawing from a broader range of repository types.

This study should also be qualified by the fact that it did not only include repositories that were conducting self-audits for internal reasons. The two data repositories in the study had both completed DSA, which is a self-audit that results in external certification. Another study could survey only repositories that are doing self-audits for internal reasons, such as gaining institutional support, conducting gap assessments, or refining their policies.

Now that TRAC has been superseded by ISO 16363, it is probable that more repositories will start using this international standard. Future studies should investigate the uses of ISO 16363, studying how the standard is utilized in different contexts. It will be interesting to see if repositories continue to use ISO 16363 for self-audits, and how and why they conduct them.

REFERENCES

About | Data Seal of Approval. (n.d.). Retrieved September 20, 2016, from http://www.datasealofapproval.org/en/information/about/

Adamick, J., & Reznik-Zellen, R. (2010). Trends in Large-Scale Subject Repositories. *D-Lib Magazine*, *16*(11/12). http://doi.org/10.1045/november2010-adamick

Ambacher, B. I. (2007). Government Archives and the Digital Repository Audit Checklist. *JoDI: Journal of Digital Information*, *8*(2), 7.

Bak, G. (2015). Trusted by whom? TDRs, standards culture and the nature of trust. *Archival Science*. http://doi.org/10.1007/s10502-015-9257-1

Catalogue of Criteria for Trusted Digital Repositories. (2009). nestor Working Group Trusted Repositories - Certification. Retrieved from http://files.dnb.de/nestor/materialien/nestor_mat_08_eng.pdf

Consultative Committee for Space Data Systems (2011). *Recommendation for space data system practices: audit and certification of trustworthy digital repositories, Recommended practice CCSDS 652.0-M-1, Magenta book, Consultative Committee for Space Data Systems*. Retrieved from https://public.ccsds.org/pubs/652x0m1.pdf.

Consultative Committee for Space Data Systems (2012). *Reference model for an Open Archival Information System (OAIS): recommended practice CCSDS 652.0-M-2, CCSDS, Washington, DC*. Retrieved from http://public.ccsds.org/publications/archive/650x0m2.pdf

Data Seal of Approval. (n.d.). Retrieved April 8, 2017, from https://www.datasealofapproval.org/en/

Dearborn, C. C., Barton, A. J., & Harmeyer, N. A. (2014). The Purdue University Research Repository: HUBzero customization for dataset publication and digital preservation. *OCLC Systems & Services: International Digital Library Perspectives*, *30*(1), 15–27. http://doi.org/http://dx.doi.org.libproxy.lib.unc.edu/10.1108/OCLC-07-2013-0022

Dillo, I., & de Leeuw, L. (2015). Ten Years Back, Five Years Forward: The Data Seal of Approval. *International Journal of Digital Curation*, *10*(1), 230–239. http://doi.org/10.2218/ijdc.v10i1.363

Donaldson, D. R., & Conway, P. (2015). User conceptions of trustworthiness for digital archival documents: User Conceptions of Trustworthiness for Digital Archival Documents. *Journal of the Association for Information Science and Technology*, *66*(12), 2427–2444. http://doi.org/10.1002/asi.23330

Downs, R. R., & Chen, R. S. (2010). Self-Assessment of a Long-Term Archive for Interdisciplinary Scientific Data as a Trustworthy Digital Repository. *Journal of Digital Information*, *11*(1). Retrieved from https://journals.tdl.org/jodi/index.php/jodi/article/view/753

DRAMBORA: About. (2015). Retrieved September 20, 2016, from http://www.repositoryaudit.eu/about/

DRAMBORA Interactive: Users. (2015). Retrieved October 24, 2016, from http://www.repositoryaudit.eu/users/

Hank, C., Tibbo, H. R., & Barnes, H. (2007). Building from Trust: Using the RLG/NARA Audit Checklist for Institutional Repository Planning and Deployment. *Archiving Conference*, *2007*(1), 62–66.

Jantz, R., & Giarlo, M. J. (2005). Digital Preservation: Architecture and Technology for Trusted Digital Repositories. *D-Lib Magazine*, *11*(6). http://doi.org/10.1045/june2005-jantz

Johnston, W. (2012). Digital Preservation Initiatives in Ontario: Trusted Digital Repositories and Research Data Repositories. *Partnership: The Canadian Journal of Library and Information Practice and Research*, *7*(2). Retrieved from http://search.proquest.com.libproxy.lib.unc.edu/lisa/docview/1315865022/86077642B7F74840PQ/2

Galletta, A., & Cross, W. E. (2013). *Mastering the Semi-Structured Interview and Beyond: From Research Design to Analysis and Publication*. NYU Press.

Innocenti, P., & Vullo, G. (2009). Assessing the Preservation of Institutional Repositories with DRAMBORA: Case Studies from the University of Glasgow. *Bollettino AIB*, *49*(2), 139–158.

Kelton, K., Fleischmann, K. R., & Wallace, W. A. (2008). Trust in digital information. *Journal of the American Society for Information Science and Technology*, *59*(3), 363–374. http://doi.org/10.1002/asi.20722

Krahmer, A., & Phillips, M. E. (2016). Communicating Organizational Commitment to Long-Term Sustainability through a Trusted Digital Repository Self-Audit. Presented at the IFLA World Library and Information Congress, Columbus, Ohio. Retrieved from http://digital.library.unt.edu/ark:/67531/metadc854117/m2/1/high_res_d/090_krahmer_en.pdf

Lawson, D., & Spies, P. B. (2004). Developing a trusted digital repository: the OCLC experience. *Vine*, *34*(1), 27–32.

Lee, C. A., & Tibbo, H. R. (2007). Digital Curation and Trusted Repositories: Steps Toward Success. *Journal of Digital Information*, *8*(2). Retrieved from https://journals.tdl.org/jodi/index.php/jodi/article/view/229

Li, Y., & Banach, M. (2011). Institutional Repositories and Digital Preservation: Assessing Current Practices at Research Libraries. *D-Lib Magazine*, *17*(5/6). http://doi.org/10.1045/may2011-yuanli

Nicholas, D., Rowlands, I., Watkinson, A., Brown, D., & Jamali, H. R. (2012). Digital repositories ten years on: what do scientific researchers think of them and how do they use them? *Learned Publishing*, *25*(3), 195–206. http://doi.org/10.1087/20120306

OpenDOAR. (2016). Growth of the OpenDOAR Database - Worldwide. Retrieved September 28, 2016, from http://www.opendoar.org/onechart.php?cID=&ctID=&rtID=&clID=&lID=&potID=&rSoftWareName=&search=&groupby=r.rDateAdded&orderby=&charttype=growth&width=600&height=350&caption=Growth%20of%20the%20OpenDOAR%20Database%20-%20Worldwide

Pejsova, P., & Vaska, M. (2012). Audit DRAMBORA for Trustworthy Repositories: A Study Dealing with the Digital Repository of Grey Literature. *The Grey Journal*, *8*(2), 96–105.

Phillips, M. E., Tarver, H., Krahmer, A., Alemneh, D., & Waugh, L. (2015). UNT Libraries: TRAC Conformance Document. UNT Libraries. Retrieved from http://www.library.unt.edu/sites/default/files/documents/digital-libraries-uploads/UNT_Libraries_TRAC_Conformance_Document.pdf

Prieto, A. G. (2009). From conceptual to perceptual reality: trust in digital repositories. *Library Review*, *58*(8), 593–606. http://dx.doi.org.libproxy.lib.unc.edu/10.1108/00242530910987082

RLG-OCLC Working Group on Digital Archive Attributes (2002). *Trusted Digital Repositories: Attributes and Responsibilities*. Retrieved from https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf

RLG-NARA Digital Repository Certification Task Force (2007). *Trustworthy repositories audit & certification: Criteria and checklist.* Retrieved from http://www.crl.edu/PDF/trac.pdf

Rosenthal, D. S. H. (2014). TRAC Audit: Lessons. Retrieved from http://blog.dshr.org/2014/08/trac-audit-lessons.html

Ross, S., & McHugh, A. (2006). The Role of Evidence in Establishing Trust in Repositories. *D-Lib Magazine*, *12*(7/8). http://doi.org/10.1045/july2006-ross

Schmidt, L. M. (2011). Preserving the H-Net Email Lists: A Case Study in Trusted Digital Repository Assessment. The American Archivist, 74(1), 257–296.

Scholars Portal. (2012). ISO16363 Audit Criteria - Trusted Digital Repository Documents. Retrieved April 2, 2017, from https://spotdocs.scholarsportal.info/display/OAIS/ISO16363+Audit+Criteria

Schultz, M. (2010). MetaArchive Cooperative TRAC Audit Checklist. Educopia Institute. Retrieved from http://www.metaarchive.org/sites/metaarchive.org/files/MetaArchive_TRAC_Checklist.pdf

Steinhart, G., Detrich, D., & Green, A. (2009). Establishing Trust in a Chain of Preservation: The TRAC Checklist Applied to a Data Staging Repository (DataStaR). *D-Lib Magazine*, *15*(9/10). http://doi.org/10.1045/september2009-steinhart

Wildemuth, B. M. (2009). *Applications of Social Research Methods to Questions in Information and Library Science*. Westport, Conn: Libraries Unlimited.

Yakel, E., Faniel, I., Kriesberg, A., & Yoon, A. (2013). Trust in Digital Repositories. *International Journal of Digital Curation*, *8*(1), 143–156. http://doi.org/10.2218/ijdc.v8i1.251

Yoon, A. (2015). *Data Reuse and Users' Trust Judgments: Toward Trusted Data Curation* (Doctoral dissertation). University of North Carolina at Chapel Hill. Retrieved from https://cdr.lib.unc.edu/record/uuid:2c2268b3-88cf-4397-b038-b39e88f80d83

APPENDIX A: INVITATION TO PARTICIPATE

Dear [participant],

I am an M.S.I.S. student at the UNC-Chapel Hill School of Information and Library Science, and I am writing my master's paper on self-audit in digital repositories. For my research study, I am interviewing information professionals working in digital repositories to learn about their experiences with self-audit using trusted digital repository standards/tools (i.e. TRAC, ISO 16363, DSA) or other resources.

I am writing to you because [referral source]. I would very much appreciate the opportunity to interview you to learn more about your methods and reasons for self-audit. Would you be willing to be interviewed for my research study?

The interview would take place remotely using the video conferencing service GoToMeeting and would be recorded. The interview should take no longer than an hour. If you are interested in participating, please contact me at [email] with times that work for your schedule. I look forward to hearing from you!

Sincerely,

Hannah Wang

APPENDIX B: PARTICIPANT INFORMATION SHEET

Thank you for taking part in this research study, *Self-Audit in Digital Repositories*. The semi-structured interview for this study will take no longer than one hour. In-person interviews will be recorded using an iPhone; remote interviews will be recorded using the video conferencing service GoToMeeting. Results will be anonymized – the names of the participants and the repositories will not appear in the final paper.

The purpose of this study is to investigate the current use of trusted digital repository standards to conduct self-audit in digital repositories. In particular, these questions will be explored in interviews and subsequent analysis:

- What evidence and tools are digital repositories using to conduct self-audits?
- How is self-auditing being conducted?
- What is the value of this work to the repository and its stakeholders?

If you have any questions pertaining to this study, please contact Hannah Wang at [email] or at [phone number]. You may also contact Helen Tibbo, faculty advisor for this study, at [email].

APPENDIX C: INTERVIEW GUIDE

General Introductions

1. What is the name of your repository/repositories?
2. What is your position in the repository? How long have you held this position?
3. What is the function of your repository within a larger institution?

Self-Audit Introduction

4. Are you currently conducting a self-audit or has it already occurred?
5. Have you conducted more than one self-audit? Do you have plans to?
6. When did you begin the self-auditing process?
7. What made your institution decide to conduct a self-audit?
8. Is your institution considering doing an external audit?

Self-Audit Tools

9. What tools or standards are being used for self-audit?
10. Why were these tools or standards chosen?
11. How have these tools or standards met your needs?
12. Are there other resources that you have found useful during this process (e.g. the use of consultants or colleagues, published reports from other repositories, published guides)?
13. Are there any resources that might make this process smoother?

Method of Self-Audit

14. Who is responsible for conducting the self-audit? If more than one self-audit has been conducted, was the same person responsible for each audit?
15. How is responsibility for the self-audit allocated?
16. Are other departments involved in the process?

Risks and Costs

17. What are the risks and costs associated with the self-audit?
18. Have there been any unexpected risks or costs?

Value of Self-Audit

19. What have been the benefits and opportunities of self-audit to the repository?
20. What kinds of repositories might benefit from this process?

Wrap-Up

21. Do you have any questions or comments?
22. Can you identify any other repositories that have conducted self-audits who might be able to be interviewed for this study?

APPENDIX D: TYPE AND SCOPE OF REPOSITORIES

Repository number	Type of repository	In academic library?	Houses content for outside institutional partners?	Standard(s) used
1	Multiple - special collections and institutional repositories	Yes	Yes	TRAC
2	Special collections repository	Yes	No	TRAC ISO 16363
3	Data repository	No	Yes	DRAMBORA DSA TRAC
4	Multiple - special collections and institutional repositories	Yes	Yes	TRAC ISO 16363
5	Multiple - special collections repositories	Yes	Yes	TRAC
6	Data repository	No	Yes	DSA TRAC ISO 16363