VIRTUAL AUTHENTICITY: AUTHENTICITY OF DIGITAL RECORDS FROM THEORY TO PRACTICE

by

CORINNE ROGERS

MAS, The University of British Columbia, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Library, Archival and Information Studies)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April 2015

Abstract

The assessment and protection of the authenticity of digital records and data are recognized as fundamental issues for the records' current use as well as for their long-term preservation and dissemination. Over the past twenty years, the matter of how to define, determine, and guarantee the endurance of authenticity has been the subject of research in all evidence-based or memory-based disciplines, including archival science, digital humanities, and law. Despite the wealth of past and current research findings, recommendations, and tools, authenticity is still discussed as an urgent problem for records and data created and maintained in traditional digital technologies as well as in emerging ones, such as cloud technologies, and embedded or wearable technologies.

This study investigates contemporary ideas about authenticity of records and data, and practices employed by records professionals. Based on the archival idea that record authenticity is assessed by establishing its identity and proving its integrity, this study identifies indicators for authenticity and categorizes them as either social or technical mechanisms. Using a mixed methods design, it measures how records professionals ensure, manage, and continuously assess record authenticity and to what extent their practices reflect the results of available research. A web-based survey reached records professionals worldwide through professional listservs, and semi-structured interviews gathered further qualitative data from a sample drawn from the survey respondents.

The results show that the standard archival definition of authenticity is not uniformly accepted or implemented in practice, and terms such as authenticity, reliability, integrity, and provenance are often used interchangeably and with little precision. They also reveal that experience plays a major role, in that professionals who are not required to authenticate records in the course of their work tend to have more confidence in technical mechanisms that those who are. The study concludes that most records professionals ensure authenticity by relying on social mechanisms but have greater confidence in technical mechanisms to authenticate records and data. In other words, records professionals, traditionally the trusted agents of record control (trustees), have frequently become the trustors, placing their trust in technology of which they may have little understanding and even less control.

Preface

This dissertation is original, independent work by the author, Corinne Rogers. Some of the results discussed in Chapter 3 were presented at the Second International Conference on Cloud Computing Security (ICCSM 2014) in Reading, UK, October 2014 and the related paper has been accepted for publication in the Canadian Journal of Information Studies. Data collection reported in Chapters 4 and 5 is covered by UBC Ethics Certificate H12-01496.

Table of Contents

List of Tables

List of Figures

Acknowledgements

This dissertation would not have been possible without the encouragement and support of many people. First and foremost I would like to thank my supervisor and friend, Dr. Luciana Duranti, for her inspiration and guidance, and her unflagging confidence in me. I would also like to thank the members of my supervisory committee: Professor Anthony Sheppard for his keen legal insight; Dr. Joseph Tennis for our many conversations about theory and methodology; and Dr. Barbara Endicott-Popovsky (and Dr. Slava Popovsky) for their encouragement and guidance on structure and process.

I would also like to thank all my professors at SLAIS who have instilled in me such respect and passion for archival work throughout the course of the Master of Archival Studies program and the Doctoral program. My fellow students and doctoral colleagues have also been a source of inspiration and support.

My studies would not have been possible without the support of the Social Sciences and Humanities Research Council Joseph-Armand Bombardier Canada Doctoral Scholarship, and the University of British Columbia Four-Year Doctoral Fellowship.

Finally, I want to acknowledge and thank my children, Liam, Colin, and Sean Whelan, for their unfailing support and constant encouragement. A special thanks is due Colin for his expertise and advice in statistical analysis.

Dedication

To my parents, Edward S. Rogers and Jean Hayes Rogers, no longer here but still present.

To my children, Liam, Colin, and Sean

1. Introduction

1.1 Overview

Authenticity is frequently identified by records professionals as a requirement, or a goal, when creating and preserving the records and data upon which modern society relies. Beyond that goal, however, identifying the attributes required to make an attestation of authenticity, or the conditions upon which authenticity can be assessed or presumed, is neither easy nor standardized. This study investigates contemporary ideas about authenticity of records and data, and practices employed in its service by records professionals – those entrusted both with managing current digital information, and preserving information no longer used by its original creator.

This introductory chapter identifies the research problem addressed by this study and situates it in the context of the literature and current research. It states the research questions and outlines the theoretical framework in which they are investigated, and introduces the overall research design. The structure of the dissertation is presented at the end of the chapter.

1.1 Identification of the Research Problem

"There can be few words in the English language as confusing as authentic."

(Duncan 2009, 97)

The concept of authenticity has fascinated philosophers and scholars for millennia. Indeed, authenticity is the subject of much debate today as it regards persons and individual identity, questioning, for example, what is the meaning of an authentic life or how can we know that the person we are engaging on the Internet is in fact the person s/he claims to be. Such a broad concept of authenticity is not the focus of this study, however. This research is concerned with authenticity not of individuals, but of the digital information products of individuals in the context of their work – the residue (or representation) of their activities written, or captured, on digital media, that is, digital records, documents and data, created and preserved in the course of business activity.

Records, defined according to traditional archival theory as documents made or received in the course of activity and set aside for further action or reference, are the raw material of archival research and scholarship (Eastwood 1994, 125; Duranti 1993, 9; Duranti and Michetti 2015). Two aspects of records research that permeate archival discourse are the determination and maintenance of authenticity. In the digital environment, research agendas in the information management communities focus on authenticity as an integral value that must be protected over time and across technological change through digital preservation (joining values of sustainability, accessibility, and understandability), broadening the scope of enquiry beyond records as defined by archival theory to include documents, data, and digital objects of all types.

Traditionally, archives have served society's need to preserve documentary material for the purposes of accountability, evidence, and memory. Whether protecting the patrimonial rights of ecclesiastical bodies, universities, or the ruling elite as in ancient

Rome or medieval Europe, or upholding the accountability of the people's government in modern democracies, archives have been and remain the trusted repositories of countries' heritage. The researcher using material preserved in an archives can do so confident in the trustworthiness of that material – the fact that it is reliable and accurate according to its circumstances of creation and use by its creator, and that its authenticity can be presumed through the chain of its custody from the creator to the trusted preserver. This confidence rested on the provenancial and contextual information available through archival description.

Sir Hilary Jenkinson believed that archival documents (i.e. records) were "authenticated by the fact of their official preservation" (Jenkinson 1937, 4). To Jenkinson, records' history of legitimate custody alone, then, was a sufficient predictor and guarantor of the trustworthiness of the material. However, the relative archival utopia of the pre-World War II era was short-lived as the volume of material destined to enter archives exploded. Michael Cook, writing 50 years later, dismissed Jenkinson's absolute faith in the documentary chain of custody (or perhaps the assumption that such chain of custody can be presumed or demonstrated): "We no longer believe, as Jenkinson did, that an archive's value in research or as legal evidence depends on our certainty that it has never left official custody" (M. Cook 1986, 129). Thus, archival institutions cannot trust the records they intend to acquire solely on the basis of their custodial history, but must test them for indications of their authenticity through studying their provenance and elements of their form (diplomatics) (M. Cook 1986, 7).

Today, digital technology has changed the way we communicate, conduct business, present our public face(s), and document our private lives. As digital communications extend or in some cases supplant print-based culture, a new literacy is evolving. Digital culture is challenging the viability and legitimacy of many well-established social and cultural norms and their associated legal frameworks (Doueihi 2011, 12). One aspect of this evolution can be observed in our concepts of trust in digital information, our re-conception of what it means for a digital object to be authentic, and how we can assess its authenticity. "Virtual authenticity is not to be explained by a transfer of a well-known and ultimately problematic category from one model to another; it is not to be restricted to a shift from the real to the virtual" (Doueihi 2011, 53). Records professionals – archivists and records managers – are embracing the new digital literacy and reexamining traditional concepts upon which their professional identity is based. These concepts concern recorded information – records, documents, data – and the attributes by which we have traditionally assessed authenticity, such as provenance, authorship, identity, and integrity.

The establishment and protection of authenticity of digital materials is generally recognized as a fundamental issue for their current use as well as in the process of their long-term preservation. Over the past twenty years, concerns about authenticity – how to define, assess and guarantee its endurance – have been and continue to be an urgent topic of research in all evidence-based or memory-based disciplines, including archival science, digital humanities, and law (e.g. Duranti and Eastwood 1995; Bearman and Sochats 1996; Duranti and MacNeil 1997; Bearman and Trant 1998; CLIR 2000; Duranti, Eastwood, and MacNeil 2003; Paul 2004; Duranti 2005a; Guercio 2005; Duranti and Preston 2008; Paul 2008; Factor et al. 2009; Giaretta 2011; Salza et al. 2012). Record authenticity is mandated in standards for records management and preservation frameworks such as ISO 15489:2001 and OAIS – Open Archival Information Systems (ISO 2001; CCSDS 2012), and the protection of authenticity is required by professional codes of conduct (cf. SAA 2011; ACA 1999). Although theoretical frameworks have been proposed for the protection of authenticity – some of them highly influential in practice, for example in the development of the DoD 5015.2 standard for records management systems (Duranti and MacNeil 1997) and in the development of Italian legislation (this research, Interview subject D073)), despite the wealth of past and current research, establishing and guaranteeing authenticity are still discussed as urgent problems yet to be solved. The authenticity issue exists for records and data created and maintained in traditional digital technologies as well as in emerging technologies, such as cloud technologies, and embedded or wearable technologies.

Much theoretical and practical research focuses on how to preserve digital records and data for purposes of accessibility and usability, to protect their evidentiary capacity and research value. There is little empirical research, however, that investigates how the findings of research on authenticity are integrated into existing practice by records professionals working with digital material at all stages of the life cycle, or how, exactly, records professionals ensure authenticity. Workplace (or work practice) studies is a subfield of sociology that investigates practical aspects of work, integrating the roles of technology, and social processes (Trace 2011). Several important ethnographic studies of workplace practice regarding recordkeeping and accountability have been reported in the archival literature (e.g. Yakel 1997; Shankar 2004), however the focus on authenticity has been indirect. The relationship among ICTs (Information and Communications Technologies), record authenticity, and accountability has been examined in case studies in public administrations, and has distinguished the role of technical and organizational safeguards for records preservation (Meijer 2003). Practitioner behavior has been studied with respect to establishing record authenticity by Eun Park in the late 1990s, and her work was published in a pilot study and subsequently in her dissertation research (E. Park 2001; E. Park 2002b; E. Park 2002a). The focus of her research was on the authenticity requirements and authentication processes in student record systems.

1.2 Hypothesis

Anecdotal evidence and case studies from previous research projects in which this author has participated (www.interpares.org, www.digitalrecordsforensics.org, www.interparestrust.org) suggest an apparent disconnect between the findings of major research initiatives on the means of establishing and protecting authenticity and the practice of many records professionals, which has not been thoroughly investigated. This study addresses such gap by exploring the relationship between the findings of major research projects on the processes of creating, maintaining, and preserving trustworthy digital material, and the practices, experience, and beliefs of records professionals.

The hypothesis of this study is that, despite clear guidance offered by archival science on the means of ensuring, managing, and continuously assessing record authenticity, a guidance reflected in the products of several large-scale, significant and influential research projects on the topic of authenticity in the context of long-term preservation, the theoretical results of these projects are not being consistently applied in practice, and in fact records professionals are often unclear about how to define authenticity, how to protect it, and how to assess it (authenticate records).

1.3 Research Questions

This study investigates the degree to which records professionals address issues of authenticity explicitly in their work with digital material, and the means by which they do so. The primary purpose is to further our understanding of how these professionals, primarily but not exclusively archivists and records managers[1], think about authenticity of the digital material for which they are responsible, and of what techniques or indicators of authenticity they use and rely on to ensure or continuously assess it. In order to do so it will explore the extent to which traditional archival models of authenticity are still employed in the digital environment and what new techniques or models may be developing. In other words, the purpose of this study is to understand what the situation *is*, and compare the answer with the theoretical expectations of *what should be*.

The research questions were generated from an identification of perceived gaps in the literature (see Literature Review), and the experience of this author in several research projects investigating digital evidence, specifically the Digital Records Forensics Project (www.digitalrecordsforensics.org), the Digital Economy Project – *The Canadian legal*

[1] The distinction, perceived or real, between records managers and archivists, has been much discussed, and different national traditions view them as parts of one unified profession or two distinct ones. In formulating the research questions, this author focused on the practice of records professional broadly – that is, any professional concerned with the nature of records and the establishment of their authenticity. Professional identity was treated as an independent variable in the course of data collection and analysis, as will be discussed in Chapter 3.

framework for evidence and the Digital Economy: a disjunction? (Sheppard and Duranti 2010), InterPARES 3 (www.interpares.org), and the current InterPARES Trust Project (www.interparestrust.org).

This study asks three broad research questions:

Research Question 1: What elements of the context, content, and structure of digital records and data do records professionals use and rely on in order to determine and manage authenticity?

Research question 2: Is the traditional model of authenticity of records used in the digital environment and if so, to what degree?

Research question 3: Is the traditional model of authenticity sufficient to support a presumption of authenticity in the digital environment over time and across technological change?

To answer these questions, seven detailed sub-questions will shape the research:

> *Sub-question 1: What are the domain definitions of, and relationships between terms such as authenticity, identity, integrity, authentication, provenance, lineage, traceability, originality as they relate to how records professionals view authenticity?*
>
> *Sub-question 2: How do records professionals approach the issue of authenticity of digital records and data in their work?*

Sub-question 3: What elements or indicators of authenticity do records professionals rely on or believe are important?

Sub-question 4: Does experience with authenticating records affect what indicators records professionals rely on or believe are important?

Sub-question 5: What empirical evidence exists of the challenges of presuming or assessing and verifying authenticity of digital material?

Sub-question 6: How might this empirical evidence enhance a practical, theoretical, or philosophical understanding of record authenticity?

Sub-question 7: What further research is indicated?

Through these questions the convergence of theory and practice in the matter of establishing, maintaining, and assessing authenticity of digital records and data can be studied.

1.4 Theoretical Foundations

This study is grounded first and foremost in the theory and methodology of archival science and digital diplomatics, which provide a model of “record”, and a means of understanding and defining record authenticity as well as the elements that comprise it (Jenkinson 1937; Eastwood 1994; Duranti and MacNeil 1997; Duranti 1998a). This model traces its origins to 17th century diplomatics, if not archival practice dating back as far as the Roman Empire and documented in the Justinian Code (Duranti 1998a, 36–40).

It was further developed and tested for the digital environment through the research of the InterPARES Project (MacNeil and Gilliland-Swetland 2005; Duranti 2005a; Duranti and Thibodeau 2006; Duranti and Preston 2008). The model is viewed in the context of a pragmatist worldview (Hookway 2015; Creswell 2009, 10–11; Feilzer 2010).

As mentioned above, the findings of research into the creation, maintenance, and preservation of authentic digital records and data do not appear to have found universal acceptance or implementation among records professionals. Authenticity itself is not uniformly defined, even if it is almost always sought. If it were, then the results of these research projects could be directly applied and no one would question the means of assessing whether a digital record or data set was authentic or not. This is not the case. Current social theories suggest that authenticity, a component of trustworthiness, is socially constructed and contextually evaluated (MacNeil and Mak 2007). In discussing the findings of the survey questionnaire and interviews, this author applies a practice lens influenced by practice theorists such as Theodore Schatzki (Schatzki, Knorr-Cetina, and Savigny 2001; Feldman and Orlikowski 2011) and actor-network theorists (Law 1992; Latour 2005; Law 2006). This author also draws on the distinction between 'hard' and 'soft' systems, as developed by Peter Checkland and his colleagues in the1960s (1999; 2007) to discuss the role of indicators of authenticity. This distinction has been used by Sztompka (1999, 4–5, 58–59) in his sociological theory of trust, and Foscarini (2009; 2010) in her investigation of the role of functional classification in modern organizations.

1.5 Structure of the Study

Chapter 2 presents a review of the literature of record authenticity that supports the theoretical foundation of the study, presents past and current research, and outlines epistemological discourses that inform new currents of thought. The literature review considers the concept of authenticity as a component of trust and trustworthiness and examines authenticity research in disciplines which share a reliance on documentary material for purposes of research, evidence, or lasting memory. The focus of the literature reviewed is on archival science and records management, but the review also looks at closely allied fields, ranging from digital forensics to information technology and law.

Chapter 3 outlines the research design, providing the theoretical foundation of the study, a rationale for the methodological approach, and a description of the methods of data collection and analysis.

Chapter 4 and 5 present the results of a web-based survey and semi-structured interviews respectively conducted by the author for this study among records professionals worldwide. The survey collected data about professional practices – what tasks are conducted most frequently and what indicators of authenticity are most relied on in the management and preservation of records and data and in the process of their authentication. These data measure variables of practice and belief, and are analyzed quantitatively. Interviews conducted as a follow-up to the survey explore its results qualitatively and in more depth.

The final chapter presents a discussion of the findings, conclusions, and suggestions for further research.

2 Literature Review

2.1 Introduction

The literature review outlines the concept of record authenticity in traditional archival and diplomatic theory and its development in contemporary archival diplomatics, laying a foundation for a discussion of the main focus of this review, that is, the literature concerning issues of authenticity of digital records and data. It presents an overview of research that has addressed the issue of digital record authenticity, either explicitly or implicitly, and provides a framework that contextualizes and substantiates the research problem. The main body of literature considered is the English-language or English-translation corpus that is the foundation of the European, North American, and Australian archival discipline, as it relates to issues of authenticity in the creation, management, use, and preservation of records and data (regardless of medium).[2] This is supported by select literature addressing concepts of documentary authenticity in the complementary disciplines of digital forensics, information technology, and law.

Scholarly writing and research regarding digital record authenticity in archival and related information disciplines parallel broad developments in information and communications technology (ICT). The development of this technology has been

[2] The term 'archival discipline' used here includes management of current records by their creator (the records management literature) as well as ongoing use and preservation of records used also by persons or organizations other than their creator. For a discussion of the historical roots of the archival and records management disciplines, and the similarities and differences between records managers and archivists, particularly with respect to digital records, see Charles Dollar, "Archivists and Records Managers in the Information Age," (Dollar 1993) and Luciana Duranti, "The Odyssey of Records Management," Parts 1 and 2 (Duranti 1998b; Duranti 1998c).

described by International Data Corporation[3] (IDC) as happening in three phases: the early digital environment, or First Platform (first generation computer platform – the mainframe environment); the Second Platform, or client-server computing architecture; and the emerging and disruptive Third Platform, that is cloud-based, embraces the semantic Web, is accessible from mobile devices, and utilizes Big Data (EMC 2013; Golden 2014). These phases are not mutually exclusive, nor are they entirely sequential, and in fact they display significant overlap. Each phase of ICT development has brought with it specific and ongoing challenges to the creation, management, and preservation of digital records and data.

Digital technology has upset the traditional systems of control that have ensured the creation of reliable records, and the means of presuming their continued authenticity over time and across technological change (Lauriault et al. 2007, 140; MacNeil and Gilliland-Swetland 2005, 21). Digital records differ significantly from paper records. Records, documents, and data created and stored on digital media are volatile and subject to loss, intentional or unintentional alteration, contamination, or corruption, even when they are still in the custody of their creator. Their authorship, provenance, or chain of custody may be difficult or impossible to determine. They may be transmitted, shared, and copied with ease. Their accessibility is subject to hardware and software obsolescence and incompatibility. Even if the creator relies on a digital record in the course of business, and maintains its unbroken chain of custody, the fragility and vulnerability of digital records demands explicit action to protect the record's authenticity. Furthermore,

[3] International Data Corporation (IDC) is a prominent global provider of intelligence for the information technology, telecommunications and consumer technology markets. www.idc.com.

reliability and accuracy are no longer directly linked to authenticity and may be compromised together or separately (Duranti 2005b, 1; Duranti and MacNeil 1997, 48; Duranti and Thibodeau 2006, 54; MacNeil and Gilliland-Swetland 2005, 21).

Records, defined as documents made or received in the course of practical activity and set aside for further action or reference, are the raw material of archival research and scholarship (Duranti 1993, 9; Duranti and Michetti 2015; Eastwood 1994, 125). In the digital environment, research agendas in information management communities focus on authenticity as an integral value that must be protected over time and across technological change through digital preservation (joining values of sustainability, accessibility, and understandability), broadening the scope of enquiry beyond records as rigorously defined by archival theory to documents, data, and digital objects of all types. Although concern for authenticity motivates much research into the nature, communication, and preservation of digital objects, establishing a concrete, testable definition of digital record has proven extremely difficult.

2.2 Theoretical Foundations

2.2.1 Defining Documentary Authenticity

The concept of documentary authenticity has ancient roots. The word authenticity derives from the Anglo-Norman, Old and Middle French, with reference to a thing (as a noun, *authenticum*, originally and frequently a legal document), or a person (as an adjective, denoting trustworthy, credible, genuine, or legally or duly qualified). Its etymon is the Latin *authenticus,* referring to documents (2nd century a.d.), persons (3rd century a.d.), and later coming to mean some*thing* or someone who is authoritative (from 8th century in British sources), or a thing that is legally valid (12th century). In Hellenistic Greek, *αὐθεντικός* meant warranted, original, authoritative (Oxford English Dictionary 2014).

Black's Law Dictionary defines 'authentic' as "Genuine; true; having the character and authority of an original; duly vested with all necessary formalities and legally attested; competent, credible, and reliable as evidence" (Black's Law Dictionary 2012).

With respect to documents only, the Oxford English Dictionary provides the following explanation of authenticity:

1. Esp. of a document: that is the origin or source of something; original, primary; not a copy. *Obs.*
    a. Of a document, artifact, artwork, etc.: having the stated or reputed origin, provenance, or creator; not a fake or forgery,
    b. Presenting the characteristics of the original; accurately reproducing a model or prototype; made or done in the original or traditional way.
2. A document of which another is a copy or transcript; an original document. *Obs.* (Oxford English Dictionary 2014).

Documentary authenticity is related variously in these definitions to originality, source, and authority.

The OED does not distinguish 'record' from 'document'. With respect to archival and records management theory, this is a crucial distinction. According to archival theory, a document is recorded information, while a record is a document made or received in the course of practical activity and set aside for future action or reference. A record is a special type of document. The definition of record authenticity, elegant in its simplicity but challenging to apply, that is at the root of archival, diplomatic, and legal theory, and has been codified in records management standards, holds that authenticity is "the trustworthiness of a record as a record, i.e. the quality of a record that is what it purports to be and that is free from tampering or corruption" (InterPARES Glossary, www.interpares.org). The Society of American Archivists defines 'authenticity' as: "The quality of being genuine, not a counterfeit, and free from tampering, and is typically inferred from internal and external evidence, including its physical characteristics, structure, content, and context." Authenticity does not automatically imply reliability of the content of the record (Pearce-Moses 2005 np; Duranti 1998a, 46). The idea of record authenticity is codified in ISO 15489, the international records management standard as follows:

> "An authentic record is one that can be proven
>
> a. to be what it purports to be,
> b. to have been created or sent by the person purported to have created or sent it, and
> c. to have been created or sent at the time purported" (ISO 2001 section 7.2.2).

These concepts are explicated in domain distinctions among history, jurisprudence, and diplomatics – namely, the objects of authenticity inquiry (primary source material, works of art, legal instruments, records, etc.) that are important to historians, legal professionals, and archivists. For the purposes of understanding and analyzing documents and records, Duranti has differentiated among three types of authenticity: diplomatic, legal, and historical. Each is distinct and independent from the other, and reflects a specific purpose or focus of trust and the trust relationship in its discipline. These distinctions do not coincide, nor are they clearly enough defined in practice (or their definitions generally accepted) to offer the possibility of a general theory of authenticity. Duranti explains:

> Legally authentic documents are those which bear witness on their own because of the intervention, during or after their creation, of a representative of a public authority guaranteeing their genuineness. Diplomatically authentic documents are those which were written according to the practice of the time and place indicated in the text, and signed with the name(s) of the person(s) competent to create them. Historically authentic documents are those which attest to events that actually took place or to information that is true (Duranti 1998a, 45–46).

From this we see that what constitutes an authentic document is conditioned by the discipline in which it is considered – and therefore the purpose the document serves. In the digital environment, finding a common understanding of "the multiple meanings and significance of authenticity" remains critical (CLIR 2000, vii), and yet continues to be elusive.

2.2.2 Traditional Archival Theory

The roots of archival theory and concepts of record authenticity are anchored in legal and administrative principles, first executed in centralized public repositories of written documents, then, with the spread of literacy, expanding into the regulated recordkeeping practices of public and private organizations, administrations, and homes (Eastwood 1994, 125; Duranti 1998c). Law and jurisprudence are the original pillars of influence that have guided the history and development of archival theory, reaching back through the centuries to Roman times. Principles from Roman law that have become part of the foundation of archival knowledge include the idea that antiquity provides records with the highest legal authority, that deposit in a public place guarantees reliability of records as witnesses of actions, and that an unbroken chain of custody ensures records' continuing authenticity (Duranti 1996a, 1). The theory of the nature of archival material derives from the analysis of the relationship between records and their producing body, that body's functions and activities, and the rights and duties of the people interacting with it – related to the theory of the state at the time, designed to accomplish the purposes of the state (Duranti 1996a, 3). Early modern archival discourse was thus cradled in the public and state archives of Europe, articulated in the influential writings of practitioners such as the Dutch trio, Muller, Feith and Fruin, and the seminal works of English theorist Sir Hilary Jenkinson. The evidentiary capacity of records was at the core of these theories, shaping archivists' understanding of authenticity and their role in protecting probative value. Archival theory and legal notions of documentary evidence remain intertwined to this day.

Archival practice was not concerned originally with the need to establish or prove explicitly records' authenticity. Rather, authenticity was an intrinsic characteristic of records, a quality of their archival nature resulting from the circumstances of their creation, maintenance, and preservation. In his seminal work, *Manual for Archives Administration*, Sir Hilary Jenkinson noted "two common features [of records] of extraordinary value and importance" upon which "they can be analyzed and tested," namely impartiality and authenticity (Jenkinson 1937, 12). These derive from their creation (records are "drawn up and used in the course of an administrative or executive transaction (whether public or private) of which [they] formed a part") and maintenance ("and subsequently preserved in their own custody for their own information by a person or persons responsible for that transaction and their legitimate successors") (Jenkinson 1937, 11). The contingencies that endow authenticity "are observable not in the document itself but in the procedures" of creation, maintenance, and preservation (Eastwood 1994, 127). While the validity of Jenkinson's theory of the inherent characteristics of archives has been vigorously debated and has been rejected by many contemporary writers (e.g. Cook 1997; Cook 2001; McKemmish 2001; Nesmith 2002), it remains a valuable link in understanding the development of archival notions of authenticity. Regardless of critiques of his ideas, Jenkinson's "spirited defence of the evidential character of records certainly remains inspirational to archivists everywhere" (Cook 1997, 25), and according to Duranti, protection of record authenticity, his "moral defence of archives," (Jenkinson 1937, 83) remains a primary function of the archivist (Duranti 1996b, 518).

2.2.3 Diplomatics

The science of diplomatics was developed in the 17th and 18th centuries to prove the authenticity, and indirectly, the reliability, of archival documents, in order to establish the existence of patrimonial rights of the church and its religious orders and other authorities, and to identify and eliminate forgeries. Diplomatic authenticity is concerned with proving that a document is what it purports to be through the study of its genesis, forms, and transmission, and its relationships with actions and persons and with its juridical context. At the core of diplomatic theory is the understanding that all records can be analyzed, understood and evaluated in terms of a system of formal elements that are universal in application and decontextualized in nature. A record is understood conceptually as a system of internal and external elements consisting of acts (determinant cause of record creation), persons (those who concur in record formation), and form, binding all together (Duranti 1997).

In classic diplomatics, trustworthiness equates with authenticity, which implies a presumption of reliability, accuracy, and legitimacy, and therefore genuineness. This inference was possible because of the highly controlled process of creation, maintenance, and preservation of ancient documents that were the subject of study of the early diplomatists. By establishing the identity of the document, its integrity was presumed. With digital records, identity and integrity are no longer linked. Diplomatics developed into a "very sophisticated system of ideas about the nature of records, their genesis and composition, their relationships with the actions and persons connected to them, and with their organizational, social, and legal context" (Duranti and Eastwood 1995, 215).

Diplomatic criticism has evolved to analyze and evaluate individual documents in terms of this system of formal elements, through which those documents can be shown to have been "written according to the practice of the time and place indicated in the text, and signed with the name(s) of the person(s) competent to create them" (Duranti, 1998, p. 46). Authenticity is evaluated by establishing the document's identity and integrity. Modern diplomatics establishes the trustworthiness of a record in terms of the three elements of trustworthiness – reliability, accuracy, and authenticity, but cannot infer from that truthfulness or legitimacy.

It studies the genesis, forms and status of transmission of individual documents in order to establish their authenticity. In research into records in digital environments, diplomatics has supported the identification of necessary and sufficient attributes of digital records. In contrast, archival theory provides information about records in their aggregations and their documentary and functional relationships. Between 1989 and 1992, Duranti published a series of articles that explained the principles of classic diplomatics and applied and adapted them to records of modern bureaucracies, extending them beyond traditional analogue records into the realm of digital records.[4] By integrating the principles and concepts of diplomatics with those of archival science, Duranti developed a conceptual model of an authentic record, regardless of medium, that is rooted in jurisprudence, administrative history, and theory (Duranti 1998a; Duranti and MacNeil 1997; Duranti 2001, 43). Archival diplomatics, used both retrospectively (to

[4] Six articles, entitled *Diplomatics: New Uses for an Old Science* (Parts I-VI) were published in *Archivaria* over the course of six issues, providing the most comprehensive examination of diplomatics available to English-speaking audiences. In 1998, the articles were published as a book of the same title (Duranti 1989a; 1989b; 1990a; 1990b; 1991a; 1991b; 1998a).

understand the nature and attributes of existing records and to assess their trustworthiness) and prospectively (to design documentary forms and procedures and to develop trusted record-making, recordkeeping and record preservation systems), has provided the theoretical foundation of two decades of research into issues of reliability and authenticity of digital records. (cf. Duranti and MacNeil 1997; Duranti, Eastwood, and MacNeil 2003; Duranti 2005a; Duranti and Preston 2008).

2.3 Early Archival Concerns: Before 1990

Our familiarity and comfort with assessing the authenticity of traditional records stems from our ability to see, touch, and hold them. Archival practices and diplomatic methods developed over many centuries of handling records on stable physical media. Similarly, common and statute law governing treatment of documentary evidence centered on the physicality of documents and the stability of the media upon which they were inscribed. In the digital world, we do not see a physical document, but a display of assembled digital components – streams of bits ordered by sets of rules interacting at different levels of the technology (operating system, transport protocols, software applications, etc.) written in a languages humans cannot directly read or understand.

The National Archives and Records Administration (NARA) accepted its first electronic records from U.S. federal agencies in 1969, mainly flat database files and ASCII records. Authenticity of these electronic records was ascertained through visual inspection of printouts (NARA 2015). In 1973 the Public Archives of Canada established a Machine

Readable Archives Division[5], following in the footsteps of the United States and Sweden. During its first three years it developed methods and standards to meet the Archives' mandate of appraisal and acquisition, processing, conservation, and public service (Naugler 1978). It was not until 1978 that Charles Dollar called for continuing retention of electronic records, evaluated, or appraised, by a dual process of technical and intellectual considerations. Dollar considered such records to have informational value only, with no legal or business value, thus distinguishing these electronic records from traditional records in a creator's fonds (Dollar 1978). This position was challenged in 1981 by the Public Archives of Canada, which called for computer-generated records to be appraised in the context of the whole of a creator's records and on the basis of the same taxonomy of values as paper records. This position subsequently gained international acceptance within the archival community following publication of the UNESCO Records and Archives Management Programme (RAMP) study authored by Harold Naugler in1984 (Duranti 1996b, 494). This study provides an overview of appraisal issues and offers guidelines for establishing preservation programs for electronic records. It notes the lack of legislative support, restrictions on transfer to archives, and the lack of programs for identifying, inventorying, and scheduling these

[5] Early literature distinguished traditional paper records from "machine readable" records – those records whose form could be recognized, accepted, and interpreted by a machine, such as a computer or other data processing device, analog and digital. The term, "machine-readable," encompassed a wide variety of storage media, including punched paper cards, magnetic discs, cassettes, paper tape, and magnetic tape (Dollar 1978, 423–430). As storage media evolved, the term "machine-readable record" gave way to "electronic record," a generic term defined as "an analogue or digital record that is carried by an electrical conductor and requires the use of electronic equipment to be intelligible by a person" (InterPARES 2012). When talking about records created and/or stored in digital computers, the term "electronic record" has gradually been replaced by the more accurate term "digital record," defined separately as "a digitally-encoded object and the metadata necessary to order, structure or manifest the object's content and form," where "digital object" is taken to mean "a discrete aggregation of one or more bit streams and the metadata about the properties of the object and, if applicable, methods of performing operations on the object" (InterPARES 2012).

records that makes their systematic acquisition difficult, if not impossible (Naugler 1984). The issue of appraisal was at the forefront of archival writing in this period; however, despite the challenges to the appraiser presented by issues of authenticity, nowhere did this literature "concern itself with the authenticity of electronic records" (Duranti 2002, 1).

As archivists grappled with the issues of value and application of appraisal criteria to electronic records, the legal status of electronic records and the circumstances of their admissibility was also a subject of intense debate. In common law countries, case law responded slowly to the increasing use of computer records, and legislation continued to adapt to reflect the new reality. Perhaps the highest profile and most influential case for archival issues concerning electronic records was *Armstrong v. the Executive Office of the President,* commonly known as the PROFS case, in 1989 (MacNeil 2000b, 77–79; Bearman 1993). This case raised issues concerning the essential characteristics of electronic records and the verification of their authenticity and determination of their reliability. As a result of the PROFS case, "judicial officers, administrators, systems designers, records keepers and researchers are reviewing their practices and the assumptions behind them, and searching for a) criteria that would allow them to determine when electronic records can serve as reliable evidence of action and decision, for b) techniques that would allow them to preserve such evidence intact, and for c) methods that would allow them to verify and prove its authenticity" (Duranti and Eastwood 1995, 213). This case served as a catalyst for several research projects into issues of creation, maintenance, and preservation of electronic records, including the nature of electronic records themselves, and their reliability and authenticity. The research projects that inform the theoretical perspective of this study, the UBC-MAS Project and the InterPARES Projects, are discussed below.

2.4 Authenticity of Digital Records: 1990 and Beyond

2.4.1 Reports and Position Papers: International Council on Archives

The authenticity of digital records developed as a critical issue in the early 1990s (cf. Duranti and Eastwood 1995; Duff 1996; Duranti and MacNeil 1997; Bearman and Trant 1998). In 1993 The International Council on Archives (ICA) Committee on Electronic Records began developing a series of products, the goal of which was to "undertake study and research, promote the exchange of experience and draft standards and directives concerning the creation and archival processing of electronic records". Three Studies resulted from this initiative: *Electronic Records Programs: Report on the 1994/95 Survey***,** *Electronic Records Management: A Literature Review,* and *Guide for Managing Electronic Records from and Archival Perspective* (Committee on Electronic Records 1997 Preface, p. 3).

Electronic Records Management: A Literature Review provided an "exhaustive review of the international literature on electronic records" and formed the foundation of the subsequent *Guide for Managing Electronic Records from an Archival Perspective* (Committee on Electronic Records 1997 Preface, p. 3). The *Literature Review* covered "the latest thinking and theories of leading experts in the management of electronic records" (Erlandsson 1997, 12), predominantly from 1992-1996, including an extensive discussion of the issues of reliability and authenticity of digital records as they were addressed in two important research projects, the Pittsburgh Project, and the UBC-MAS Project.

The *Guide* describes the implications of electronic records management for archives from the legal, organizational, human resources and technological perspectives, and proposes strategies for operationalizing this work. Among its findings were recommendations that the archives be involved in the entire life cycle of electronic systems in which records are made or received and retained and "ensure that records creators create and retain records which are authentic, reliable, and preservable" (Committee on Electronic Records 1997, 8). The Guide adopts the position that an organization's main purpose in creating and keeping records is to provide evidence of activities and transactions, to which end electronic records must be created reliable and preserved authentic. These twin concepts – reliability and authenticity – are the foundation of accountability.

> The reliability of a record is its ability to serve as reliable evidence. Basically, a record can be no more reliable than it was at the instant of its creation. Therefore, direct responsibility for reliable records is that of the records creator. However, the archives should inform and guide creators on best practices for producing reliable records. Authenticity refers to the persistence over time of the original characteristics of the record with respect to context, structure and content. An authentic record is one that retains its original reliability (Committee on Electronic Records 1997, 24).

The protection of authenticity depends on maintaining intellectual and technological control over the records. Intellectual control entails describing them according to archival standards and including contextual information "sufficient to define the provenance, context, and structure of the records whenever they are not explicit in the records themselves." Technological control must be exercised over any migration or transformation of the records. (1997, 33–34).

At the XIVth International Congress on Archives in Seville, Spain in 2000, the ICA recognized the importance of preserving authentic electronic records and called upon National Archivists to provide leadership. In 2001 the ICA established a working group within the Committee on Archival Legal Matters to prepare a report identifying "the issues that archivists and records keepers must keep in mind to ensure the authenticity of electronic records" (ICA Committee on Archival Legal Matters 2002, 4). The working group consulted the Committee on Electronic Records, and published its report in 2002. The report does not provide recommendations, but gives a snapshot of practice, legislative work and research at that moment, noting a lack of cooperation or links between legislator and researcher (ICA Committee on Archival Legal Matters 2002, 10).

The report adopts a position of jurisdictional neutrality, and embraces the definition of record authenticity put forward in the international records management standard, ISO-15489-1 (International Standards Organization 2001). The requirement for authenticity is linked to four reasons for creating archives: to prove legal rights, to serve as instruments for the administration of an organization, and to serve as cultural heritage and as one of the preconditions for social and political accountability. Authentic documents are “reliable not only at the moment when they are created but remain reliable for a long time to come” (ICA Committee on Archival Legal Matters 2002, 6).

The working group surveyed the Committee on Archival Legal Matters to determine the state of awareness of record authenticity in the archival profession and in national legislation on records/electronic records. It found that there was no uniformity across the profession with respect to the understanding of the concepts of authenticity, reliability, and validity. Generally, authenticity was understood to mean that a record is what it purports to be, and reliability was equated with trustworthiness. National legislation in the countries of the respondents did not define the terms in question with any consistency. Authenticity was mentioned most frequently in laws relating to evidence, electronic signatures, and e-commerce.

The report concludes that the preservation of authentic electronic records should be a critical priority for records professionals, and offers several observations, including that: 1) archivists and archival institutions are still “in the paper age” with too little understanding or concern for issues of authenticity of electronic records; 2) lack of agreement about terminology, and careless use of terminology leads to confusion; and 3) guidelines on preserving authentic electronic records are needed. It calls on UNESCO to take the initiative in promoting training programs, surveying “the world on the status of the authenticity of electronic records,” promoting agreement on terminology, supporting the development of guidelines on preservation, convening a conference of high ranking world government representatives, and developing criteria and models for preservation of digital cultural heritage (ICA Committee on Archival Legal Matters 2002, 10–11).

In 2004, a second report prepared for UNESCO and the ICA was published "to address the global status of authenticity of electronic records, with particular attention to developing countries." The central question asked was "what measures are necessary for records and archives professionals, especially in developing countries, to ensure the authenticity of electronic records…" (Millar 2004, 4). Challenges to authenticity were presented as recurring themes, including the low profile of record keeping, the focus on IT-oriented approaches to creation, management, and preservation of electronic records, the absence of technical or operational standards for management of electronic records, the absence of sustained educational initiatives, and the need for a strategic approach to capacity building (Millar 2004, 8). The eleven recommendations resulting from the consultative exercises that addressed that question were not detailed with respect to ensuring authenticity of records (in contrast with the specific recommendations and guidelines offered by research projects such as InterPARES), but high level strategic priorities and actions for UNESCO, the profession, and the ICA to undertake in response to the identified challenges.

2.4.2 The Council on Library and Information Resources

The Council on Library and Information Resources (CLIR) published a set of position papers in May 2000 by experts from different domains of the information resources community. The papers addressed the question: What is an authentic digital object? In the introduction to the collection the authors recognized that "authenticity" in recorded information connotes "precise, yet disparate, things in different contexts and communities." The goal of the report was to bring together different communities of practice to arrive at a common understanding of key concepts and terms regarding authenticity. This involved exploring the "meaning and significance of content, fixity, consistency of reference, provenance, and context." The report published the perspectives on authenticity of five professionals: a digital librarian, a documentary editor, a special collections librarian, a document theorist, and a computer scientist, asking each to address the nature of a digital object from his/her perspective (CLIR 2000, vi). The view closest to that of an archivist is outlined below.

Clifford Lynch, in his contribution to the CLIR report, distinguished philosophical (social) and computational (technological) constructs in determining authenticity and integrity. According to Lynch, distrust of the digital is forcing exactitude on concepts of authenticity and integrity, yet the result is abstract and elusive, defying testable definitions. Furthermore, distrust of the digital environment appears to be balanced by faith and optimism about the potential for technological solutions – the "magical arsenal [that] has solved the problems of certifying authorship and integrity" (Lynch 2000, 33). Lynch highlights the role of integrity in the determination of authenticity in the digital environment, something that this author has found to be a pervasive theme. "It is an interesting, and possibly surprising, conclusion" claims Lynch "that in the digital environment, tests of integrity can be viewed as just special cases and byproducts of evaluations of authenticity" (Lynch 2000, 41).

2.4.3 The University of Pittsburgh Project

Richard Cox led the project at the University of Pittsburgh that developed and tested a set of functional requirements for recordkeeping in electronic environments. The project focused on recorded transactions providing evidence. Undertaken over a three-year period, the project produced a set of 19 functional requirements for electronic evidence based on literary warrant. While the project did not address authenticity explicitly, the requirements supported trustworthiness and accountability. Requirements included compliance with legal and administrative obligations, retention rules, documented policies and procedures, evidence that records were created in the ordinary course of business, evidence of the archival bond, fixity, unique identity, accuracy and quality control, provenance information, evidence of integrity and auditability (Duff 1996; Erlandsson 1997).

2.4.4 The Preservation of the Integrity of Electronic Records – UBC-MAS Project

A very different approach to that of the consultative reports discussed above was taken by researchers at the University of British Columbia. The Preservation of the Integrity of Electronic Records was a three-year research project (April 1994-March 1997) carried out at the University of British Columbia, funded by the Social Sciences and Humanities Research Council of Canada (SSHRC), under the direction of Principal Investigator, Luciana Duranti and Co-Investigator, Terry Eastwood, and with the support of Research Assistant, Heather MacNeil.[6] One of the project's strengths was its focus on identifying and defining the byproducts of information systems and protecting the integrity of records (those byproducts which constitute evidence of actions) in those systems on purely theoretical grounds. This distinguished it from other projects whose research foci fell within specific legal or programmatic frameworks. The premise was that the

[6] The results of the Project are available at the Project website, http://www.interpares.org/UBCProject/intro.htm#BIBLIOGRAPHY.

identification of the criteria, techniques, and methods needed to solve the problems posed by the use of electronic information systems for carrying out business “cannot derive from purely pragmatic or ad hoc decisions but must be rooted in principles and concepts that can be applied in different situations and various contexts” (Duranti and Eastwood 1995, 214). The theoretical foundation was provided by principles of diplomatics integrated with principles of archival science and interpreted within the framework of electronic systems (Duranti and MacNeil 1997, 47).

The objectives of the project were:

- to establish what a record is in principle and how it can be recognized in an electronic environment;
- to determine what kind of electronic systems generate records;
- to formulate criteria that allow for the appropriate segregation of records from all other types of information in electronic systems generating and/or storing a variety of data aggregations;
- to define the conceptual requirements for guaranteeing the reliability and authenticity of records in electronic systems;
- to articulate the administrative, procedural, and technical methods for the implementation of those requirements; and
- to assess those methods against different administrative, juridical, cultural, and disciplinary points of view (Duranti and Eastwood 1995, 215; Duranti and MacNeil 1997, 47).

The perspective of the UBC Project was from the point of view of the records creator, specifically a corporate body. While an agency is using its records it has a direct interest in “making and maintaining reliable and authentic records in order to carry out its activities.” Once the records are no longer used, that circumstantial guarantee of trustworthiness no longer exists, and transfer to a neutral third party is essential (Duranti and MacNeil 1997, 57–60).

The first step of the project was to define terminology – what exactly was meant (and could be operationalized) by the terms ‘integrity’, ‘reliability’, and ‘authenticity’. The precision with which these and other concepts were analyzed and defined is another defining characteristic of the UBC project and the subsequent InterPARES projects. In the first published progress report, Duranti and Eastwood wrote: “The first step in the study was to refine more clearly the concept of *integrity*, which was internally broken down in two concepts, *reliability* and *authenticity*. This amounts to saying that preserving the integrity of records means ensuring that they are created reliable and maintained authentic” (1995, 215, italics mine).

The meaning of the concepts of reliability and authenticity were derived from diplomatics: reliability is the authority and trustworthiness of records as proof and memory of the activity, their ability to stand for the facts they are about. Reliability can be assessed by degrees, based on the accumulated information about the level of control over the procedure of the record’s creation (the body of rules governing the making, receiving, and setting aside of records, and competence of persons involved), and the degree of completeness of the record’s form (that the record possesses all the elements of intellectual form necessary for it to be capable of generating consequences). Traditional indicators of reliability include one or more dates (linking the document to its author and the fact observed to its observer) and a signature (which assigns responsibility for the record and its content, and makes of the record a fact to be observed.) The more rigorous and detailed the rules, the more established the routine, the more reliable the record will be. Reliability is the responsibility of the creator of the record, through the record's form and procedure of creation, and the trustworthiness of the persons involved in its creation. A record can never be adjudged more reliable than at the moment of its creation (Duranti and MacNeil 1997, 54).

Authenticity is defined as the trustworthiness of a record as a record – that it is what it purports to be and is free from tampering or corruption (Duranti 2001, 44). It refers to "the maintenance of a record's reliability through its transmission, use, and preservation over time. A record is authentic when it can be proved to be that which it is claimed to be at some point in time after its creation… Authenticity is provided to a record by the controls established on its transmission and preservation. In contrast to reliability, authenticity cannot be assessed by degrees: a record is either authentic or not" (Duranti and Eastwood 1995, 216). Authenticity and reliability are linked in the following way: "Authenticity … is protected and guaranteed through the adoption of methods that ensure that the record is not manipulated, altered, or otherwise falsified after its creation, that is, the record is precisely as reliable as it was when made, received, and set aside" (Duranti and MacNeil 1997, 56). It was in preservation and custody that the research team found the greatest difference between analogue and digital records: while the authenticity of analogue records is protected by keeping them in the same form and state of transmission as when created and set aside, the vulnerability of digital records and rapid obsolescence of hardware and software demands that they be copied and migrated over time through "self-authenticating processes of reproduction… and conversion" (Duranti and MacNeil 1997, 57).

There were two categories of research findings: specific methods for ensuring reliability and authenticity of electronic records, and management issues concerning the maintenance and preservation of reliable and authentic records. The team found that reliability and authenticity are best ensured by embedding procedural rules in the overall records system and by integrating business and documentary procedures, and by establishing agency-wide control. Procedures that strengthen the archival bond (e.g. classification, registration, and record profiles) provide the best guarantee of reliability and authenticity, and preservation of these qualities is only possible if the management of the electronic and non-electronic components of the records system is integrated. The team recommended that the life cycle of managerial activity directed to the preservation of the integrity of electronic records be divided into two phases: control of the creation of reliable records and maintenance of authentic active and semi-active records, and preservation of authentic inactive records. A separation of duties between the records creator (who assumes primary responsibility for their reliability and authenticity while they are needed for business purposes) and the records preserver (who assumes responsibility for their authenticity over the long term) provides the best assurance of the integrity of electronic records. Reliability, governed by the creator, is ensured by procedural and technological controls over persons, process of creation, and definition of record forms. Authenticity is "guaranteed by the adoption of procedural and technological methods aimed at ensuring their proper identification in context (administrative and documentary), and their secure transmission and maintenance," and once inactive, it must be protected "by physically transferring them to a neutral third party and implementing intellectual control through archival description" (Duranti and MacNeil 1997, 57–62).

Theory was operationalized in a collaboration between the UBC research team and the U.S. Department of Defense Records Management Task Force that saw the hypotheses of the UBC project expressed as activity models and entity relationship diagrams, and then translated into mandatory functional requirements for records management application software (DOD 5015.2 STD) (Duranti, MacNeil, and Underwood 1996; Thibodeau and Prescott 1996; MacNeil 2000b, 142 chapter 4, note 25). The validity of traditional archival and diplomatic concepts was therefore tested and found to provide a "powerful and internally consistent methodology for preserving the integrity of electronic records" (Duranti and MacNeil 1997, 64).

2.4.5 International Research on Permanent Authentic Records in Electronic Systems (InterPARES)

The longest running, continuously funded research[7] into the preservation of authentic digital records has been the InterPARES Project at the University of British Columbia. InterPARES has developed knowledge essential to the long-term preservation of authentic records created and/or maintained in digital form, and provided the basis for standards, policies, strategies and plans of action capable of ensuring the longevity of such material and the ability of its users to trust its authenticity. InterPARES is international in scope, and supported by an interdisciplinary process that has included a

[7] InterPARES has been funded through all four phases by the Social Sciences and Humaniteis Research Council of Canada (SSHRC).

wide range of academic and professional fields, from sciences and the arts, to computer engineering and law (Duranti and MacNeil 1997; Duranti 2005a; Duranti and Preston 2008).

InterPARES has been carried out in three completed phases, and a fourth phase is in progress. The first phase, InterPARES I (1999-2001), sought to address the problem of assessing and maintaining authenticity of records (primarily born digital textual records in databases and document management systems) when they come into archival custody. InterPARES 1 was organized around four domains of inquiry for inactive electronic records, the first of which developed the conceptual requirements for preserving authentic electronic records and the identification of elements necessary to maintain their authenticity over time. The concepts of reliability, authenticity, record, and electronic record adopted and developed in the UBC Project formed the basis of inquiry. Research was conducted from the point of view of the preserver and the life-cycle model of administrative and legal records generated in databases and document management systems (Duranti 2001, 50; Duranti 2005a, 12–18; Duranti 2007, 113).

The Authenticity Task Force explained the rationale for establishing conceptual requirements for assessing the authenticity of electronic records. It recognized that the records relied upon by their creator in the usual and ordinary course of business are presumed to be authentic. This is true in archival theory and in jurisprudence (where it is the foundation of the business records exception to the rule prohibiting hearsay evidence from being admissible at trial). However, in the digital environment, records are at risk of intentional or unintentional alteration, which may be difficult to determine. The Task

Force further distinguished electronic records that exist as created, and those that have undergone change of some kind (for example format change or migration). Both types are considered authentic if relied upon by their creator. However, the authenticity of electronic records is threatened whenever the records are transmitted across space or time, necessitating the means for assessing and maintaining authenticity to support the presumption that records continue to be as claimed and free from corruption or undocumented modification (MacNeil and Gilliland-Swetland 2005, 22, 49).

The findings of the Authenticity Task Force were both conceptual and methodological. Conceptual findings provided requirements for authenticity, defined the concept of authentication, and introduced the concept of the presumption of authenticity. The Task Force found that, to assess the authenticity of an electronic record, the preserver must be able to establish its identity and demonstrate its integrity. This represents a development in the way the terms ‘authenticity’ and ‘integrity’ are related from the previous (UBC-MAS) research. In InterPARES I, and subsequent phases, the identity of a record refers to the attributes that uniquely characterize it and distinguish it from other records, while the integrity of a record refers to its wholeness and soundness, that is, to the fact that it is complete and uncorrupted in all essential respects. An important finding of the research was that “complete and uncorrupted in all essential respects” does not necessarily require the record to maintain the same bit structure, but means that the message the record is meant to communicate in order to achieve its purpose is unchanged. The preserver must assess the authenticity of records transferred from their creator. Thus a presumption of authenticity is an inference based on evidence about how the records have been created and maintained. Evidence may come from the creator, or through further analysis to verify authenticity, such as comparison of the records with copies preserved elsewhere (redundancy), forensic analysis, testimony of a third party, or analysis of audit trails (MacNeil and Gilliland-Swetland 2005, 47–51).

The Task Force developed benchmark requirements, that give reasonable assurance of authenticity prior to transfer of records from their creator to the trusted preserver (trusted recordkeeping), and baseline requirements that support the production of authentic copies of electronic records that have been transferred to the preserver (trusted custodianship). The benchmark requirements included:

- identification of fundamental information that establishes a record's identity and allows for demonstration of its integrity, explicitly expressed and inextricably linked to the record (may appear on face of record or in metadata),
- evidence of access privileges that show the assignment of authority and capacity to carry out administrative action accompanied by exclusive technical capability to exercise such responsibility,
- establishment and implementation of procedures to prevent, discover, and correct loss or corruption of records (regular backups of both files and systems...)
- establishment and implementation of procedures to guarantee the continuing identity and integrity against media deterioration and across technological change,
- establishment and control of documentary forms (down to the level of record elements) associated with procedures either according to juridical requirements or institutional policy.

The creator must also specify details governing authentication of records, establish procedures to identify the official record from among multiple copies, and establish and implement procedures to determine what documentation must be removed and transferred to preservation with the record (i.e. what information is required to establish and maintain identity and integrity).

The baseline requirements to support the production of authentic copies require that:

- procedures and systems used to transfer, maintain and reproduce embody adequate and effective controls to guarantee integrity and identity, including
    - unbroken chain of custody
    - security and control procedures implemented and monitored
    - content unchanged after reproduction,
- activity of reproduction must be documented, including
    - date of reproduction and name of responsible person
    - relationship between records acquired from creator and copies produced by archivists
    - impact of reproduction process on form, content, accessibility and use
    - details of any elements not fully and faithfully reproduced,
- description of all technological changes are included as part of archival description (a collective attestation of authenticity of records in the archival group and all their interrelationships) (MacNeil and Gilliland-Swetland 2005, 204–219).

The Task Force found several deficiencies in the electronic systems they observed with respect to creating, maintaining and preserving records, as defined by archival diplomatics. For example, electronic systems are often designed to manage data rather than records – that is, fixity requirements for records did not exist. Identity information is often implicit in the records, with the consequence that key indicators of identity may be lost when the records are transferred out of the record creating or record keeping system. Indifference of records creators to issues of authenticity were also common, replaced by confidence in the technology to protect the authenticity of the records (MacNeil and Gilliland-Swetland 2005, 52). Limitations of diplomatics as an analytical tool were also discussed by the Task Force - a discussion that paved the way to the second phase of the InterPARES (MacNeil 2004).

InterPARES 2 (2002-2007) returned to the perspective of the records creator. In addition to dealing with issues of authenticity, it delved into the issues of reliability and accuracy during the entire lifecycle of records, from creation to permanent preservation. The project was organized in three research domains: digital records creation and maintenance; authenticity, reliability, and accuracy of digital records in the artistic, scientific, and governmental sectors; and methods of appraisal and preservation. These domains were supported by four cross-domains that modeled the records life cycle and continuum (developing the Chain of Preservation model and the Business-Driven Recordkeeping Model), investigated the role of metadata (description cross-domain), structured the relationship between creators and preservers through policy (policy cross-domain) and studied the terminology that underpinned relevant issues across disciplines (terminology cross-domain). The focus of research was on records produced in complex (dynamic and interactive) digital environments in the course of artistic, scientific and governmental activities (Duranti and Preston 2008).

The Domain 2 Task Force, investigating authenticity, reliability and accuracy of digital records, carried out case studies in the artistic, scientific, and governmental sectors. Building on the work of InterPARES I, the Task Force was immediately confronted with the challenges of diverse domain understanding of what is meant by the terms ‘record’ and ‘authenticity’ in the three areas of investigation, and the fact that the structure and function of digital entities created in art and science often did not resemble those in legal or administrative contexts. It was cognizant of the fact that the diversity encountered in the case studies also reflected lines of thought about the constructed nature of authenticity developing in the postmodern archival literature. It found that, while the benchmark requirements were useful for measuring a presumption of authenticity, they could be difficult to apply or adapt depending on the nature of the creator’s records, and in some cases were not sufficient to preserve the kinds of authenticity valued by the creator. It also found in several disciplines limited definitions of authenticity that related it most closely to integrity. Frequently authenticity was presumed from the circumstances of record creation, or linked to technological methods of authentication. Within the sciences, for example, the term ‘authenticity’ is rarely used, although information about identity, captured in metadata, integrity, ensured through authentication and security measures, and provenance, or lineage, is crucial (Roeder et al. 2008, 141–163).

Scientific disciplines do not normally use the word ‘authenticity’ when describing datasets, although the fundamental archival concepts are often addressed, either implicitly (trusted source) or explicitly (data lineage, integrity). They are more concerned with issues of completeness, reliability, accuracy, and integrity. Many have issues of legacy datasets that have been digitized. In these situations, if the source of the original data can be assumed trustworthy, then the data acquired are presumed reliable and accurate (Hackett, Underwood, and Eppard, n.d., 33–41).

In the field of Geography and Geomatics, authenticity is assessed through analysis of data lineage, which is one of at least seven elements comprising ‘spatial data quality’. Data lineage information records the chain of transmission of a dataset from the moment of data collection. It is the history of a dataset from collection through stages of compilations, corrections, conversions, transformations (Hackett, Underwood, and Eppard, n.d., 31–32). In scientific fields generally, accuracy of data receives the most attention, with primacy given to data quality, which includes the concept of authenticity, (normally articulated as data provenance or lineage) (Roeder et al. 2008, 133–137). Metadata are means of attesting to and assessing a dataset’s authenticity – authenticity is linked to a clear lineage recorded in the accumulating metadata surrounding the data.

Is the distinction between ‘data’, ‘information’, ‘document’, and ‘record’ important? Data is the smallest meaningful unit of numbers, symbols, letters, or words that can be understood and manipulated. Data is the ‘raw material’ that can be processed to become information. Archival science defines ‘document’ as information recorded on a medium in a fixed and stable form, and ‘record’ as a document made or received in the course of practical activity and set aside for future action or reference. While the nature of ‘data’, ‘documents’ and ‘records’ are different, in the digital environment where data, documents, and records coexist, the issues related to their management coincide (Duranti and Rogers 2012, 3; Krementz 2009). Todd suggests that archivists must now be prepared to assess the authenticity of objects at the “sub-record level” (Todd 2006, 182) in the form of digital components of records (which may themselves be records), or digital data. One of the affordances of digital technology is the ability to compile, combine, recombine, mine, and analyze vast datasets. “Big data” includes scientific and academic datasets, census and other data collected by government on its citizens, health, surveillance, and commercial data. Access to data and datasets provides the possibility for validation of previous experiments, as well as the possibility of conducting new research. This has expanded the horizons for scientific research, allowing heretofore unimaginable advances in knowledge such as the mapping of the human genome. The resultant datasets "constitute a critical national resource, one whose value increases as the data become more readily and broadly available" (Lauriault et al. 2007, 125; See also Gantz and Reinzel, 2011). The value of these data depends on their quality, which for scientific data includes authenticity (Lauriault et al. 2007, 125).

The preservation of authentic datasets of information collected through observation, computation, or experiment is of increasing concern (National Science Foundation 2005, in Lauriault et al. 2007, 132, n. 32). These data may be historical recordings of natural events that can never be replicated or recollected, may concern models for complex computations, such as climate change models, or be experimental, reproduceable only at prohibitive cost, or not at all. Scientists give primacy to data quality, which they equate with authenticity, and base on provenance or lineage, and traceability, expressed through metadata or data-quality parameters. As stated previously, the term "authenticity" is not often used, despite the discussion of qualities of identity and integrity through concepts of data provenance and data lineage. Lineage is represented in an audit trail that provides the data with assurances about its source or pedigree, and fitness for use (Lauriault et al. 2007, 153).

The trustworthiness of official statistics relies on citizen confidence that they are independently produced and free from bias or political interference. Statistics are based on data collected through a variety of government and research agencies. Increasingly, governments are making large datasets available for public scrutiny and analysis through official programs of open data. A comparison of open data policies in national and regional jurisdictions across North America (US and Canada) enacted from 2009 through 2014 show, however, that specific quality controls are generally lacking. For example, open data policy recommendations such as publishing metadata, making available information about the data creation process, sharing of code or publishing open source, and requiring the use of unique identifies – all critical mechanisms for establishing authenticity, provenance and data quality – are addressed in a very few, if any, jurisdictions (Sunlight Foundation 2014a; Sunlight Foundation 2014b).

In government, concepts of authenticity, accuracy and reliability are seldom addressed directly. Concerns about authenticity in the electronic environment tend to be generic, and difficult to address because of imprecise terminology, which as used in the governmental sector in discussing digital records is at times vague or inconsistent. This is particularly true for words like "authenticity," "accuracy" and "reliability," which are not technical terms in general parlance, but words with common sense, everyday meanings. The research team found that the concept of authenticity was frequently equated with integrity. The conclusion for the government sector was that, although concern for authenticity of records was high, the use of terminology was loose. Authenticity was often presumed, particularly in instances where authentication techniques are employed (Roeder et al. 2008, 126–133).

One of the key products of InterPARES 2 is the set of Guidelines for creators and preservers that operationalize and further develop the Benchmark and Baseline Requirements that were a product of InterPARES 1. The development of the Guidelines was led by the Domain 2 Task Force in an iterative process that sought consensus from archival scholars, practicing archivists and specialists in the arts, sciences, and government. One recommendation that arose from the case studies for developing the Guidelines suggested avoiding using the terms authenticity and reliability, while clarifying what records must have to be authentic and reliable – these terms, although precisely defined in archival science, mean different things to different creators, and if they are used at all, they are often confused or conflated (Roeder et al. 2008, 133). Despite the reservations of the Task Force that the Guidelines could not claim to exhaust all preservation-related issues (Roeder et al. 2008, 185–186), these Guidelines continue to be highly referenced and used (as judged by the frequency of their download from the InterPARES website.)

The Description Cross-Domain Task Force examined the crucial role of recordkeeping metadata in the creation of authentic records and the maintenance of their authenticity over time and across technological change. Their premise was that detailed and trustworthy metadata were key to the creation of reliable and preservation of authentic digital records (Gilliland 2008, 335). Metadata describing records and the actions taken on them are key to establishing and assessing authenticity (Gilliland 2008; Gilliland and McKemmish 2012). Metadata are the machine- and human-readable assertions about information resources that allow for physical, intellectual and technical control over those resources. Users create and attach, and then maintain and preserve metadata, either automatically and/or manually, when maintaining their digital records, documents, and data. These metadata may be technical, administrative, or descriptive. They codify and track the identity and integrity of the material over time and across technological change. The importance of recordkeeping metadata has been acknowledged since the 1990s (e.g. Hurley 1995), but in practice metadata frequently still remain underused and misunderstood (Isaza 2010).

The Description Cross-Domain Task Force undertook the development of a metadata schema registry, and identified all the elements of metadata arising from the Chain of Preservation model across the life cycle of records (Gilliland 2008). In the most comprehensive analysis of metadata requirements in relation to reliability and authenticity to date, the Task Force concluded that deficiencies and challenges in the current state of metadata needed to be addressed within both a custodial and non-custodial recordkeeping model and a life cycle and continuum world view of records. Areas identified for future research included the degree to which recordkeeping metadata could contribute to archival description, and the extent to which automated tools can utilize metadata in description and use (Gilliland 2008, 345–350).

InterPARES 3 (2007-2012) built upon the findings of InterPARES 1 and 2, as well as other digital preservation projects worldwide, to put theory into practice, applying the results of the previous two phases through case studies with small and medium-sized organizations, or those with limited resources, and general studies. One general study built on the work of the Description Cross-Domain of InterPARES II and attempted to develop an application profile for authenticity metadata based on the benchmark and baseline requirements as articulated in the Chain of Preservation model (Tennis and Rogers 2012b; Tennis and Rogers 2012a). This work is ongoing.

2.4.6 Authenticity in Related Digital Preservation Research Projects

Because of the cross-disciplinary nature, sweeping scope, and staggering cost of digital preservation, research is often carried out by national and international alliances of universities, libraries and archives, government agencies, business and industry. Each alliance is defined by its particular epistemic perspective and purpose. However, cooperation and collaboration, if not always agreement, are constants across the entire research community. There are also major national initiatives undertaken by national archives and/or libraries, such as those in Australia, the United States, and Denmark.

Meaningful engagement with digital information resources requires predictability and comprehensiveness, interoperability, transactionability, and preservability (Lavoie and Dempsey 2004). Digital preservation is partly a technical problem, but more importantly, it is "one component of a broad aggregation of interconnected services, policies, and stakeholders which together constitute a digital environment" (Lavoie and Dempsey 2004). Preservation research can be classified according to its particular focus: the development of standards, frameworks, and repository systems (e.g. OAIS); defining and using/sharing metadata schemas (e.g. PREMIS, OAI); the nature of digital objects (e.g. InterPARES, InSPECT); technologies of preservation (e.g. preservation-aware storage); and file formats and object identification (e.g. JSTOR, JHOVE). All of these projects share a common goal, that of preserving digital objects that can be trusted, although not all of them approach authenticity explicitly. Of note, and discussed below, are OAIS and CASPAR, both of which are connected in different ways to InterPARES. A comprehensive summary of preservation research from the early 1990s through the 2000s is found in Anne Gilliland's book, *Conceptualizing 21st-Century Archives* (Gilliland 2014).

The Open Archival Information System (OAIS) Reference Model is a high-level model and the benchmark for digital preservation systems, addressing all aspects of long-term preservation of digital information: ingest, archival storage, data management, access, dissemination, and migration to new media and forms. The reference model does not dictate means of implementation, but prescribes requirements to ensure that an OAIS-compliant repository is "an organization of people and systems that has accepted the responsibility to preserve information and make it available for a Designated Community." An OAIS archive, therefore, is situated in the context of a user community and answerable to that community. The Reference Model describes the external environment, the functional components or internal mechanisms, which collectively fulfill the preservation responsibilities, and the information objects that are ingested, managed, and disseminated by the OAIS. Developed in 2002 by the Consultative Committee for Space Data Systems, the OAIS is now an approved ISO standard (ISO 14721:2003) and has undergone several revisions, the most recent in 2012 (CCSDS 2012).

This latest revision addresses authenticity requirements more directly than previous revisions; however, as it is a high level standard, it does not dictate how authenticity is to be ensured or protected. It defines authenticity as “the degree to which a person (or system) regards an object as what it is purported to be. Authenticity is judged on the basis of evidence”. This is consistent with the InterPARES definition (Giaretta et al. 2009, 69). Part of the necessary evidence is provided by Provenance Information, which tells the origin of the source of the Content Information, documents changes to it and the chain of custody since creation. Authenticity, a stated objective of long-term preservation, is the responsibility of the repository to protect (CCSDS 2012, 1.9–1.14).

The Alliance for Permanent Access (APA) is a major research consortium of European libraries, archives, universities, government agencies, and businesses which aim “to develop a shared vision and framework for a sustainable organizational infrastructure for permanent access to scientific information” for designated communities of users (www.alliancepermanentaccess.org/). Digital resources considered for preservation include natural science and social science datasets, government, health, and economic data submitted to national data archives conforming to the OAIS standard. This shifts the focus of authenticity requirements from the record or digital object in general to the authenticity needs of a specific community of users.

As international and interdisciplinary in scope as InterPARES, APA encompasses several research projects studying the assessment of authenticity of data. One such project is APARSEN (Alliance for Permanent Access to the Records of Science in Europe Network), launched in 2010, and designed to bring together work in digital preservation carried out across Europe. APARSEN defines success as establishing “coherence and general direction of travel of research in digital preservation, with an agreed way of evaluating it and the existence of an internationally recognized Virtual Centre of Excellence" (Alliance for Permanent Access 2012).

APARSEN builds on another project under the APA umbrella, CASPAR (Cultural, Artistic, and Scientific Knowledge for Preservation, Access, and Retrieval). CASPAR developed an Authenticity Conceptual Model that is OAIS-compliant, technology neutral, and domain-independent (Lamb 2009). The model consists of an Authenticity Protocol, applied to an Object Type, and comprising Authenticity Steps (Reference, Provenance, Fixity, Context, Access/Rights) (Guercio 2008; Guercio and Michetti 2009a; Guercio and Michetti 2009b; Giaretta 2011, 209–210) Authenticity Protocols (APs) are defined as "procedures to be followed in order to assess the authenticity of specific type of Digital Resource (DR)." CASPAR conducted its research based on certain assumptions about digital preservation: that it is not enough to preserve just the bits, but also information and knowledge; that preservation is a process of transforming and enriching content through different technological strategies to adapt it to new constraints of rendition and playability, to preserve its intelligibility and (re)usability, and to ensure its integrity and authenticity (Salza et al. 2012; Guercio and Michetti 2009a; Guercio and Michetti 2009b; Guercio 2008).

Early in 2012 APARSEN released a report on the implementation and testing of domain-specific authenticity protocols. This comprehensive report begins with a "State-of-the-Art" outline of related projects in digital preservation research – first on the list is InterPARES, followed by CASPAR. These three projects are highly connected in purpose and complementary in approach. APARSEN adopts the CASPAR definition of authenticity, which is general and high level, and the theoretical underpinnings of InterPARES, and has formalized an authenticity management model, based on the principle of performing controls and collecting authenticity evidence in connection to specific events of the digital object's lifecycle. This allows the assessor – preserver or user – to trace back all the transformations the digital object has undergone since its creation and that may have affected its authenticity (Salza et al. 2012, 8).

2.4.7 Exploring New Models of Record and Record Authenticity

Through the 2000s the concept of record was revisited (cf. Lemieux 2001; Yeo 2007; Yeo 2008), and with it, the interrelated concepts of risk, authenticity, and trust. In the late 2000s and into the 2010s, emerging issues associated with the continuing advance of digital technology further complicated recordkeeping and archival practice. Authenticity remains a critical issue in research into digital preservation and access, with a number of major projects funded by the European Union through their EU Framework Programme (cf. Giaretta 2011; Strodl, Petrov, and Rauber 2011). Issues of trust and confidence in the Web are the subject of computer science research (cf. Cofta 2007b; Cofta 2013). The failure of record trustworthiness in the digital environment has been attributed as a significant factor in national banks crises (cf. Lemieux 2001), and in the global financial crisis (cf. Tonkiss 2009; Gurría 2009).

The literature spans not only the technological developments that have brought so much change to records professions and records-related issues, but significant developments in archival worldview. This is reflected most clearly in the theoretical archival literature, where the rise of critical, hermeneutic, or pragmatic epistemologies (Hjørland 2008) resulted in new interpretivist concepts of record (cf. Lemieux 2001; Lemieux 2014; Yeo 2007; Yeo 2008), of archival functions (cf. Cook 2001; Cook 2012; Nesmith 2002), and of custodianship (the continuum model) (cf. Upward 1996; McKemmish 2001; Upward 2005). Different articulations of the concept of 'record' continue to emerge, arising from the particular challenges of increasingly complex digital technological infrastructures (Lemieux 2001; Yeo 2007; Yeo 2008; Duranti 2009; Duranti and Endicott-Popovsky 2010; Lemieux and Limonad 2011; Thibodeau 2013; Lemieux 2014).

Also, "[t]he meanings of 'authenticity' are relative to the concept of authentic that is held by different disciplines" (Lauriault et al. 2007, 140). This idea has been explored in recent literature (MacNeil and Mak 2007; Duncan 2009; Mak 2012). At the root of these explorations is the concept of authenticity as a social construction dependent on the context or discipline within which authenticity is defined, interpreted, and required. If one subscribes to the view that digital resources are "in a continuous state of becoming" as they are created, used, migrated, preserved, and accessed over time, then so too is the nature of their authenticity (MacNeil and Mak 2007, 26). In both cases, questions remain about how to define the necessary elements of authenticity within a given context, and how to assess them.

2.5 Authenticity on Trial

In legal discourse, authenticity is part of the foundation upon which the admissibility of documentary and real evidence is based, proven through the authentication of documents at the moment they are introduced as evidence. According to Duranti, legally authentic documents are those that bear witness, i.e. may be admissible as evidence, through the guarantee of their genuineness by a recognized authority during or after their creation (Duranti, 1998, 47). Authenticity is related to and supportive of the concept of "best evidence." Authentication serves to establish the identity of the record and its relevance to the issues in the proceeding, while the best evidence rule demands proof of the integrity of the contents (MacNeil 2000b, 46–48).

Traditionally, documentary best evidence is satisfied by production of the original document, thus linking the legal concept of authenticity to the status of transmission, that is, the degree of perfection, of a record. However, the historical notion of authenticity, whereby an authentic record is one which attests "to events that actually took place or to information that is true" (Duranti 1998a, 46), focuses attention on the contents of a document and their truthfulness. Mapped to legal practice, historical authenticity is seen in moments when a document is considered "for the truth of its contents" (Sheppard and Duranti 2010, 27) This is tested not through authentication or establishment of best evidence, but in the consideration of the hearsay nature of the document, and is linked to reliability. These are nuanced approaches to understanding the meaning of authenticity of documents and the purpose for which authenticity is considered important.

In common law legal systems, a document must be identified – that is, authenticated – before it can be entered into evidence. If a document is admitted for a hearsay purpose, the truth of its contents can be tested and weighed (Scanlan 2011, 7). The authenticity of a document may be conceded or disputed. If it is disputed, the proponent must offer proof of its authenticity. Authentication of documentary evidence can be proven through testimony, expert analysis, non-expert opinion, or, in the case of public documents or other special types, circumstances of record creation and preservation (*Federal Rules of Evidence* 2014, Rule 901).

Determining authenticity and trustworthiness of a document should be, at least in theory, straightforward endeavors. But what happens when the technology with which facts and acts are recorded changes so dramatically that the resulting “document” – an object that has been commonly understood and readily recognized for centuries – is no longer recognizable? What happens when the volume of material presented in discovery threatens to overwhelm the system’s ability to review and analyze it?

Digital communications technologies have profound implications for the administration of justice. Traditional issues of jurisdiction, validity and enforceability of agreements, and rights and obligations in transactions are now being examined through documentary evidence that is often digital and may bear little or no resemblance to traditional documentary evidence, but is still subject to traditional rules of admissibility. It is not the substance of business conducted electronically that raises new questions, but the process, and it is concerning the resulting documents that these questions arise (Daughtrey 2000).

Our legal system regulates and constrains human behavior through laws and rules. Evidence law in the common law tradition has developed over centuries to ensure that only relevant, material, and authentic evidence is presented to decision-makers. In respect of documentary evidence, common law and statutory rules developed that allowed evidence to be tested for authenticity, reliability, and accuracy, and so trusted or rejected as proof of a matter at issue. In the digital environment this is no longer a straightforward matter.

Until the end of the 20th century, these rules were predicated on mechanical writing technology by which a writer created a document by inscribing a message directly onto a physical medium and signing it, thereby signifying its authenticity (the prototype being a hand- or type-written document on paper). Authentication of traditional documents was based on established methods such as direct testimony and physical inspection. The rules have worked well for traditional documents. The law of evidence has also been flexible in accommodating analog electronic technologies. However, the law is severely challenged by documents produced, transmitted, and stored in digital form. “We now live in an age of information complexity, and that fact has largely destroyed our existing evidentiary scheme” (Paul 2008, 21).

Research has shown that digital evidence is precipitating an “authenticity crisis” in North American courts (cf. Paul 2004, 2008; Chasse 2011; Sheppard and Duranti 2010). In a constant state of technological evolution, the law of evidence must adapt to changing societal norms and expectations and technological advances. It continues to do so in the face of rapidly developing computer and network technologies. The law has tended to approach new media from the perspective of analogy – comparing the traditional to the new, and looking for commonalities. The use of “appropriate existing and historical analogies” has guided the development of rules governing documents produced by new technologies (Takach 2003, 13), and the “functional equivalent approach” has been used to map purposes and functions of paper-based evidence to new media and forms (Davies 2008, 3).

Computer technology cannot be compared solely to paper-based communication technologies. Certainly for most computer users the parallel may easily be drawn – the computer is used to produce documents in the traditional sense through office software – word processing, spreadsheets, and presentations whose function and importance to the creator is in the presentation seen on the screen or printed to paper. However, even these "traditional" documents carry with them traces (metadata) attached by the creating technology that provide contextual evidence of creation, transmission, and alteration. Some of these traces may be easily identified by those with sufficient digital literacy to understand the technology, but many of them require the special knowledge of experts in computer forensics. Some metadata associated with office software may be easily viewed, but other metadata are more difficult to uncover. Log files created in the course of Internet banking and commercial transactions, for example, are not seen by the parties in those transactions but can provide critical evidence of their existence. Other traces may be found that give evidence of files that have been deleted. This creates evidence governed under the rubric of forensic science, while exhibiting many of the characteristics of documents.

It has been said that the most challenging aspect of offering digital evidence to a court is to establish its authenticity (Grimm, Ziccardi, and Major 2009, 365). The problems posed to admissibility by what may now be considered *traditional* digital evidence have not yet been solved, but technology is not waiting for the law to catch up. New challenges are developing with the increasingly ubiquitous use of social media for personal and business purposes (Strutin 2011, 228). In 2011 the case of *State of Connecticut vs Eleck* (WL 3278663, Conn. App. 2011) considered the admissibility of Facebook evidence and noted "that while the emergence of social media evidence does not necessarily require new rules of evidence, 'circumstantial evidence that tends to authenticate a communication is somewhat unique to each medium,'" and that "precedent cases where emails, chat logs, and texts were properly admitted [were] based upon their supporting and unique metadata and other circumstantial evidence that provide 'identifying characteristics'" (Patzakis 2012 np). One obvious, but by no means the only, question, is: what constitutes admissible digital evidence – that is, how can we define and assess the identity and integrity of material resulting from rapidly developing technology? While the court may recognize that proper authentication depends upon appropriate "identifying characteristics," these remain to be consistently identified and theoretically justified.

In general, authenticity of digital evidence is considered on a case by case basis (Mason 2012, 132). The often-cited authority on the issue of authentication is Edward Imwinkelried (Fosmire 2006, 1–2; Imwinkelried 2005 np) who has identified eleven foundational points that should be established in order to authenticate digital business records. These eleven points confirm that the business uses a computer, has developed procedures for use and safeguards for accuracy, maintains its reliability and security, determines that it was working properly at the material time, and that a witness can recognize and verify the data at issue produced by the computer. This list was cited in the seminal case, *In re: Vinhee* 2005 WL 3609376 (B.A.P. 9th Cir. Dec. 16, 2005) in which the issue was the authentication of the data themselves – "showing that the data properly represented what it claimed to show." According to Fosmire, however, this list is "somewhat out of date" (2006, p.3).

Legal scholars continue to debate the extent of the challenges being created by digital evidence. They have tended to adopt one of two lines of argument. The first argument presents the view that the law of evidence is flexible, and legal practitioners can approach any legal challenge posed by new technology and related business practices through a judicious combination of the common law, private contracts, technological solutions, and law reform . "Time and again, courts have shown a receptiveness to understand the new technology, a keen ability to tackle the various dynamics and themes of computer law, and a willingness to use their significant latitude to craft appropriate rules for a specific technologically driven problem" (Takach 2003, 4). Many legal scholars, attorneys, and judges who write about digital evidence, however, believe that greater understanding of the nature of digital materials and the affordances of digital technologies is crucial to avoid a miscarriage of justice. They call for continuous review and reform of the laws that govern admissibility of digital documentary evidence (Chasse 2007; Mason 2010; Fisher 2004; Paul 2008; Facciolo 2010). One writer predicts "rough justice" resulting from a poor understanding of the concepts of digital evidence, which requires a conceptual leap "tantamount to a paradigm shift" (Mason 2008 np). Furthermore, some legal scholars fear that vagueness and uncertainty of statute and common law as they apply to digital records means that litigants do not know how best to prepare their positions, and court decisions lack consistency (Chasse 2007, 147). Each argument is illustrated with case law that support its position.

Challenges to the authenticity of digital evidence can include a claim that the records in question were altered, manipulated, or damaged; that the reliability of the technology is in question; that the identity of the author is in question; that evidence from social networking sites is unreliable; that the person who had access to a device, and therefore might have sent or received a particular message is at issue; or that an unauthorized person used a password, PIN, or click through (Mason 2012, 112). A recurring question in litigation is whether the new medium requires different authentication procedures for admissibility in court. In *People v Patterson* (1999 New York Court of Appeals) traditional methods of proof were deemed applicable to new and developing technologies, subject to "the obligation and need for responsible accuracy and careful reliability" not sacrificed to "the whims and weaknesses of fast moving and rapidly changing technology" (Crusco 2014).

Conceptually, digital evidence may be authenticated by proving the provenance of the proffered evidence through organizational criteria; through digital forensics testing the characteristics and content of the material; or through testing any signatures, seals, or time stamps (Mason 2012, 119). There are both technical and organizational considerations relating to authenticity. Technical considerations include an analysis of the risks of the chosen method of preservation, and defense of the choice should it be questioned. Also to be considered are how the identity of digital evidence will be established (e.g. name of author, date of creation, place of origin, subject matter) and how its integrity will be proven (e.g. evidence of changes, use of time stamps and demonstration of their accuracy, metadata showing transmission and establishment of the reliability of that metadata, use of data logs). Organizational considerations include circumstantial evidence of integrity through information about policies and procedures and how they were created (and evidence that they were followed), security and access controls, credible metadata, audit trails and reports, and intrinsic elements of record form (Mason 2012, 123–131).

2.6 Digital Forensics and Information Assurance

Digital diplomatics, developed and refined over the years of the InterPARES research, is ideally suited to the analysis of authenticity of digital *records* as defined by archival science, but is limited when the subject of analysis is broadened to include digital objects that may not satisfy that precise definition (Duranti and Endicott-Popovsky 2010, 2; MacNeil and Gilliland-Swetland 2005, 52). Archivists are now creating research alliances with digital forensics practitioners in order to develop and extend the applicability of digital diplomatics in the field of digital preservation and the focus on authenticity, reliability, and accuracy (Duranti 2009; Kirschenbaum, Ovenden, and Redwine 2010; John 2012). For the archivist, digital forensics offers a new lens through which to view records, documents, and data, and assess their authenticity. Archival repositories are motivated to adopt digital forensics to help process digital information for several reasons. These technologies can support description and context, integrity, version detection, and identification and protection of authenticity (John 2012, 11).

Digital forensics consists of the acquisition of a digital system and the extraction of data from it, analysis of the acquired data and their preservation and presentation. The working definition proposed by the first Digital Forensics Research Workshop (DFRWS) over a decade ago stated that digital forensic science is “The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations” (Palmer 2001). The processes of acquisition and analysis are dominated by technical issues, but presentation depends on policy and legal requirements (Carrier 2003b, 3). Two fundamental issues that digital forensics must deal with are complexity and quantity. These problems derive from the nature of digital technology, and therefore are common to all information domains that deal with digital material. The complexity problem is that digital objects at the lowest level of their existence are streams of bits – series of 0s and 1s. These are not understandable by humans without the intervention of layers of technology through which the data are translated (Carrier 2003a, 2). Part of determining authenticity depends on assurance of integrity of each abstraction layer. Digital forensics offers archival science a more granular and nuanced understanding of integrity. While archivists have defined integrity simply as the quality of being complete and unaltered in all essential respects, focusing on the logical manifestation of the record, digital forensic scientists distinguish several levels of integrity at both the physical and the logical level – at the level of the bit stream, the data, the computer, or the system. However, not all layers need to be or can be maintained without change throughout the life of the object. Analyzing the object through abstraction layers offers the possibility of a more nuanced view of authenticity.

The second problem is that of quantity. Faced with terabytes or more of data, digital forensics specialists, archivists, scientists, and trusted recordkeepers in all domains need to be able to group data by layers, type, or other means in order to analyze them and assess their authenticity. This has been referred to variously as “information inflation” (Paul and Baron 2007, 1), or the “digital tsunami” (Lemieux and Baron 2011, 2).

Digital forensics is also at the core of the information assurance industry, where its primary objectives are continuity of operations and availability of service. Forensic techniques are also becoming more widely used in enterprise risk management (Casey 2007, 49–50). The concept of forensic readiness, defined as the ability to maximize an organization’s potential to use digital evidence while minimizing the cost of an investigation, emphasizes prevention and detection over post-incident investigation. Rowlingson outlines ten steps to enterprise forensic readiness that bear close resemblance to an implementation plan for a systematic records management program. Although he does not provide a definition of digital evidence, business records are clearly a subset of the digital material to which he refers. “Digital evidence is required whenever it can be used to support a legal process. … To succeed in a legal process, it is therefore essential that the organization has actively gathered the evidence it is likely to require. Moreover, it is vital to have the capability to process evidence cost-effectively, and to have suitably trained staff who know how to ensure potential evidence is preserved” (Rowlingson 2004, 3). A well-organized records management program will ensure that human-generated business records, at least, will meet these requirements.

Authenticity can be seen as an important component of information assurance and security (IAS). The National Institute of Standards and Technology (NIST) defines information security as “The protection of information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction in order to provide confidentiality, integrity, and availability”, and information assurance as “Measures that protect and defend information and information systems by ensuring their availability, integrity, authentication, confidentiality, and non-repudiation. These measures include providing for restoration of information systems by incorporating protection, detection, and reaction capabilities” (National Institute of Standards and Technology 2013). IAS has been described as a multidisciplinary knowledge domain (Cherdantseva & Hilton 2013), and a business-wide issue that extends far beyond the IT department (ICC Belgium 2013)

However, authenticity has not always been included explicitly in computer security models. The first conceptual computer security model was the CIA-triad, composed of confidentiality, integrity, and availability. Since its introduction in the mid-1980s security experts have challenged the adequacy of the model and proposed extensions (Cherdantseva & Hilton 2013). In 1998 Donn Parker introduced six foundation elements essential to information security. To the CIA-triad he added utility, authenticity, and possession. He found integrity, the characteristic of being complete and whole and free from corruption or manipulation, to be insufficient without the assurance also of authenticity, or “conformance with reality” (Parker 1998; Kabay 2013). Most recently, Cherdantseva and Hilton have proposed a reference model for information assurance and security that extends the CIA-triad to the IAS Octave: confidentiality, integrity, availability, privacy, authenticity and trustworthiness, non-repudiation, accountability and auditability (Cherdantseva & Hilton 2013).

While most of the digital forensics literature focuses on practical and technical aspects of practice, there are articles spanning the last fifteen years by practitioners and scholars that stand out for their explicit recognition of parallels between the disciplines of digital forensics and archival/records/information management (cf. Rowlingson 2004; Ferguson-Boucher and Endicott-Popovsky 2008; Irons 2006; Lemieux and Baron 2011). These authors' ideas touch on issues of appraisal, records management, and the application of principles of diplomatics, and suggest fertile ground for further research. They are, as yet, the exception – lone voices from the digital forensics perspective embracing archival and records management principles. Clearly, however, this is beginning to change, inspired by projects such as the Digital Records Forensics Project, Records in the Cloud, and InterPARES Trust at the University of British Columbia, and collaboration between the School of Library, Archival and Information Studies at UBC and the Center for Information Assurance and Cybersecurity at the University of Washington (Duranti and Endicott-Popovsky 2010).

Irons recognizes the interdisciplinary background of digital forensics, with its grounding in forensic science, computer science, criminology, law, mathematics, audit and business. He then adds to that panoply records management, suggesting that the two are "compatible disciplines and areas of study" and notes how they can mutually benefit from interaction, but states "There remains very little published on the discussion of the potential implications of computer forensics for records managers or how computer forensics can enhance the records management discipline.” Irons makes explicit the parallels and complementarity of digital forensics and records management in his analysis of the principles of computer forensics in the context of record characteristics of authenticity, reliability, integrity and usability. Most of his arguments present the benefit of digital forensics to records management, although he recognizes that many skills of records management, such as metadata expertise, functional classification, and digital preservation, can benefit forensic investigative techniques. Perhaps more interestingly, he suggests that digital forensics could benefit from the application of theoretical models of records management; however, he does not elaborate further on this point (Irons 2006, 107–110).

If digital forensics specialists are just discovering archival science, archivists have known for some time of their affinity to the forensic discipline. “If the historian is the lawyer in the court of history, then the archivist is the forensic scientist,” wrote Elizabeth Diamond more than fifteen years ago. She notes the parallels between the two disciplines: each has the job of acquiring, preserving, arranging and making accessible “impartial evidence,” and clarifying the meaning of that evidence through “its own distinct knowledge and methodology” (Diamond 1994, 142).

Any definition of authenticity is based on assumptions about the value of the object and how that value can be expressed, assessed, and preserved. Communities of practice each ground their assessment of authenticity in the use to which they put recorded information – a record with the capacity to serve as evidence of an action or transaction, a source that bears witness to the past, or data for use in replicable and verifiable experiments. In order to understand records as being what they purport to be, we must first understand what we want or need them to be. “Quality records can only be realized after prior consideration has been given to our own expectations of what the purpose and intention of the records are” (Duncan 2009, 115).

2.7 Studies of Practitioner Behavior and Authentic Records

Few studies have been conducted on the behavior of records professionals in ensuring, maintaining, and assessing record authenticity. An exploratory pilot study on practitioners’ concepts of authenticity in their work activity was conducted in 1998. Park noted that while questions about authenticity of electronic records had been the subject of archival and preservation research, a systematic investigation of practitioner behavior had not been undertaken. She asked: What does the concept of authenticity mean to practitioners? How do practitioners define the concept of authenticity? And, is the concept of authenticity understood differently in different professional domains? Among her results, she found that while practitioners were highly aware of the concept of authenticity in both paper and electronic records, less than half have been required to authenticate records. Park compared treatment of paper records with treatment of electronic records, and used content analysis to study the use of terminology. Practitioners do not perceive a difference between paper and electronic records with respect to authenticity, although they recognize that the means of authenticating records will be different (E. G. Park 2001). She concluded that research and practice were far apart, and work was needed to bridge the gap. This study preceded her doctoral research, which studied authenticity requirements in university student record systems (E. Park 2002a).

The relationship between ICTs, authentic records, and accountability was examined in an empirical study of accountability forums and public administrations (Meijer 2003). Meijer found that authenticity of records is protected by a combination of technological, organizational (division of tasks), and institutional (norms, values, cognitive scripts) safeguards. Accountability fora need authentic digital records to reconstruct actions and decisions of government officials and organizations and are willing to rely on perceived or stated safeguards, and only question the authenticity of records if they are confronted with clear evidence of tampering.

Little has been done since these studies to map the knowledge gained through research into the practice of records professionals. It is this author's premise that, despite strides in knowledge and awareness of digital records issues among records professionals, and complex research into authenticity models as part of preservation research, the gap between research and practice still exists and may be widening.

2.8 The Role of This Study

Evaluating authenticity lends a measure of confidence, stability, and fixed reference points – that is, evidence of trustworthiness (MacNeil 2001, 42). Metrics of authenticity are one measure of confidence. An assessment of authenticity relies on both structural assurances and situational normality (McKnight and Chervany p. 37-38). Current research is actively pursuing models of authenticity measures (Salza et al. 2012; Guercio and Salza 2013), secure provenance (Hasan, Sion, and Winslett 2007; Lu et al. 2010), and preservation-aware storage (Factor et al. 2009).

What is lacking, however, is a measure of how professionals are handling authenticity of digital records on a day-to-day basis. Park's work of more than a decade ago demonstrated that research and practice were far apart, and the continued research focus on and concern about digital records' authenticity would suggest that this has not changed. It is important to ask what practitioners believe is important to ensure and protect authenticity and what means or indicators they are using and relying on.

3 Research Design

3.1 Introduction

Research designs are procedures for collecting, analyzing, interpreting, and reporting data in research studies (Creswell 2011, 54). The development of a research design involves a complex hierarchy of decisions and commitments to the pursuit of knowledge. At the outset, research begins with a "real-life issue that needs to be addressed, a problem that needs to be solved, a question that needs to be answered" (Crotty 1998, 13). Crotty suggests that the first questions to ask in designing a research project are: 1) what methods and methodologies will be employed in the proposed research, and 2) how will the choices be justified. This leads through a thought process that moves from the specific (research question) to the general (worldview), where each element is informed by the one above it (Figure 1, based on (Crotty 1998, 2)).

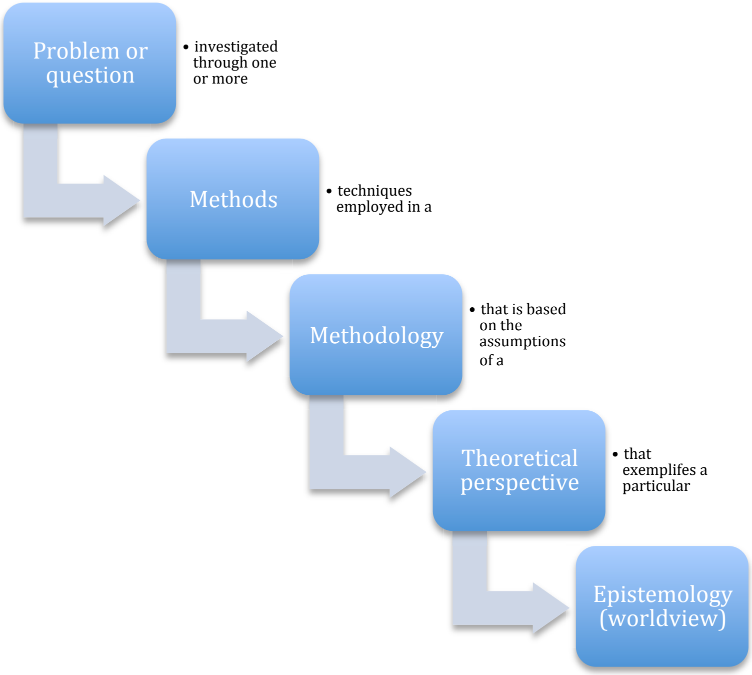

Problem or question
• investigated through one or more
Methods
• techniques employed in a
Methodology
• that is based on the assumptions of a
Theoretical perspective
• that exemplifes a particular
Epistemology (worldview)

Figure 1: Elements of research design

The description of the research, then, begins with a detailed account of the methods of data collection and analysis. Each method chosen (e.g. participant observation) is a technique of data collection or analysis appropriate to a certain strategy or plan of action – a methodology (e.g. ethnography). In turn that methodology is reflective of, and justified by, a particular theoretical perspective or philosophical approach that informs the methodology – an approach to understanding the human and natural world (e.g. symbolic interactionism). This theoretical perspective derives from an epistemological point of view, a set of assumptions about the nature of reality and "how we know what we know" (e.g. constructionism) (Crotty 1998, 1–8).

Most discussions of research design begin at the level of epistemology (cf. Bryman 2004; Creswell 2009; Creswell 2011). How can a method of data collection be known or justified separately from a particular worldview and theoretical framework? These elements of research design are intertwined and inextricably linked by the intention of the research expressed in the research questions, as well as the assumptions underlying it. Thus we see that the four elements also inform one another in a move from the general to the specific (Crotty 1998, 4; Creswell 2011, 39) (Figure 2).

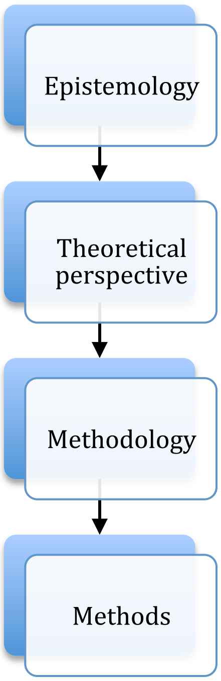


Figure 2: Elements of research design

While the methods and methodology will be chosen on the basis of the research question and what the research is intending to explore, explain, demonstrate, or prove, the explanation of research design once those decisions are made flows most clearly in a top down manner. In this way, for example, the researcher can explain the logic of the research question in terms of the epistemic and theoretical perspective of the researcher.

This study is undertaken in a pragmatist paradigm (epistemology or worldview), using the theoretical perspective of archival diplomatics, interpreted through the lens of practice theory. It employs a mixed methods sequential methodology, using quantitative data collection (survey-questionnaire) and descriptive statistical analysis and narrative analysis, followed by qualitative data collection (semi-structured interviews) and narrative analysis employing the technique of constant comparison.

3.2 Worldview – Pragmatism

Research involves philosophical assumptions – ontological and epistemological – about the nature of the world and the way in which we can know the world, as well as choices of distinct methods or procedures. These are the "basic set of beliefs that guide action," consciously or unconsciously (Guba 1990, 17 quoted in; Creswell 2009, 6). Social science researchers agree that such assumptions, paradigms (Kuhn 1970), or worldviews (Creswell 2009) influence the formation of research questions, and the design and methodology of research. Although terminology varies among research methodologists, it is possible to talk about two dominant philosophical traditions, positivism and interpretivism (Bryman 2006, 11–16; Brannen 2005, 7). According to Kuhn, these paradigms are incommensurable. Quantitative research strategies are generally associated with positivist traditions, while qualitative research strategies are associated with interpretivist traditions. Quantitative researchers, operating from the position of statistical generalizability, criticize qualitative research for being too contextually driven, unrepresentative and not generalizable. Qualitative researchers, on the other hand, judge quantitative research as overly simplistic, reductionist, and incapable of capturing the rich meaning inherent in personal experience (Brannen 2005, 7). In practice, social science research cannot be so neatly dichotomized (Crotty 1998, 14–17; Brannen 2005, 7).

Over the past decade there has been a resurgence of interest in social science research in another paradigm or philosophical worldview – that of pragmatism. Pragmatism emphasizes the research problem above the methods, and employs a pluralistic approach to acquiring knowledge about the problem (Creswell 2009, 5–10). Pragmatism "sidesteps the contentious issues of truth and reality, accepts, philosophically, that there are singular and multiple realities that are open to empirical inquiry and orients itself toward solving practical problems in the 'real world'" (Feilzer 2010, 8). It has garnered interest particularly within organizational and informational studies as a viable alternative to the 'paradigm wars' that pit positivism against anti-positivism (interpretivism), and has been promoted as a useful paradigm for conducting Information Systems (IS) research that studies information systems in organizational settings (Goldkuhl 2004, 13).

The American philosophy of pragmatism was formulated in the latter decades of the 19th century and early decades of the 20th century by Charles Sanders Peirce (1839–1914), William James (1842–1910) and John Dewey (1859–1952). Although each philosopher had his own particular perspective, there were two common threads. The first was the pragmatist maxim – a rule for clarifying the contents of hypotheses by tracing their 'practical consequences'. The second was a shared rejection of a Cartesian, atomistic approach to the norms that govern inquiry (Hookway 2015).

Goldkuhl's description of pragmatism for IS research is inspired by the classic American pragmatists and applies well to this study. Pragmatism's foundation is in empiricism broadened beyond a "pure orientation" (p. 13) to observation of a given reality based in human action. The primary unit of analysis, therefore, is action, and this primary concern for action guides the research inquiry. This focus on action as the basic unit of analysis leads to a study on what humans do – what actions they take for a particular purpose, and how those actions can be described and categorized. Actions can be distinguished as social or instrumental (material). Instrumental actions are mediated through material tools (e.g. technology) or immaterial tools (e.g. linguistic concepts). Further interest is placed on how these social and instrumental actions are managed in the context of practice – a network of actions related and combined in some meaningful way ("embodied, materially mediated arrays of human activity centrally organized around shared practical understanding" (2004, 17, citing Schatzki 2001, 2)). Furthermore, the meaning of an idea or concept is defined in terms of the practical consequences of that idea or concept. Pragmatism avoids abstract conceptualism that does not link the concept to an actable world (Goldkuhl 2004, 13–26). For example, in the present study, the concept of "record authenticity" can be examined in terms of the consequences of a record being authentic, based on related actions, undertaken as part of professional practice.

Research enquiry in a pragmatist paradigm, then, focuses on observation of people's actions and their means of acting within a community of practice, and the contexts and results of those actions. This suggests that the following types of questions may be asked:

- what is being done? (what *action* is performed)
- who is doing? (who is the *actor*)
- what is done? *result*
- when is it done? *time-context of the action*

- towards whom is it done? *receiver of the action*
- what should this action lead to? *what are the intended effects/purposes*
- what was unanticipated during the action? *unintended effects arising from the action*
- how the action aided? *what instrument is used*
- what is transformed in the action? *what is the base to be transformed in the action*
- who initiates the action? *assigner* to the action
- what initiative is there to this doing? *assignment* governing the action
- what kind of *knowledge* is used in the action?
- what kind of result is valued as good? *norms* governing the action
- who decides what is good? *norm-framers*
- what is *learned* through the action (ibid.)

The three main research questions guiding this study can be discussed from the perspective of this pragmatist worldview.

1. What elements of the context, content, and structure of digital records and data do records professionals rely on in order to determine their authenticity?

    a. This question encompasses the idea of action (ensuring and/or assessing authenticity) undertaken through social and material instruments (indicators of authenticity and their relation to context, content, and structure) in the context of professional practice in order to achieve a goal (record authenticity). Whether a record can be assessed as authentic or not has practical consequences (anticipated and unanticipated) for the organization or individual who uses and/or relies on it.

   b. Sub-questions may be posed that investigate in more detail the actors involved (those who have an interest in ensuring or maintaining authenticity as well as those who have an interest in using records they can presume authentic) and the time context of ensuring and assessing authenticity.

2. Are the traditional archival models of authenticity of records still used in the digital environment and to what degree?

3. Are the traditional archival models of authenticity of records considered sufficient to support a presumption of authenticity in the digital environment over time and technological change?

   a. These two questions address the kind of knowledge that is needed and used, how results are judged as good, and who makes that determination.

3.3 Methodology

This study employs a mixed methods approach. “Mixed methods research has been hailed as a response to the long-lasting, circular, and remarkably unproductive debates discussing the advantages and disadvantages of quantitative versus qualitative research as a result of the paradigm ‘wars’” (Feilzer 2010, 6). Pragmatism eschews methodological orthodoxy in favour of methodological appropriateness, recognizing that different methods are appropriate for different situations (Patton 2014, 72). As such, it is the philosophical underpinning of a mixed methods design.

A mixed methods research design combines elements of qualitative and quantitative research approaches (e.g. use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for achieving breadth and depth of understanding and corroboration (Johnson, Onwuegbuzie, and Turner 2007, 123). This design was chosen for several reasons: quantitative and qualitative data collection and analysis can complement one another (Greene, Caracelli, and Graham 1989, 259), enhance credibility of findings, and provide a more complete account, greater contextual understanding, and richer description of the topic of inquiry (Bryman 2006, 105–107). It can also facilitate selection of research participants, as will be described below.

There are several ways to construct a mixed methods design. They may be fixed or emergent. In a fixed design, such as the one used for this study, the decision to employ quantitative and qualitative data collection techniques was made at the outset because of the nature of the research purpose and questions. The quantitative and qualitative data collection may proceed simultaneously or sequentially. A fixed, explanatory sequential design was adopted for this study (see Figure 3). The design occurred in two phases, or strands. It began with the collection of quantitative data from a broad cross-section of the research population. The research subjects and questions for the qualitative strand that followed were respectively identified and articulated through the first strand. The protocol was derived from the results of the quantitative data collection, in order to explore those data in greater depth in response to the research questions. Therefore the two strands interacted at several points in the research process. Each strand was of equal importance or priority within the design. The rationale for this design was that the quantitative data were meant to provide an overview of professional practice in ensuring, presuming, assessing, verifying, and protecting authenticity of digital records and data, while the qualitative data were intended to offer greater depth of understanding of the practice of individual professionals.

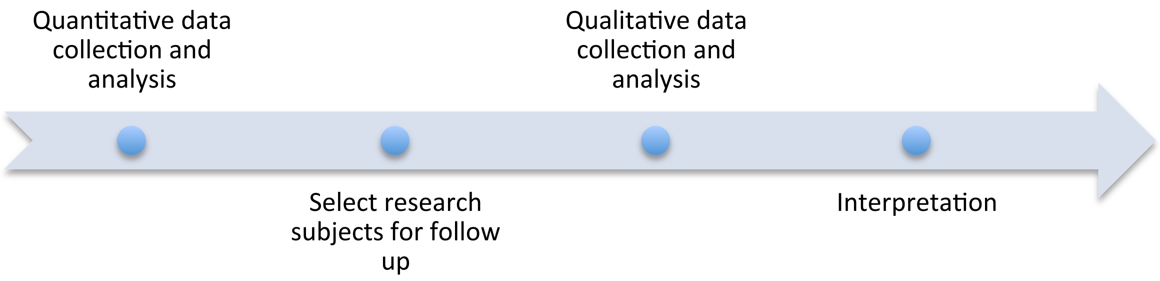
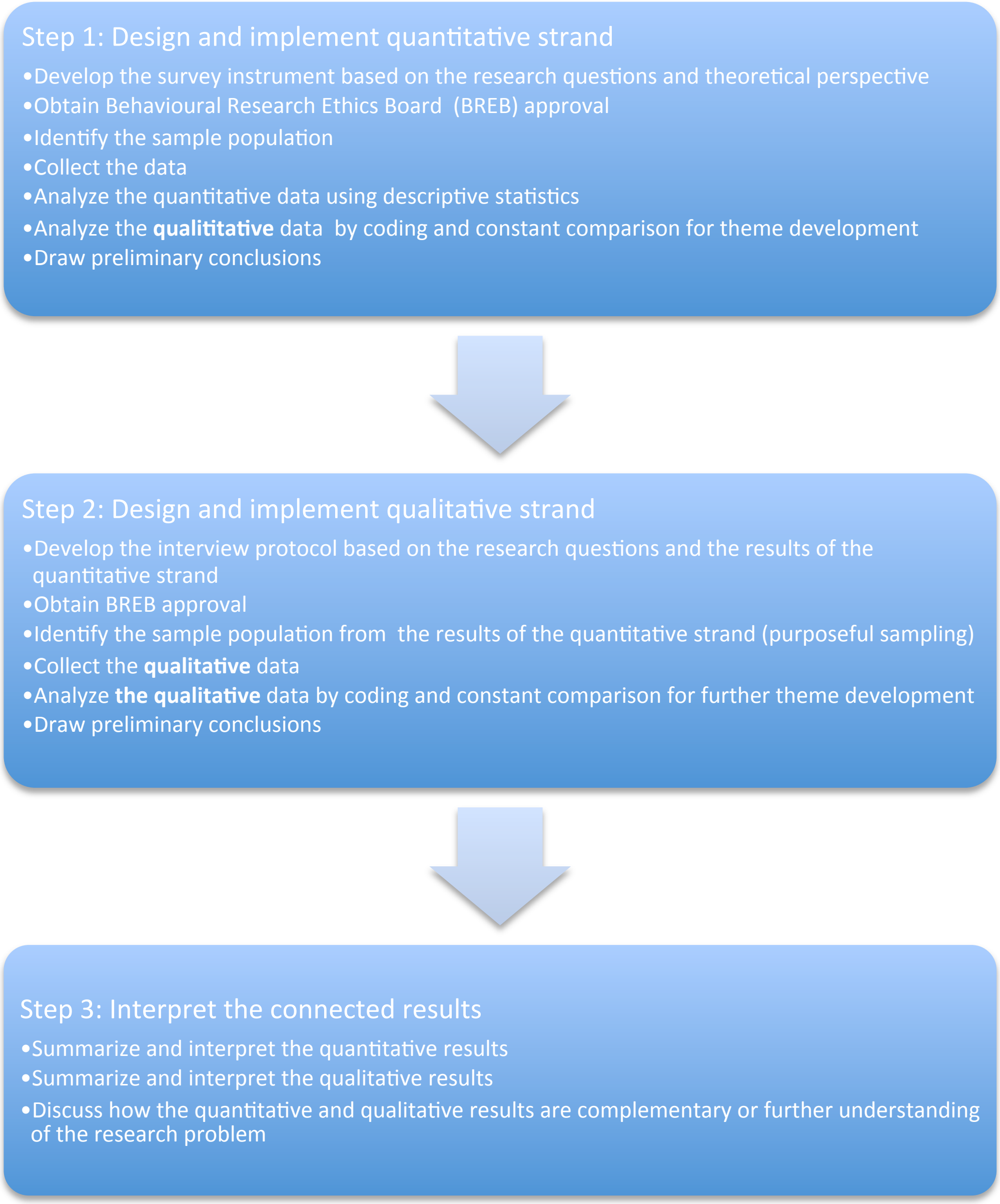

Quantitative data collection and analysis
Select research subjects for follow up
Qualitative data collection and analysis
Interpretation

Figure 3 - Explanatory sequential mixed methods design

3.4 Methods

This study uses both quantitative and qualitative methods of data collection and analysis. Quantitative data were collected through a survey, a questionnaire consisting of closed and open questions that was sent to professional listservs. The data from the closed questions were analyzed using basic statistical methods, while the data resulting from the narrative questions were coded based on a process of open coding using constant comparison to identify emerging themes (Bryman 2004, 404). Prior to coding the answers to the open-ended questions in the first strand a basic set of initial codes was developed based on definitions of record authenticity developed in the context of archival diplomatics, and this set was expanded and refined through the coding process by a process of constant comparison. Qualitative data were collected through semi-structured interviews that were developed based on questions arising from the answers to the questionnaire in order to explore its results in greater depth. These data were also coded using the codes developed in the first strand and analyzed using a process of narrative analysis. The interpretation was done using analysis from both strands. The stages of the research process are outlined in Figure 4 below (based on Creswell 2011, 84).



Step 1: Design and implement quantitative strand

- Develop the survey instrument based on the research questions and theoretical perspective
- Obtain Behavioural Research Ethics Board (BREB) approval
- Identify the sample population
- Collect the data
- Analyze the quantitative data using descriptive statistics
- Analyze the **qualititative** data by coding and constant comparison for theme development
- Draw preliminary conclusions

Step 2: Design and implement qualitative strand

- Develop the interview protocol based on the research questions and the results of the quantitative strand
- Obtain BREB approval
- Identify the sample population from the results of the quantitative strand (purposeful sampling)
- Collect the **qualitative** data
- Analyze **the qualitative** data by coding and constant comparison for further theme development
- Draw preliminary conclusions

Step 3: Interpret the connected results

- Summarize and interpret the quantitative results
- Summarize and interpret the qualitative results
- Discuss how the quantitative and qualitative results are complementary or further understanding of the research problem

Figure 4 - Stages of the research process

3.4.1 Quantitative Strand – Survey

3.4.1.1 Survey Tool

The survey was run using Qualtrics software, an online survey tool provided by the University of British Columbia Faculty of Arts. Qualtrics is a private research software company based in Utah, U.S.A. This tool is approved by UBC for collection of non-sensitive research data.

3.4.1.2 Research Population and Research Frame

The research population chosen for the survey was records and information professionals. This is a very broad population that includes archivists, records managers, information managers, data preservation or curation professionals, metadata librarians, to name but a few. Their subjects of professional focus may be organizational or individual records, documents, scientific or social data, images or audio-visual material, or simply ‘information’. However, professional interests coalesce around common concerns that are discussed in a variety of discipline-specific listservs. The sample frame chosen, therefore, was that of professional listservs of particular interest to archivists and records managers. The intended scope was international, reaching English-speaking professionals working across sectors, jurisdictions, and legal and regulatory systems.

As the first strand of a mixed methods study, the survey respondents became the sample frame from which interview subjects were chosen. The survey asked respondents if they would be willing to participate in an interview to further explore their responses. The interview candidates were chosen from those who agreed. This is discussed in the section 3.4.2.

3.4.1.3 Survey Panel

Listservs were chosen based on the following criteria: they were

- moderated;
- subscribed to or hosted by national/international archival and records management bodies or research alliances focusing on records-related issues;
- open to domain professionals and interested others; and
- containing active discussions of records management or archival issues involving both theory and praxis.

Listservs devoted to archives, records management, digital preservation, and research data were chosen from North American, European, and Australian scholarly, professional and research communities. The estimated reach of these listservs was several thousand individuals. The complete annotated list is found in Appendix 1.

3.4.1.4 Data Analysis

Raw data were exported from the Qualtrics survey tool and prepared for analysis in Excel. Analysis was conducted in Excel and Stata Data Analysis and Statistical Software (www.stata.com). Details of the analysis follow in Chapter 4.

3.4.1.5 Limitations of Web-based Surveys

Web-based surveys conducted through professional listservs are convenient ways in which to reach a specialist population quickly, but they do have limitations. Primary among these limitations is the inability of guaranteeing a representative sample. Individuals self-select to join a listserv, and not all members of a listserv will choose to participate. Public listservs have a particular focus and purpose, but are open to all interested individuals who may or may not have professional domain knowledge about the target topic. Even when using professional listservs (where the members can be reasonably assumed to share a degree of common purpose, training, and responsibilities), respondents opt in to the survey, and while all respondents may be members of the target population, not all members of that population are members of, have access to, or read these listservs. Results may be practically significant, inviting interesting discussion of the questions, but as there can be no assurance of a truly random sample, results may be difficult to generalize, nor can validity be objectively measured. However, as an indicator of general practice, such surveys provide useful information.

3.4.2 Qualitative Strand - Interviews

3.4.2.1 Qualitative Interviews

Qualitative interviewing is a technique of data collection that encompasses research interview styles ranging from semi-structured interviews guided more or less strictly by the researcher, to unstructured conversations between researcher and subject. They have in common their foundation in discourse or conversation, in "interviewer-respondent interaction," as opposed to "stimulus-response" models derived from survey questionnaire techniques (Mishler 1986, 16). Research interviews vary in the degree of structure, openness of purpose, and emphasis on exploration or theory building versus hypothesis testing. Mishler distinguishes between interview techniques grounded in survey research, which derive from quantitative procedures, and interviews as language-centred, context-relevant verbal interchanges between researcher and respondent. The interviews conducted in the course of this research were based on the results of the questionnaire and are therefore of the first type, although the opportunity provided by one-to-one real-time conversation allows for exploration of meaning and context with respect to participants' experience and beliefs.

3.4.2.2 Research Sample

Purposive sampling was used to identify interview subjects. Purposive sampling is a strategic approach intended to establish correspondence between the research questions and the sample (Bryman 2004, 334). The sample was drawn from the survey respondents to maintain coherence in the data, and sampling continued until theoretical saturation was achieved. Theoretical saturation is achieved when no new or relevant data seems to be emerging in desired categories, which are themselves well developed (Bryman 2004, 305, citing Strauss and Corbin 1998).

3.4.2.3 Data Analysis

Interviews were recorded and then transcribed. Analysis was conducted on the transcriptions in TAMSAnalyzer, a free, open source, qualitative data analysis software tool created by Dr. Matthew Weinstein, a Professor at the University of Washington, Tacoma. TAMSAnalyzer is available at http://tamsys.sourceforge.net.

3.5 Theoretical Perspectives

The issue of theoretical perspectives has been left till the end because of the nature of mixed methods design. Several theoretical lenses have been brought to bear on this study, depending on methods of data collection and analysis. When looking at record authenticity, two perspectives (at least) are available:

- The first is the perspective of archival diplomatics, that recognizes records as definable entities that can be reified according to a generalized conceptual model;
- The second sees records as complex, socially constructed representations, part of human activity systems.

These two views can each be brought to bear in this study, as will be discussed in Chapter 6. This is the essence of the pragmatist philosophy that attempts to settle apparently conflicting points of view by tracing their respective practical consequences (James 1907 Lecture II).

Theory is a contemplation or speculation about, a view of, or a perspective on the nature of something, that, when associated with disciplined knowledge building, results in a set of propositions empirically or experimentally derived (Eastwood 1994, 123). If the purpose of this study is to examine the actions of records professionals with respect to a particular focus of their practice – that is, the actions of ensuring, maintaining, and assessing the authenticity of records, then the starting point for such an inquiry must be a view or perspective on the nature of record authenticity. This theory provides a framework for understanding records and their various qualities, and for understanding archival and records management practice. At the highest abstraction, archival theory posits that records attest to facts and acts, and their trustworthiness is dependent on the circumstances of their generation and preservation (Eastwood 1994, 126). This postulate is expressed in theoretical conceptualizations of record and record authenticity. This study adopts the model developed from the perspective of archival diplomatics and operationalized through the first two phases of InterPARES. Such model guided the quantitative strategy – the development of the survey questionnaire, and by extension, informed the qualitative aspect – the interview protocol.

This author is aware of the potential challenges of adopting this perspective. "Traditional notions of impartiality and authenticity have become the lightening rod of criticism of traditional [archival] ideas" writes Eastwood (2010, 18). Critics of traditional archival theory reject the possibility of certain knowledge or truth inherent in records and assessments of one true view of authenticity (cf. MacNeil and Mak 2007; Duncan 2009; Mak 2012). However, this author believes it is possible to respect and adopt the traditional view as a lens for exploring and interpreting current practice. The goal is not to force current practice to adhere to a theoretical ideal, or to test the validity of that ideal, but to chart the course from this ideal to current reality and ask questions about it. This study charts the path from the ideal to the real, examining the gap between theoretical and practical research findings about the nature of record authenticity and professional practice which is the “less formal theory that … reflect[s] the untidy reality” (Buckland 1994, 348).

This leads to the question of analysis, and the theoretical lens through which analysis is done. The qualitative strand of the study explored some of the complexities of the topic that were not apparent through the survey questionnaire. While the theory of archival diplomatics provides the framework for understanding record authenticity and its mechanisms (the unit of analysis being the record), the focus of the study on professional practice and belief (the focus of analysis being actions on the record) demands a framework for understanding practice. Many social theories espouse the complexity of real life social situations. Three theoretical perspectives have particular resonance with this author in the context of this study and for future research. They are soft systems methodology (Checkland and Scholes 1999; Checkland and Poulter 2007), practice theory (cf. Schatzki, Knorr-Cetina, and Savigny 2001), and actor network theory (Law 1992; Latour 2005; Law 2006).

Changing worldviews and purposeful action are behind soft systems methodology (*Peter Checkland on the Origins of SSM* 2012). Soft-systems thinking would approach the issues of authenticity through evaluation of the social contexts of record creation, use, and preservation (e.g. the presence of policies governing records and record systems), and the collection of provenance information, and the social means or indicators by which authenticity is ensured. It recognizes the “irreducible complexity” (Checkland 1999, 90), messiness and heterogeneity (Law 2006) of real world situations. A hard-systems approach, in contrast, takes a structured problem-solving approach (Checkland

1999, 130–136), an algorithmic approach that seeks metrics as an evaluative tool. Cryptographic validation techniques, for example, are a product of hard systems thinking, particularly persuasive to advocates of computer technology for purposes of security, control, and information assurance. This study explores the dichotomy between and influence of hard- and soft-systems thinking in discussing indicators of authenticity: is one approach generally prevalent over the other? is one approach favored by certain records professionals? and so on.

“Practice,” according to Schatzki, shares equal consideration with concepts such as “structure”, “system,” and “action,” among others. However, there is no unified practice approach, although investigation of practice generally consider practices as arrays of human activity. A common belief is that phenomena of study occur within and are components of fields of practice. Practice theorists acknowledge that shared skills or understandings underpin practice (Schatzki, Knorr-Cetina, and Savigny 2001, 10–14). Actions are performed within a context of practice, and are determined by the practice of which they are a part (Goldkuhl 2004, 17).

The concepts of actor-network theory (ANT) are persuasive to providing another means to analyze the complexity of practice. Our intuition tells us that the world is a messy place. Simplification (abstraction) does not always help to understand mess, but actor network theory suggests that a disciplined lack of clarity might (Law 2006, 2). It allows a view of what is hidden, absent, or inconsistent in the face of the expected or typical (Law 2006, 10–11). ANT is particularly valuable in providing thick descriptions in case studies. While the ideas of ANT are not developed in the present study, its potential to provide a means to understand the complexity and “messiness” of current practice in ensuring and assessing authenticity of records is recognized for future research.

3.6 Summary

The research design is developed and conducted within a worldview of pragmatism that recognizes the appropriateness of quantitative and qualitative data collection and analysis techniques – namely, a mixed methods strategy. Data collection is sequential, with the quantitative strand preceding the qualitative strand. The research is conducted from the theoretical perspective of archival diplomatics, viewed through a lens of practice theory recognizing the presence of both soft and hard systems thinking and the complexity described by actor network theory. Figure 5 is a schematic of the research design, showing the rationale from the general to the specific. While the diagram appears to present a hierarchy, the research process is one of recursive and iterative development.

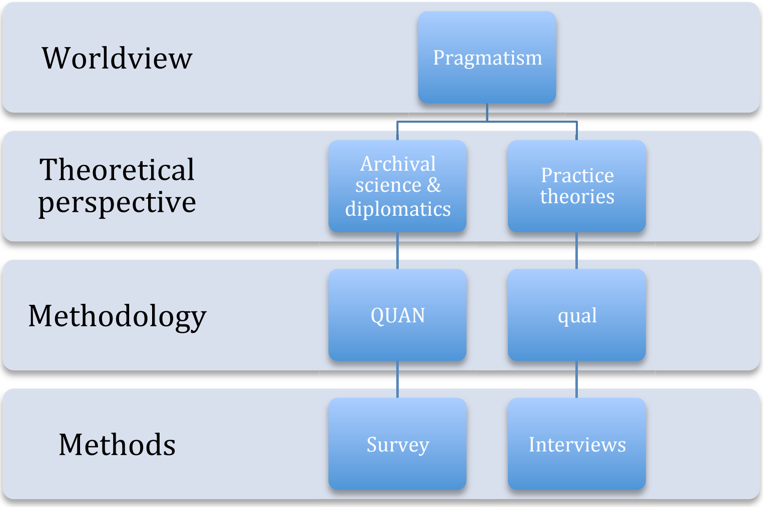


Figure 5: Schematic of research design

4 Survey of Records Professionals

4.1 Introduction

The purpose of the survey was to explore the relationship between practice and belief among records professionals – that is, what records professionals rely on in their work and whether their practice matches their belief or trust in specific authenticity indicators. The survey is found in Appendix 2.

The survey questions were designed to measure practitioners' notions of authenticity of digital records and data. They gathered basic information about the extent to which records professionals are concerned about the authenticity of digital records and data, how they ensure, assess, and/or protect authenticity, and the level of importance that they place on specific indicators of authenticity, both traditional indicators (e.g. records policies and procedures) and technological indicators (e.g. cryptographic validation).

Problems with terminology across sectors, disciplines, and countries are frequently cited in the literature. The survey sought to draw out individuals' conceptual understanding of authenticity through the design of the questions. Definitions of 'authenticity' and 'record' were not provided. This was done intentionally to avoid biasing or leading the responses based on one particular definition. The questions were determined based on domain knowledge, theory and praxis. They were reviewed by the author's dissertation committee and by several practitioners recommended by the committee, and revised accordingly prior to finalizing and posting the survey.

Data collection for this study began with a web-based survey questionnaire that ran from March 3-May 1, 2014. The survey was posted on major English-speaking archival and records management listservs, reaching professionals in North America, Europe, Latin America and Australasia.

The survey consisted of 17 questions of three types, organized in two sections. Section 1 included six basic demographic questions that established the respondent's type of employment (e.g. archivist, records manager, systems engineer) and the sector in which they work (e.g. government, telecommunications, cultural industries), location (country), age (range), level of education, and discipline of their degree(s). This allowed the results to be segmented based on employment type, job sector, predominant legal system (common law, civil law, or pluralistic), country (e.g. Canada, United Kingdom) and so on. Section 2 included nine questions about respondents' responsibilities and work practices, and their opinions regarding specific work practices in respect of guaranteeing, maintaining, and assessing authenticity of digital records. More specifically, these questions were intended to measure respondents' main professional responsibilities, the means they used to ensure authenticity, what metadata they routinely applied or relied on for what purpose, whether they have ever been called upon to make a formal attestation of authenticity in the course of their work and if so, what indicators had been most important in that attestation, and whether their organization explicitly defines authenticity in its policy instruments. Respondents were asked to rate the importance of each component in a 3- or 5-point Likert-style scale. These questions were further supported by two open-ended opinion questions asking respondents to give their own definition of authenticity and identify the indicators they felt were most important.

4.1.1 Indicators of Authenticity – Social and Technical

The set of indicators used was developed from the perspective of archival science and informed by the literature review. This set has been categorized into social (S) and technological (T) indicators of authenticity (Table 1).[8] This distinction is investigated throughout the analysis here and explored further in the follow-up qualitative interviews, discussed in the next chapter.

Social indicators are instruments developed by an organization to support the creation, management, or preservation of records (e.g. classification schemes, retention and disposition plans, policies and procedures documents). They are based on domain knowledge, and created and implemented by the intention of human actors (records professionals, management, legal counsel, etc.). They may or may not be present within a given organization; they may be mandatory or voluntary in their application or use, and even when mandatory, they may be circumvented or adapted, as Foscarini showed in her study of central banks (Foscarini 2009). They include the foundational instruments of archival and records management practice: policy instruments, classification schemes or file plans, retention and disposition schedules, and archival description or other descriptive measures (which may be captured in varieties of descriptive metadata).

[8] In *Burdens of Proof,* Jean Francoise Blanchette investigates one cryptographic technique, digital signatures, as part of a technical as opposed to social foundation for authenticity and a new evidentiary regime (Blanchette 2012). The categories proposed in this study develop this concept.

Technical indicators are non-discretionary in their creation – that is, they are the result of a work process or state change in the records (e.g. system metadata capturing date created, and date modified), are algorithmically generated or implemented by the technological (e.g. computer, network) components of the overall record system (e.g. checksums, audit logs), are created to manage and control system access and security, or are created by a third party as specifications to a part of the technological system (e.g. documentation about software). Technological indicators may be used to control the records, but are more focused on controlling the system in which the records reside. They include audit logs, access controls and security measures, cryptographic validation techniques, and system metadata, as well as technical documentation. Indicators were listed in the questions in no particular order (although in the same order in each question that asked about them).

No.	Indicator	T or S
1	Written policies and procedures governing the management of the records system	S
2	Documentation about the record system (design, operation, management, etc.)	S
3	Written policies and procedures governing digital records	S
4	Information about the software used to create and manage the digital records	T
5	Information about changes made to the digital records over time, (e.g. migration, normalization, etc.)	T
6	Information about actions taken to preserve the digital records	T
7	Classification scheme and/or file plan	S
8	Retention and disposition schedules	S
9	Archival description	S
10	Access controls/security measures	T
11	Audit logs	T
12	Cryptographic validation techniques (e.g. digital signatures, hash digests, etc.)	T
13	Standardized metadata	T

Table 1 - Indicators of authenticity

4.2 Results and Analysis

4.2.1 Section 1: Data Preparation and Demographic Data

The survey received 441 responses and 293 completions (66.44%). The data were prepared for analysis in two stages. First, test results and incomplete results were deleted. Respondents were not required to answer all questions. A result was deemed incomplete if respondents answered only demographic questions and the second section of the survey remained unanswered. Because of the broad distribution of the survey to records professionals of all types and across knowledge sectors, the data in several categories were too few to be meaningful. In the next stage of data preparation, categories in question 1 (employment), 2 (knowledge sector), and 3 (country) were reviewed and consolidated for analysis as follows.

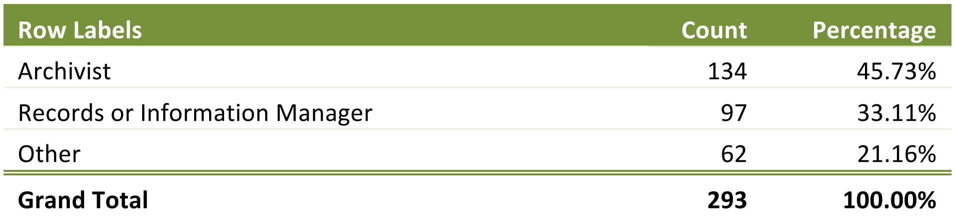
4.2.1.1 Employment Type

In the first question, respondents were asked to state their professional identity from a list of standard records-related disciplines (Table 2).

Employment categories	Count	Percentage
Archivist	125	42.66%
Compliance/privacy officer	3	1.02%
Conservator/curator	4	1.37%
Educator (e.g. professor, instructor, or trainer in an information field)	15	5.12%
Records or Information Manager	92	31.4%
Other	54	18.43%
Grand Total	**293**	**100.00%**

Table 2 - Employment categories – raw data

Responses in all categories were reviewed, and categories ‘Compliance/privacy officer’, ‘Conservator/curator’, and ‘Educator’, each of which received fewer than 20 responses, were merged with the ‘Other’ category. Respondents identifying as ‘Other’ included IT personnel (6), librarians (9), one lawyer, and several information sub-specialties, several of which were recategorized. This resulted in the following categories for analysis (Figure 6):



Row Labels	Count	Percentage
Archivist	134	45.73%
Records or Information Manager	97	33.11%
Other	62	21.16%
Grand Total	**293**	**100.00%**

Figure 6 - Final employment categories

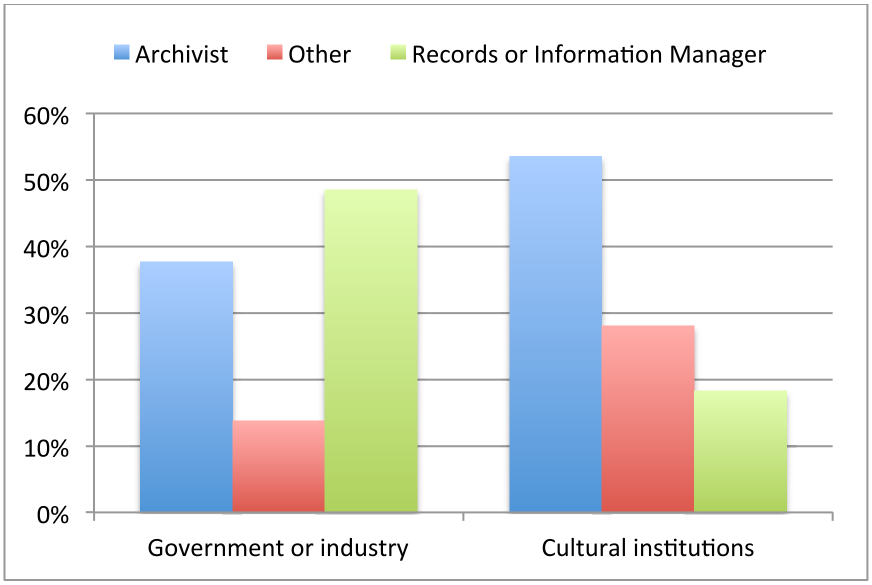
4.2.1.2 Knowledge Sector

Categories for knowledge sector were condensed from the North American Industry Classification System (Statistics Canada and Standards Division 2012) (Table 3).

Knowledge sector	Count	Percentage
Arts and museums	7	2.41%
Construction and manufacturing	3	1.03%
Educational services	52	17.87%
Finance	8	2.75%
Government or public administration	100	34.36%
Health care	5	1.72%
Information and culture (including libraries and archives)	72	24.74%
NGO	4	1.37%
Other	5	1.72%
Professional scientific and technical	22	7.56%
Resources	3	1.03%
Transportation	2	0.69%
Utilities	8	2.75%
Grand Total	**291**	**100%**

Table 3 - Knowledge sectors – raw data

As with the employment categories, many were simply too small to be meaningful for analysis and so were condensed. Knowledge sectors most represented were government and public administration, followed by information and cultural industries (including libraries and archives, broadcast and telecommunications), educational services, and professional, scientific, and technical services. For meaningful analysis, sectors were consolidated into two categories: ‘Cultural institutions’ and ‘Government and industry’. Figure 7 shows the distribution of employment type according to knowledge sector.

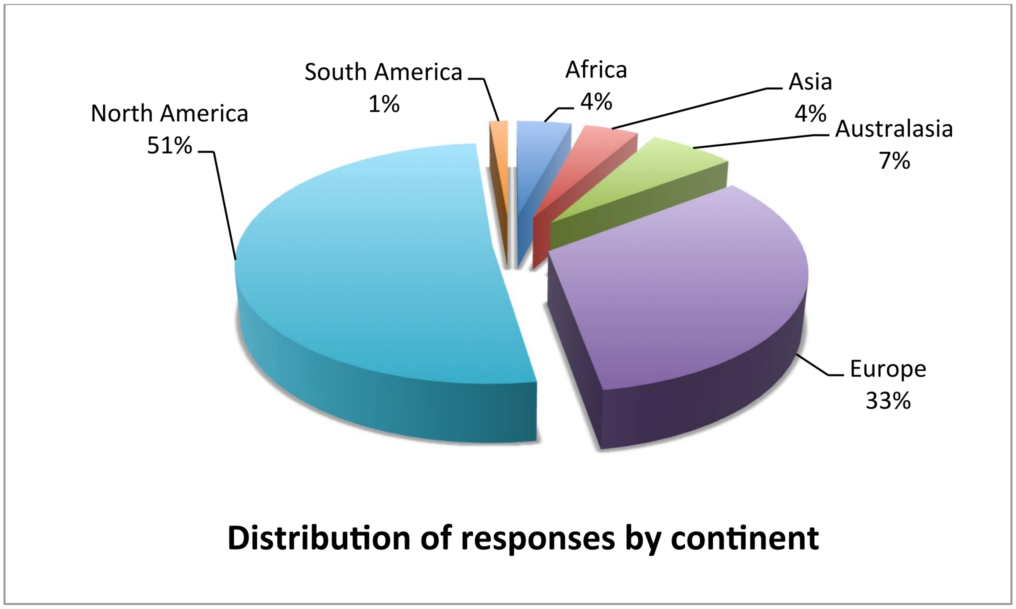


n=291
Figure 7 – Employment type by knowledge sector

4.2.1.3 Location and Legal System

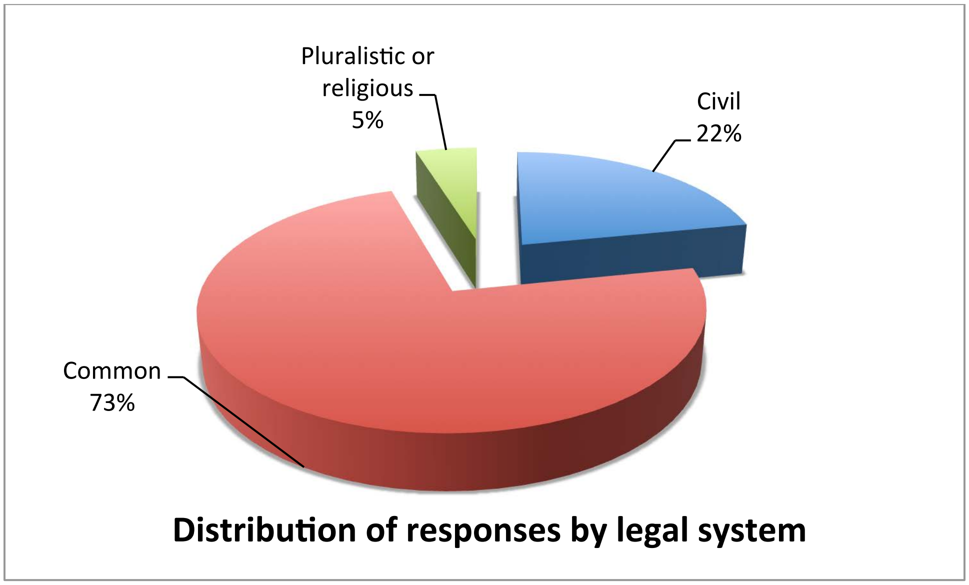
One hundred sixty-two respondents reported their location. By analyzing IP addresses and cross-referencing with email and narrative information, the countries of all but one respondent were identified. Responses were received from forty-six countries and territories on six continents (see Appendix 3).

Figure 8 shows the distribution of responses by continent. Most responses were from North America, with 34 (20.21%) responses from Canada, and 52 (30.14%) from the U.S. Of respondents from Europe, all but six were from member states of the European Union (Isle of Man, Russian Federation, Turkey, and Switzerland).



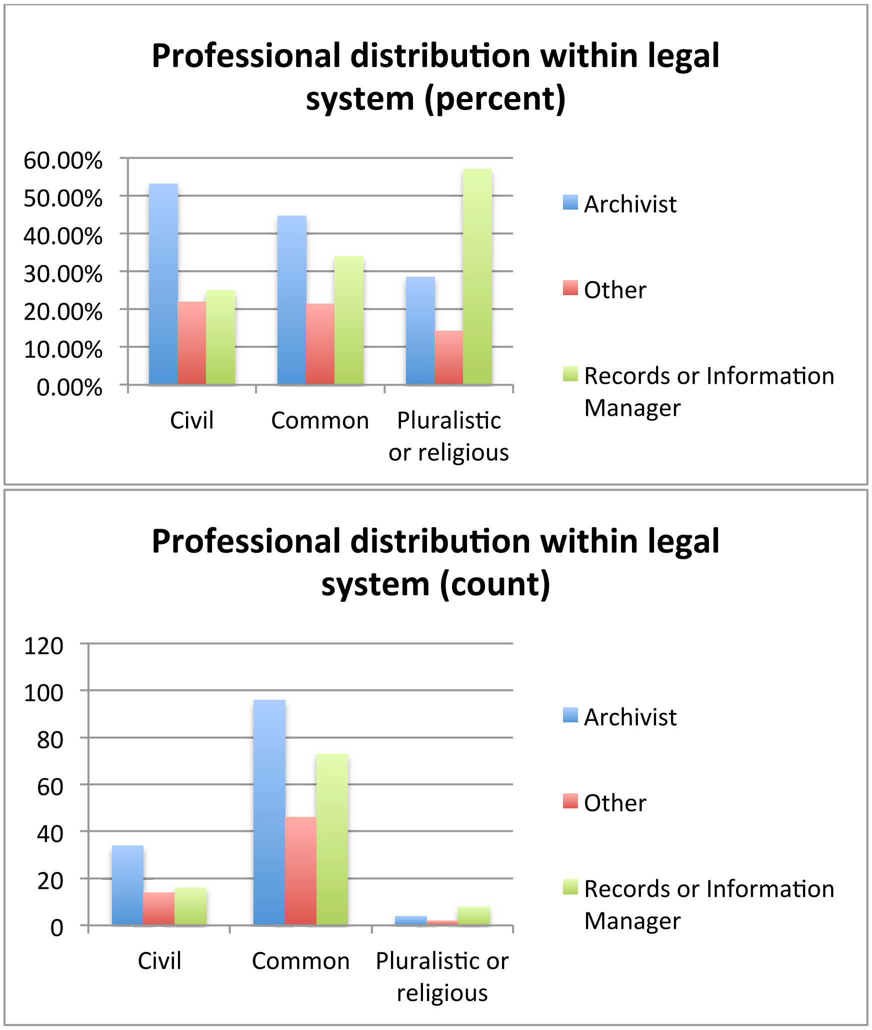
n=292
Figure 8 - Geographic distribution

Countries were also grouped for further analysis by predominant legal system (common law, civil law, pluralistic) (Wikipedia 2014) (see Appendix 3). The majority of respondents are working in common law legal jurisdictions (Figure 9).



n=293
Figure 9 - Distribution by legal system

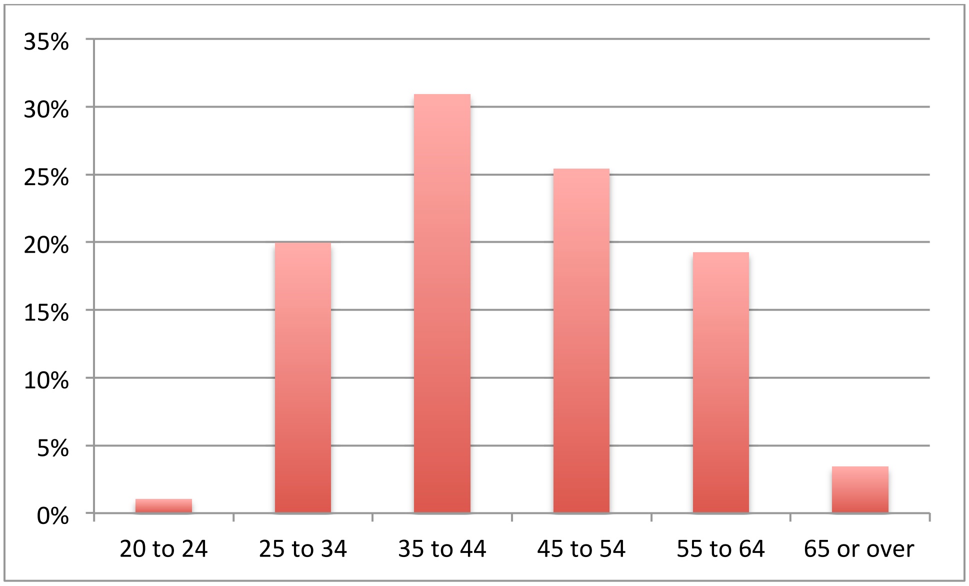
The split by profession in each legal system is shown in Figure 10. In common law countries the number of records managers and archivists is more balanced among survey respondents than in civil or pluralistic/religious legal systems. This may reflect the development of the profession in civil and common law jurisdictions, where the separation between records manager and archivist is seen most clearly in North America, the UK and Australia/New Zealand. Responses from pluralistic/religious systems are too few to draw any meaningful conclusions.



n=293
Figure 10 – Professional distribution by legal system (percent and count)

4.2.1.4 Age Range

The age distribution indicates that the majority of respondents were active career professionals (Figure 11).



n=291
Figure 11 - Age distribution

4.2.1.5 Education

The majority of respondents (66%) reported that the highest level of education received was that of a Master's degree (MA, MAS, MLIS) (see Figure 12). Respondents were asked about their field(s) of study, with possible responses being archival science, library and information science (LIS), computer science, law, history, and other. Most respondents listed degrees in several disciplines. The highest concentration of degrees was in archival science or LIS with an archival concentration (46% of respondents) and history (34% of respondents) followed by LIS (29% of respondents), computer science (7% of respondents), and law (1% of respondents). Two of the three respondents who had degrees in law also had degrees in archival science; the third listed no other degrees.

Degrees identified as 'Other' reflected a wide variety of humanities and social science disciplines. There were few degrees in formal or applied sciences represented.

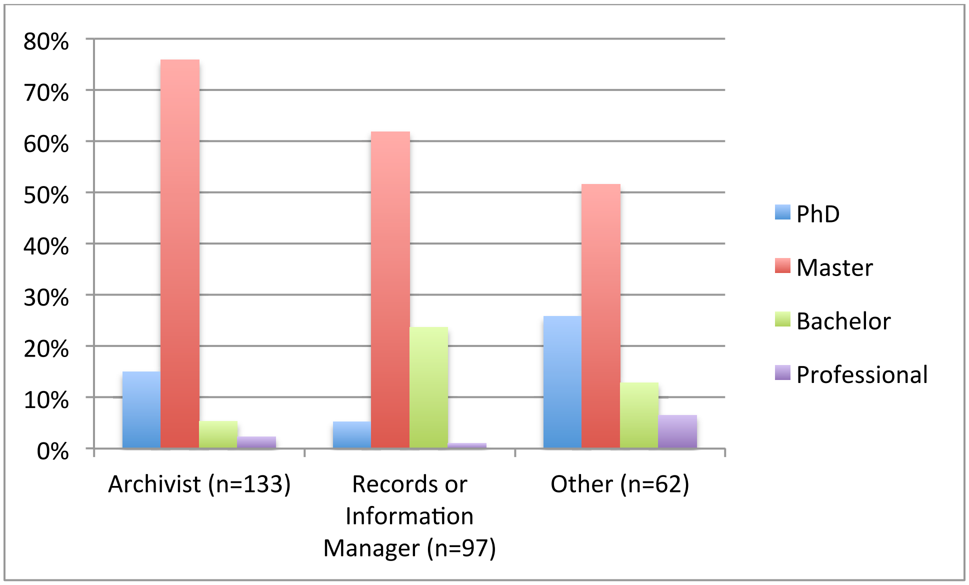


Figure 12 – Degrees by employment type

4.2.2 Section 2: Work Practice and Beliefs

4.2.2.1 Introduction

The main focus of the survey was to explore the relationship among practice, experience, and belief with respect to the use of and value placed on a set of indicators of authenticity. Questions measured who used different indicators most or least in the course of their normal work to ensure authenticity, how these indicators were used in the event of a formal attestation of authenticity, what indicators were *believed* to be most important in attesting to authenticity, and how frequently social versus technical indicators were invoked in use and in belief (a discussion of social versus technical indicators follows).

Results were segmented by employment type: archivists, records managers, and other, and by experience attesting to authenticity.

A key distinction between respondents is whether or not they have been required to make a formal attestation of authenticity in the course of their work. This allows a determination of whether work practice differs between those who have, and those who have not been required to authenticate a body of records. The differences between beliefs and practice is also explored by asking respondents who had not ever had to authenticate records what they think they would use if required to authenticate records.

At the end of this section, open-ended questions asked for respondents' own definition of authenticity, and what they considered essential indicators of authenticity. These questions give a view of the possible extent to which research findings have influenced professional work practice, and the ongoing validity (or otherwise) of traditional models of record authenticity.

Overall, results showed that social indicators were most commonly used to ensure authenticity in daily work, but that technical indicators were relied on in the process of authenticating records. This was consistent across employment categories. Experience affected belief in the value of indicators, with those who had never had to authenticate records placing more faith in technical indicators than those who had experience.

4.2.2.2 Work Tasks and Functions

Respondents were asked to rank ten tasks or functions according to the frequency with which they undertook them. Tasks were chosen based on common archival and records management functions:

- conduct retrieval and access;
- monitor or enforce security or access privileges;
- monitor or enforce privacy of personal information;
- monitor or enforce compliance with recordkeeping regulations or policies;
- conduct preservation or curation;
- design systems for storage and management or records;
- design information or records policies;
- manage records or information;
- manage or design metadata; or
- other.

Seventy-seven percent (76.55%) of respondents said they managed records or information often or very often, followed by retrieval and access (67.46%), managing or designing metadata (55.83%), and designing information or records policies (51.21%). The least frequently undertaken activities were monitoring or enforcing privacy of personal information (29.76%) and conducting preservation or curation (35.07%) (Figure 13)[9].

[9] Responses ‘never’ and ‘rarely’, and ‘often’ and ‘very often’ have been combined to more easily visualize the data in charts, unless doing so alters the results. When data is presented in tables, no summation has been done.

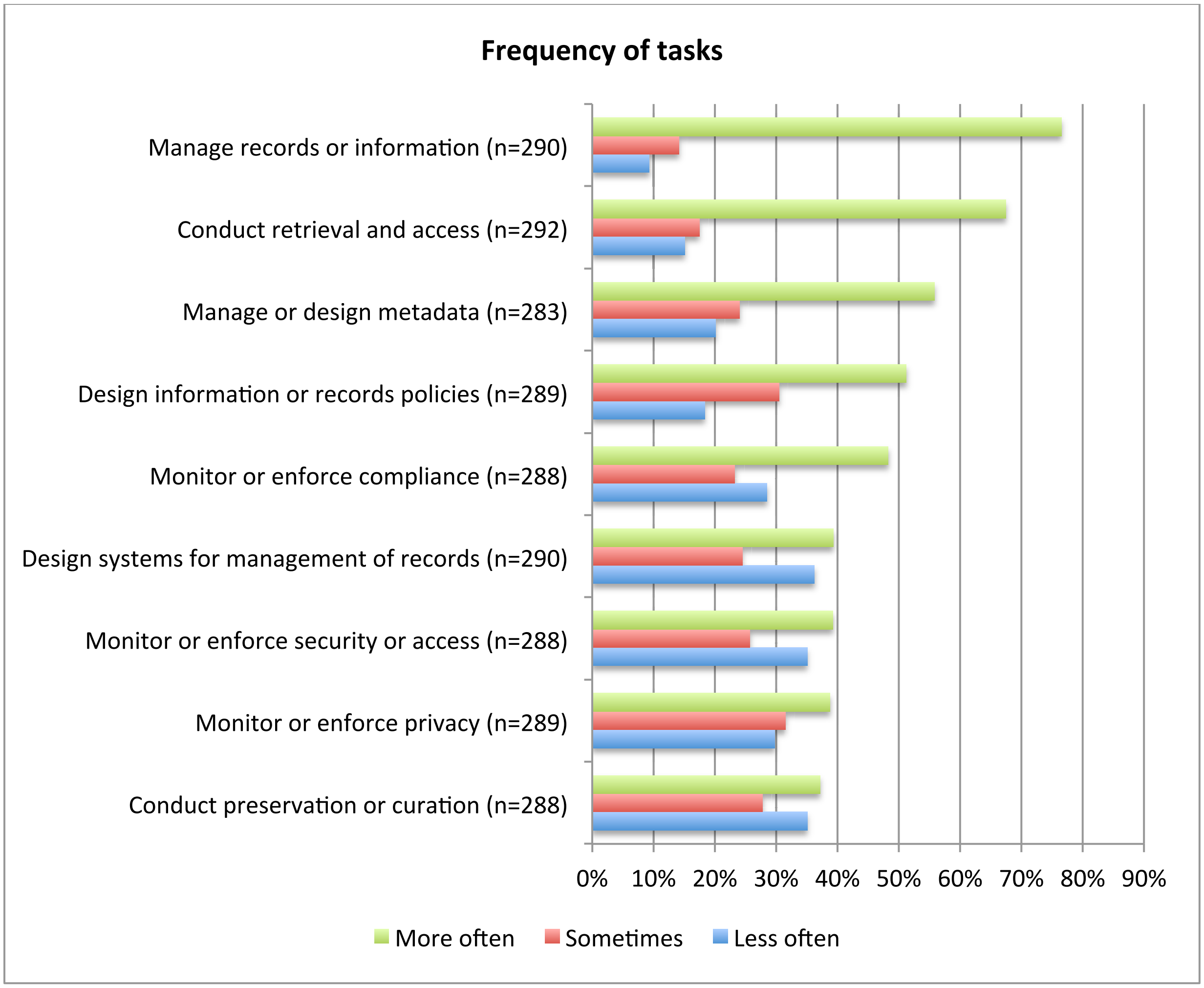


Figure 13 – Frequency of tasks undertaken[10]

Segmenting these data by employment type shows the difference between duties of archivists, records managers, and other (Figure 14). As expected, these differences reflect professional practice; however this chart highlights priorities of the group that does not self-identify as either archivist or records manager. The chart is ordered from most to least frequent task performed by 'other', and shows that while members of this category

[10] The size of respondent groups differs slightly for each variable because of the structure of the survey, which asked respondents to rate the frequency of each variable as a separate question. Responses were not mandatory to advance through the survey. Percentages are calculated for each variable based on the number of responses for that variable.

perform the same tasks as archivists and records managers, they design records systems and conduct preservation or curation more frequently than do their colleagues. This may reflect specialized positions or unique professional identities.

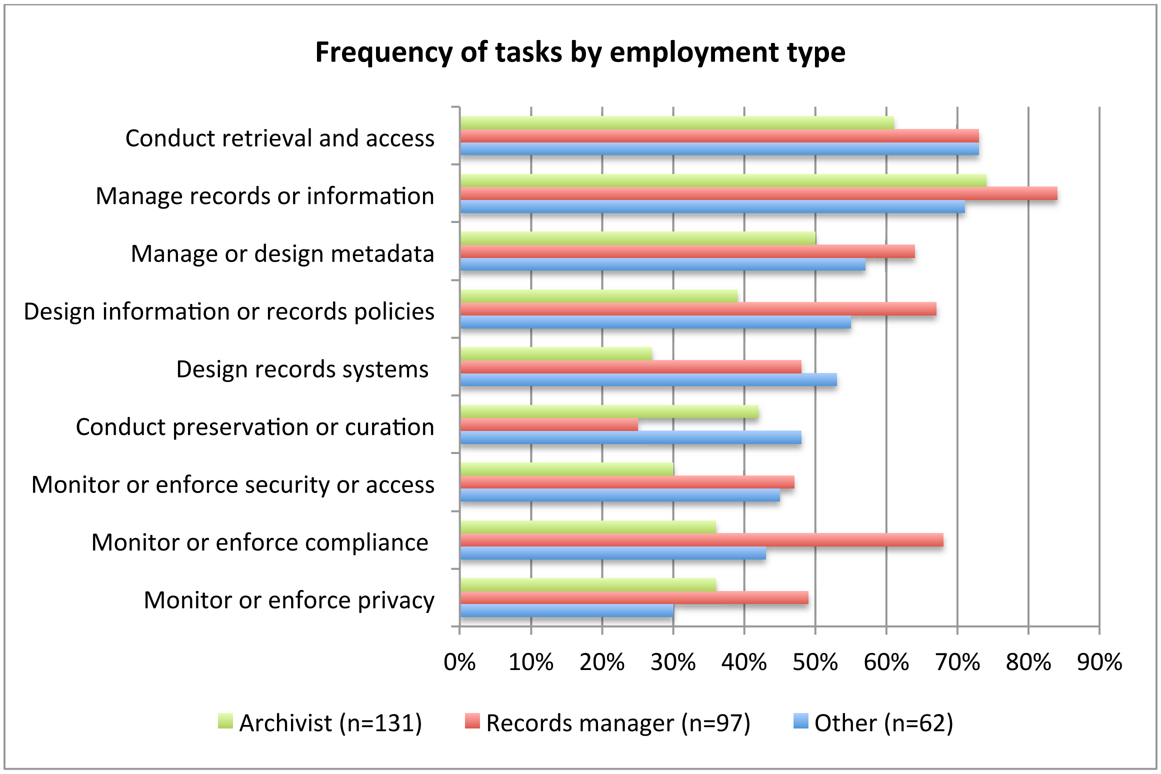


Figure 14 - Frequency of tasks by employment type

4.2.2.3 *Work Practice to Ensure Authenticity*

Respondents were then asked to rate indicators of authenticity according to how frequently they relied on them to ensure authenticity in the course of their work. As will be shown, this is at variance with respondent's expectations or beliefs about important indicators of authenticity.

Respondents reported using social or traditional indicators of authenticity most frequently to ensure authenticity. The use of a classification scheme was the most commonly used or relied on indicator (61.29%), followed by policies and procedures governing the records system (60.07%) and the records themselves (54.77%). Least often used or relied upon were cryptographic validation techniques (never or rarely used 60.71% of the time), audit logs (never or rarely used 51.1%) and information about preservation actions (never or rarely used 29.6%). The data presented in Figures 15 and 16 consolidates the two least often response choices (never, rarely) and the two most frequent response choices (often, very often) in order to more clearly show the results. The full data are presented in Table 4.

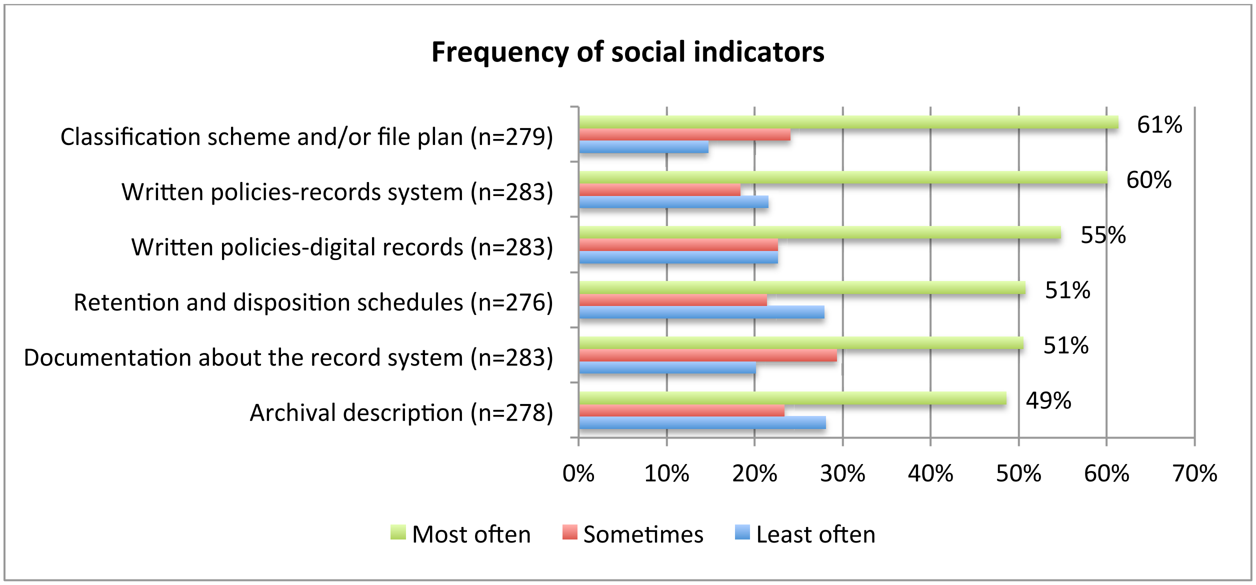


Figure 15 – Frequency of social indicators relied on or applied

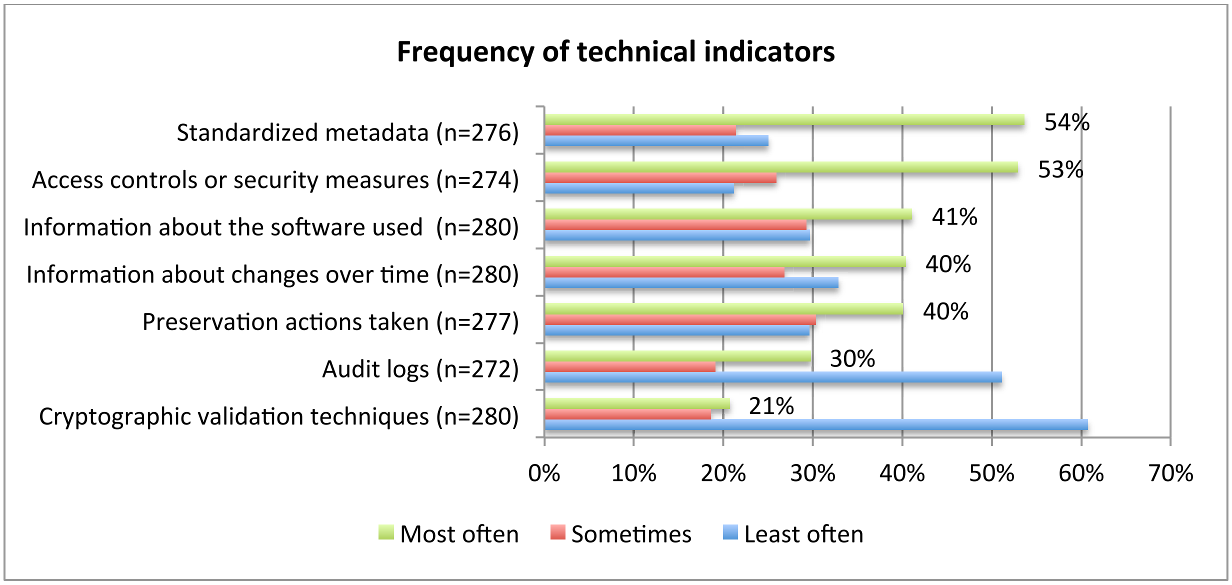


Figure 16 – Frequency of technical indicators relied on or applied

When you create or manage digital records, how often do you rely on or apply the following for authenticity (least used – most used):							
T/S	Indicator	Never	Rarely	Sometimes	Often	Very often	Total
T	Cryptographic validation techniques (n=280)	43.21%	17.50%	18.57%	10.71%	10.00%	100.00%
T	Audit logs (n=272)	25.37%	25.74%	19.12%	17.28%	12.50%	100.00%
T	Preservation actions taken (n=277)	13.00%	16.61%	30.32%	18.41%	21.66%	100.00%
T	Information about changes made to the records over time (n=280)	12.86%	20.00%	26.79%	21.07%	19.29%	100.00%
T	Information about the software used (n=280)	11.07%	18.57%	29.29%	23.57%	17.50%	100.00%
S	Archival description (n=278)	12.23%	15.83%	23.38%	23.02%	25.54%	100.00%
S	Documentation about the record system (n=283)	7.42%	12.72%	29.33%	26.15%	24.38%	100.00%
S	Retention and disposition schedules (n=276)	12.68%	15.22%	21.38%	22.46%	28.26%	100.00%
T	Access controls or security measures (n=274)	11.31%	9.85%	25.91%	25.91%	27.01%	100.00%
T	Standardized metadata (n=276)	11.59%	13.41%	21.38%	26.45%	27.17%	100.00%
S	Written policies and procedures-digital records (n=283)	8.48%	14.13%	22.61%	28.98%	25.80%	100.00%
S	Written policies and procedures-records system (n=283)	10.60%	10.95%	18.37%	31.80%	28.27%	100.00%
S	Classification scheme (n=279)	5.73%	8.96%	24.01%	27.24%	34.05%	100.00%

Table 4 - Frequency of indicators of authenticity in practice

4.2.2.4 Role of Employment Type in Choice of Indicators Used

As we have seen, social or traditional indicators of authenticity are relied upon more than technical indicators in ensuring authenticity in the course of normal work practice across employment types. To what extent do records managers, archivists, and other records professionals use certain indicators of authenticity more than others as a result of their work functions? Some indicators, for example archival description, we would not expect to see used with any frequency by records managers. This author wanted to measure which indicators were more commonly used by certain professionals than others in order to compare work practice. Figure 17 shows the distribution of frequency (often or very often) of indicators used by employment type.

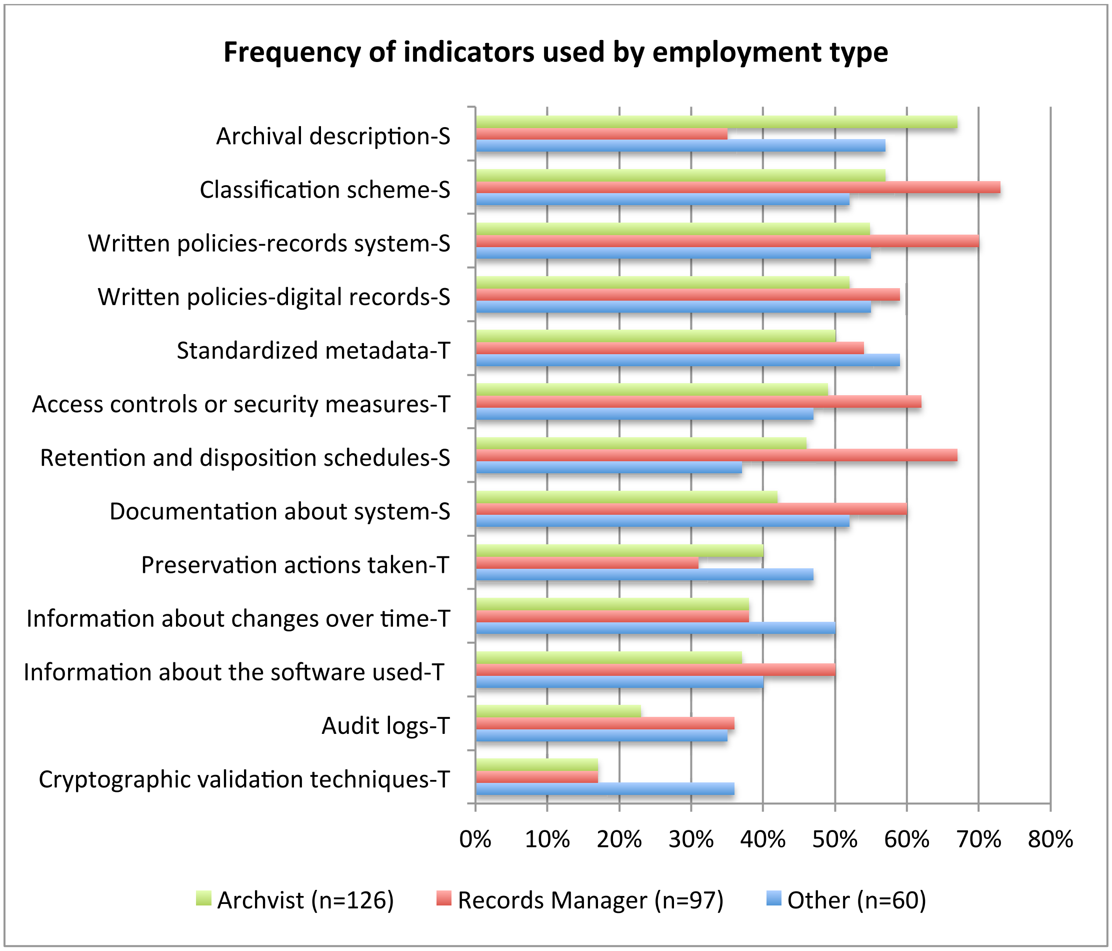


Figure 17 – Frequency of indicators used by employment type

This chart shows some expected differences (e.g. archivists rely on archival description more than other employment types; records managers use retention and disposition schedules more frequently). Other differences are less obvious (e.g. frequency of use of audit logs, cryptographic validation, system documentation). The degree to which these differences are statistically significant can be checked with a chi-square test for independence based on employment type. The results are reported in detail here for the indicator used most and least commonly across all employment types, classification code and cryptographic validation techniques, respectively (a summary of this significance test for all variables is found in Table 5.)

The results do not give evidence of a statistically significant difference (at standard confidence levels, $p < 0.05$) for the most commonly reported indicator of authenticity, 'classification scheme' across all employment types (p-value equals 0.101). (see Table 5). However, comparing archivists and records managers directly (setting aside the "other" group) using a difference of means test[11] for frequency of use of classification code confirms a statistically significant difference in practice (p=0.010). Classification codes are an essential tool in managing records, and establishing and expressing the archival bond, which is essential to the determination of a record (Duranti 1997) and thus used more frequently by records managers than by archivists in establishing authenticity.

Classification code and/or file plan						
	Never	**Rarely**	**Sometimes**	**Often**	**Very often**	**Total**
Archivist (n=123)	7.30%	8.10%	27.60%	27.60%	29.30%	100.00%
Records or Information Manager (n=96)	4.20%	5.20%	17.70%	31.20%	41.70%	100.00%
Other (n=60)	5.00%	16.70%	26.70%	20.00%	31.70%	100.00%
Pearson chi2(8) = 13.3143 Pr = 0.101						

Table 5 - Frequency of use of classification, chi square

[11] Means were calculated by treating the frequencies as an ordinal scale: never=1, through to very often=5.

Conducting the same significance tests for the least common indicator, ‘cryptographic validation techniques’, a chi-square test returns a p-value of 0.001, a result that is statistically significant (see Table 6). By using the difference of means test for frequency we can determine which groups differ. The difference is not found to be between archivists and records managers (difference of means test: p = 0.895). Here, the ‘other’ category determines the difference. The difference of means comparing archivists and ‘other’ is significant (p = 0.003). Similarly, the difference of means between records managers and ‘other’ is statistically significant (p = 0.004). From this we can conclude that the range of records professionals who do not self-identify as either archivists or records managers rely more heavily on cryptographic validation techniques to support authenticity requirements.

How often rely on for authenticity: Cryptographic validation techniques						
	Never	Rarely	Sometimes	Often	Very often	Total
Archivist (n=125)	49.60%	15.20%	18.40%	5.60%	11.20%	100.00%
Records or Information Manager (n=96)	39.60%	22.90%	20.80%	15.60%	1.00%	100.00%
Other (n=59)	35.60%	13.60%	15.30%	13.60%	22.00%	100.00%
Pearson chi2(8) = 27.4785 Pr = 0.001						

Table 6 - Frequency of use of cryptographic validation

A summary of these significance tests for the use of all indicators of authenticity in work practice based on employment category is given in Table 7. The data in the second column (Employment – all) shows the degree of difference across all three employment categories. Those cells highlighted green show a statistical significance ($p < 0.05$). The greatest differences occur in the use of archival description, retention and disposition, cryptographic validation, and audit logs. This does not tell us, however, among which employment category these difference occur. Differences were checked between archivists and records managers (columns 3-5). The cells highlighted green show differences that are statistically significant. As previously mentioned, archivists rely much more heavily on archival description (mean difference > 0), while records managers rely more than do archivists on classification (mean difference < 0.0). Further, the data show statistically significant differences in the use of written policies (governing the records and the record systems) and retention and disposition schedules (social indicators), and documentation about the record system, access and security controls, and audit logs (technical indicators).

	Employment – all	Archivists vs. Records managers		
Indicators	chi-square (p-value)	mean difference, frequency of use	standard error	p-value
Archival description	0.000	1.24	0.15	0.000
Cryptographic validation	0.001	-0.02	0.15	0.895
Retention & disposition	0.001	-0.54	0.15	0.001
Audit logs	0.002	-0.63	0.16	0.000
Classification	0.101	-0.38	0.15	0.010
Written policies governing the records	0.127	-0.39	0.15	0.010
Documentation about record system	0.154	-0.42	0.15	0.005
Standardized metadata	0.190	-0.13	0.16	0.405
Information about changes made	0.247	-0.15	0.15	0.331
Access & security controls	0.282	-0.32	0.15	0.036
Written policies governing record system	0.394	-0.42	0.15	0.006
Information about software used	0.477	-0.23	0.15	0.131
Preservation actions	0.735	0.02	0.16	0.898

Table 7 - Significance tests for authenticity indicators used based on employment type

4.2.2.5 Required Attestation of Authenticity

Of particular interest was what percentage of respondents had been required to make an attestation of authenticity, or authenticate a record or body of records, in the course of their work, and whether this affected respondents' use and belief in specific indicators. Authenticity is deemed to be an essential characteristic of records, but how often is it actually questioned and assessed in practice, and how is it tested when questioned?

Only thirty percent (30.38%) of respondents overall answered that they had been required to authenticate records for at least one purpose in the course of their work. The breakdown by employment type is shown in Table 8.

Required attestation of authenticity	N	Y	Total
Archivist (n=134)	74.63%	25.37%	100.00%
Records or Information Manager (n=97)	69.07%	30.93%	100.00%
Other (n=62)	59.68%	40.32%	100.00%
Total count	**204**	**89**	**293**
Total percentage	**69.62%**	**30.38%**	**100.00%**

Table 8 - Requirement to make attestation of authenticity

There are several reasons that one may require records to be authenticated. Respondents were asked about the circumstances for which they were required to authenticate records: 1) giving testimony to the authenticity of records in a court proceeding or administrative hearing, 2) to authenticate records in e-discovery or a legal hold process, or 3) in response to a research or reference request. Respondents could also choose: 4) 'Other' and give an explanation, or 5) state that they had never been required to guarantee or attest to authenticity. Only 4.8% of respondents had been required to give testimony in a court proceeding, and 10.2% had been involved in e-discovery or legal hold. Reference requests were the most common (18.1%).

The data were then segmented by purpose of authentication according to employment type (Figure 18), sector (Figure 19), and age group (Figure 20). Not surprisingly, reference and research requests were the most common purpose for authenticating records across employment types. This is standard practice for archivists, records managers and other record professionals who provide access to records. E-discovery is now a common occurrence associated with any trial involving electronic records (most trials) in common law jurisdictions.

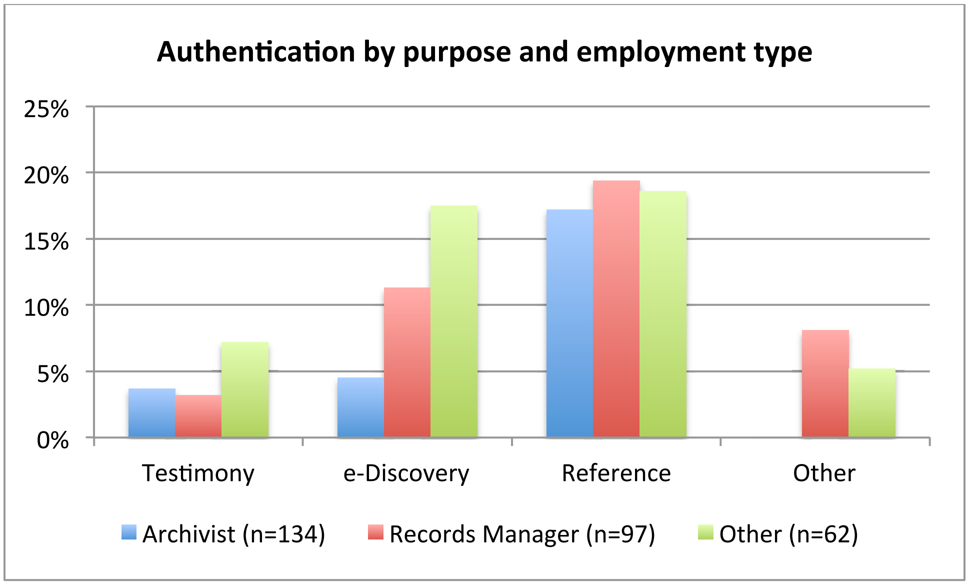


Figure 18 - Authentication by purpose and employment

By sector, records professionals employed in government or industry reported performing a higher percentage of authentications for each of the stated purposes than their colleagues in cultural industries.

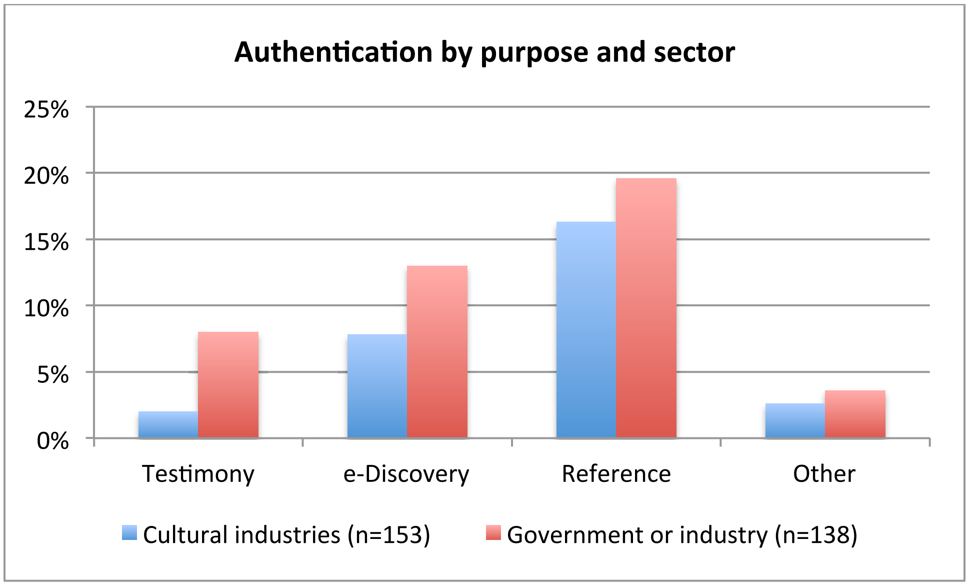


Figure 19 - Authentication by purpose and sector

These functions were distributed across age groups from 25 through 65 and older. Testimony, however, was weighted more heavily to an older demographic, suggesting greater authority ascribed to longer experience.

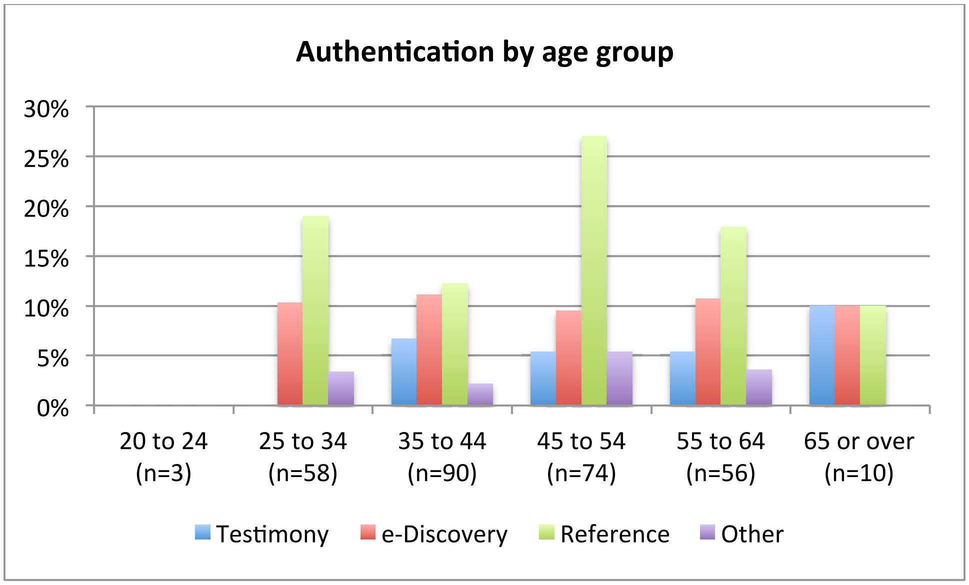


Figure 20 – Authentication by age group

4.2.2.6 Practice, Experience, and Belief

Data so far have shown what indicators of authenticity records professionals use to ensure authenticity, the percentage of professionals who have had to authenticate records, and for what purpose. The next questions address the main focus of this survey, which is to explore the relationship between practice, experience, and belief with respect to authenticity indicators. This was done in four ways: 1) for all respondents: by measuring the difference between what indicators respondents rely on to ensure authenticity versus what indicators respondents believe are most important to authenticate records (work practice vs. belief); 2) for all respondents, looking at this difference segmented by employment type; 3) for respondents who have been required to attest to authenticity: by measuring the difference between what indicators these respondents have used and what they would use (experience vs. belief); and 4) for all respondents: measuring the difference between what indicators respondents rely on to ensure authenticity versus what indicators respondents believe are most important to authenticate records, segmented by experience (work practice vs. belief, based on experience).

4.2.2.6.1 Work practice vs. belief (all respondents)

First, work practice as compared with beliefs about the relative importance of indicators of authenticity shows a significant disconnect. Table 9 shows the frequency of indicators *used* by respondents in the course of their work (work practice), ordered by most frequent to least frequent. The top three are classification schemes and written policies of the records and the records systems. These indicators are part of the social foundation of authenticity, developed from long tradition and theory (although as interview subject D103 noted, classification schemes can become extremely technical when they are automatically generated and linked to semantic ontologies). The three indicators relied on or used the least in work practice are cryptographic validation techniques, audit logs, and information about preservation actions – all technical indicators.

WORK PRACTICE - ALL RESPONDENTS			
Rank	**Indicators**	**Most often**	**T/S**
1	Classification scheme and/or file plan (n=279)	61.29%	S
2	Written policies-records system (n=283)	60.07%	S
3	Written policies-digital records (n=283)	54.77%	S
4	Standardized metadata (n=276)	53.62%	T
5	Access controls or security measures (n=274)	52.92%	T
6	Retention and disposition schedules (n=276)	50.72%	S
7	Documentation about the record system (n=283)	50.53%	S
8	Archival description (n=278)	48.56%	S
9	Information about the software used (n=280)	41.07%	T
10	Information about changes over time (n=280)	40.36%	T
11	Preservation actions taken (n=277)	40.07%	T
12	Audit logs (n=272)	29.78%	T
13	Cryptographic validation techniques (n=280)	20.71%	T

Table 9 - Ranking of indicators by use, all respondents

In contrast, Table 10 shows that the top three indicators chosen when respondents were asked to state what they *would use* to attest to authenticity (belief) reveal a technical preference: information about changes to the records over time, access and security measures, and preservation actions taken on the records over time. In fact, the indicator afforded the most importance in practice, classification, is given some of the least importance in a hypothetical authentication situation, and one of the indicators ascribed the least importance in practice, information about preservation actions, is highly regarded hypothetically.

BELIEF - ALL RESPONDENTS			
Rank	**Indicator**	**Most important**	**T/S**
1	Information about changes over time (n=248)	94.35%	T
2	Access controls or security measures (n=274)	87.80%	T
3	Preservation actions taken (n=248)	87.10%	T
4	Documentation-record system (n=247)	84.62%	S
5	Written policies-digital records (n=248)	83.87%	S
6	Information about the software used (n=247)	79.76%	T
7	Written policies-records system (n=247)	78.14%	S
8	Audit logs (n=245)	75.92%	T
9	Standardized metadata (n=243)	68.31%	T
10	Cryptographic validation techniques (n=245)	65.71%	T
11	Classification scheme (n=245)	64.49%	S
12	Retention and disposition (n=244)	63.52%	S
13	Archival description (n=247)	51.42%	S

Table 10 - Ranking of indicators by belief, all respondents

These data show that social indicators are used the most, but technical indicators are believed to be more important in establishing authenticity. This inverse relationship is shown clearly by comparing the average rank of social and technical indicators on the two different batteries (Table 11). Technical indicators rank on average 9.1 out of 13 in actual usage, but 5.6 out of 13 in expected usage, while social indicators are more highly ranked in practice than belief (4.5/13 versus 8.7/13).

	WORK PRACTICE	BELIEF
TECHNICAL INDICATORS	9.1	5.6
SOCIAL INDICATORS	4.5	8.7

Table 11 – Average rank of technical and social indicators

All indicators are afforded more weight in the hypothetical event than in practice. Furthermore, with the exception of the technical indicator, standardized metadata, the differences between technical indicators used in practice to ensure authenticity and those deemed important in the process of assessing authenticity are consistently greater than the differences between the social indicators. This again supports an inference that records professionals may believe in the promises of technology for protecting authenticity even when they do not make use of these technical indicators.

4.2.2.6.2 Work practice vs. belief segmented by employment type

Significance tests discussed in section 4.2.2.4 measured the degree of independence between frequency of use of each indicator and employment type. This can now be compared with significance tests of independence based on belief in the importance of indicators and employment type (Table 12). The only statistically significant differences that appear in what respondents believe to be important concern retention and disposition and archival description.

Significance tests - all respondents – by employment type	Work practice	Beliefs
	chi-square	chi-square
Indicator	p-value	p-value
Archival description	0	0.003
Retention and disposition schedules	0.001	0.001
Cryptographic validation techniques (e.g. digital signatures, hash digests, etc.)	0.001	0.111
Audit logs	0.002	0.272
Classification scheme and/or file plan	0.101	0.238
Written policies and procedures governing digital records	0.127	0.372
Documentation about the record system (design, operation, management, etc.)	0.154	0.17
Standardized metadata	0.19	0.234
Information about changes made to the digital records over time, (e.g. migration, normalization, etc.)	0.247	0.363
Access controls/security measures	0.282	0.127
Written policies and procedures governing the management of the records system	0.39	0.914
Information about the software used to create and manage the digital records	0.477	0.273
Information about actions taken to preserve the digital records	0.735	0.137

Table 12 – Comparison of practice and belief based on employment

As in Table 7, the green highlighting indicates a statistically significant difference based on employment type. However, here we find no difference in the perceived value of cryptographic validation and audit logs as indicators of authenticity. These technological indicators are valued equally across employment types in a hypothetical situation.

4.2.2.6.3 Work practice vs. belief segmented by legal system

It was interesting, and somewhat surprising to discover that in this sample, at least, there was no statistical difference among professionals in different regions or legal systems based on the questions asked. This was explored further in the interviews that followed.

4.2.2.6.4 Experience vs. belief – respondents who have been required to authenticate records

Among respondents who have been required to authenticate records (Group Y), questions measured the degree to which they *had used* individual indicators in attesting to authenticity (experience – Figure 21), and what they believed they *would rely on* in future instances (beliefs – Figure 22). The top two indicators used were technical indicators: information about access controls and security measures (speaking to the integrity and reliability of the records), and information about changes made to the records over time (speaking to integrity). The third indicator most used was policies and procedures governing the records.

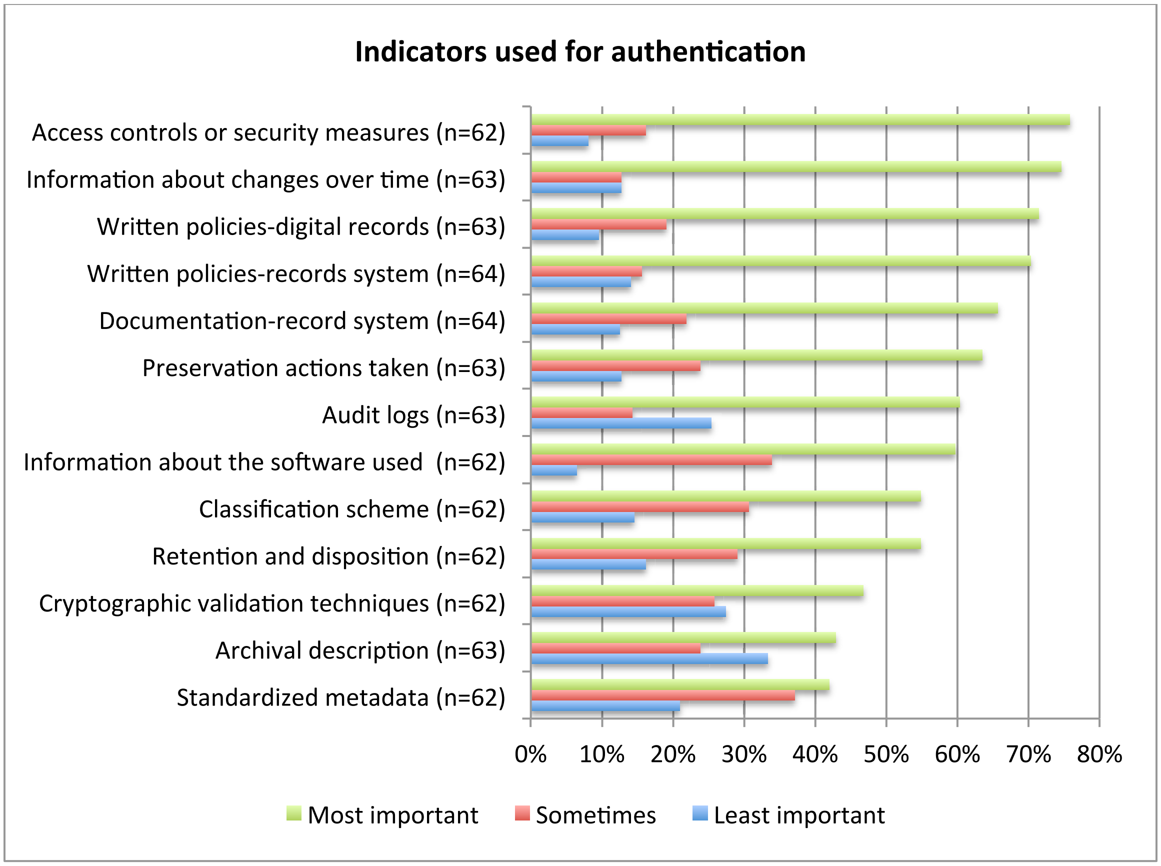


Figure 21 – Frequency of indicators used when authenticating records (Group Y)

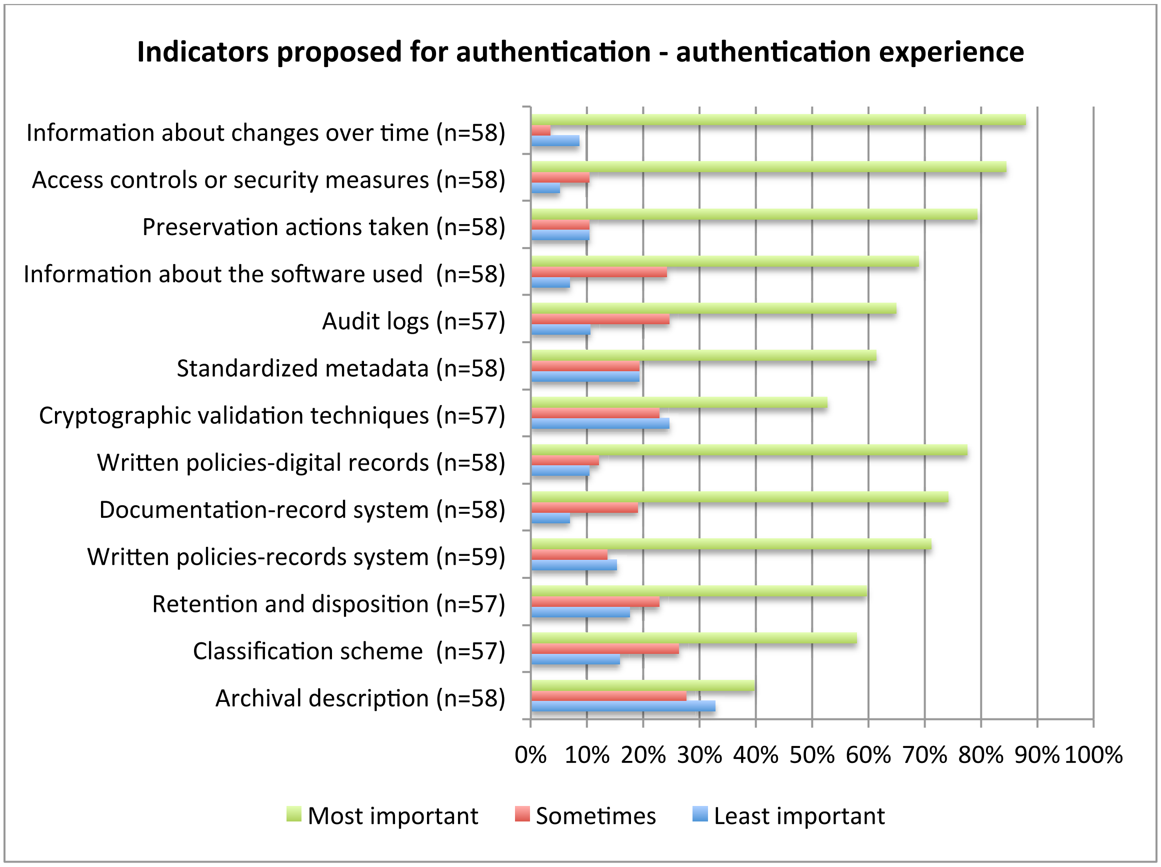


Figure 22 – Proposed use of indicators – experience with authentication

This group cited access controls or security measures, information about changes to the records over time, and written policies governing digital records as the top three indicators chosen when required to authenticate records. As shown previously, the least commonly used were cryptographic validation techniques, archival description, and classification. However, respondents with experience authenticating records indicated slightly different preferences when asked what they *would* use if required to make an attestation of authenticity, as opposed to what they had used in the past (i.e. belief over experience). The top three indicators were technical, dealing with the records explicitly

(information about preservation actions, access and security controls, and information about changes over time.)

Respondents who had not been required to attest to authenticity in general inflated the value of all indicators by as much as 10% in comparison with those who had (see Figure 23). For example, 88% of respondents who had authenticated records reported that information about changes over time was most important, while 96% of those who had not authenticated records believed it was most important. Cryptographic validation techniques were considered most important by 53% of respondents who had authenticated records, but by 70% of respondents who had not.

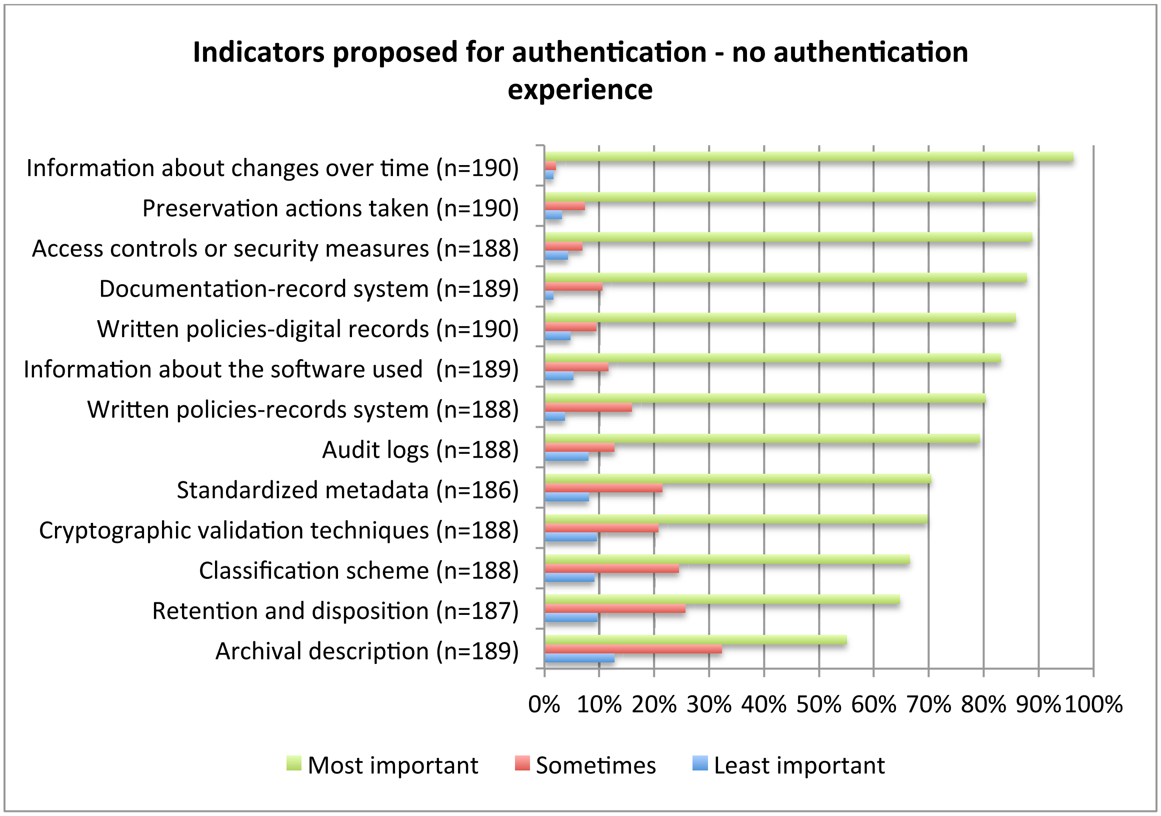


Figure 23 - Proposed use of indicators – no required authentication

4.2.2.6.5 Work practice vs. belief (all, segmented based on experience)

Similar to the examination above of whether differences between work practice and belief varied by employment type, we can now test whether these differences vary by whether or not a respondent has had to make an attestation of authenticity. As above, these differences are tested using a chi-square (Table 13).

Significance tests – all respondents – by experience	Work practice	Beliefs
Indicators used in work practice	chi-square-experience	chi-square-experience
Indicator	p-value	p-value
Written policies and procedures governing the management of the records system	0.157	0.004
Archival description	0.725	0.011
Documentation about the record system (design, operation, management, etc.)	0.146	0.027
Cryptographic validation techniques (e.g. digital signatures, hash digests, etc.)	0.025	0.032
Information about changes made to the digital records over time, (e.g. migration, normalization, etc.)	0.824	0.048
Classification scheme and/or file plan	0.277	0.077
Standardized metadata	0.521	0.085
Written policies and procedures governing digital records	0.181	0.102
Audit logs	0.348	0.105
Information about actions taken to preserve the digital records	0.559	0.118
Information about the software used to create and manage the digital records	0.097	0.14
Retention and disposition schedules	0.077	0.337
Access controls/security measures	0.079	0.488

Table 13 - Comparison of practice and belief based on experience

Highlighted values show a difference between respondents who have and have not authenticated records in the importance they afford in practice and in belief between the highlighted items.

4.2.2.7 Cryptographic Validation Techniques

A number of algorithmic or cryptographic techniques are available to help ensure integrity or security of data, and thus support a presumption of authenticity. The survey attempted to drill down into specific cryptographic validation techniques that were used in practice for the purpose of ensuring or assessing authenticity at a moment in time or over time. Cryptographic validation techniques have emerged as an indicator that is treated inconsistently between practice, experience and belief. As previous questions have demonstrated, cryptographic validation techniques were the least frequently used or relied on of the authenticity indicators in work practice (relied on very often by 20.71% of respondents) (Figure 24).

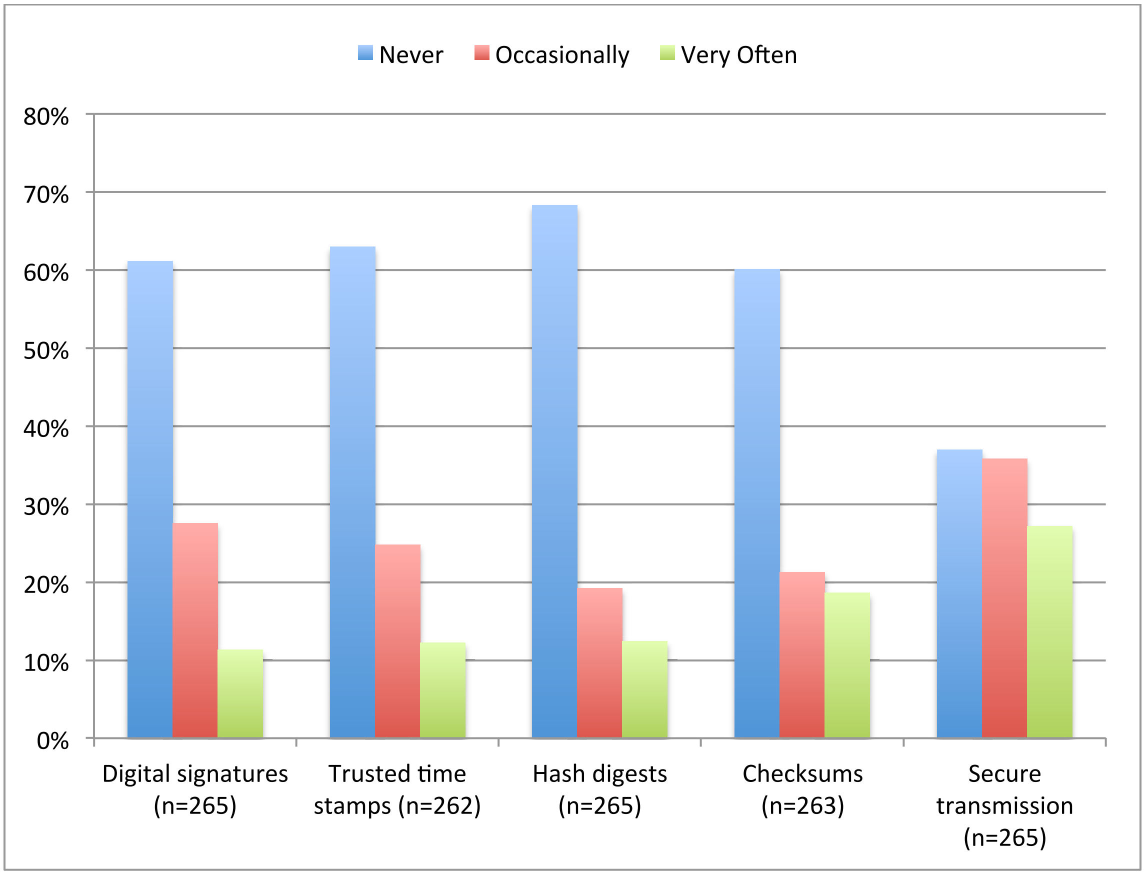


Figure 24 - Frequency of use of cryptographic techniques

Cryptographic validation techniques were relied on slightly more often by respondents who had been required to authenticate records, although they were among the least commonly used indicators. However, *belief* in their importance was significantly higher among all respondents, and particularly so among respondents who had never been required to attest to authenticity of records (Figure 25 – Group N are respondents who have never been required to authenticate records; Group Y have been required to do so).

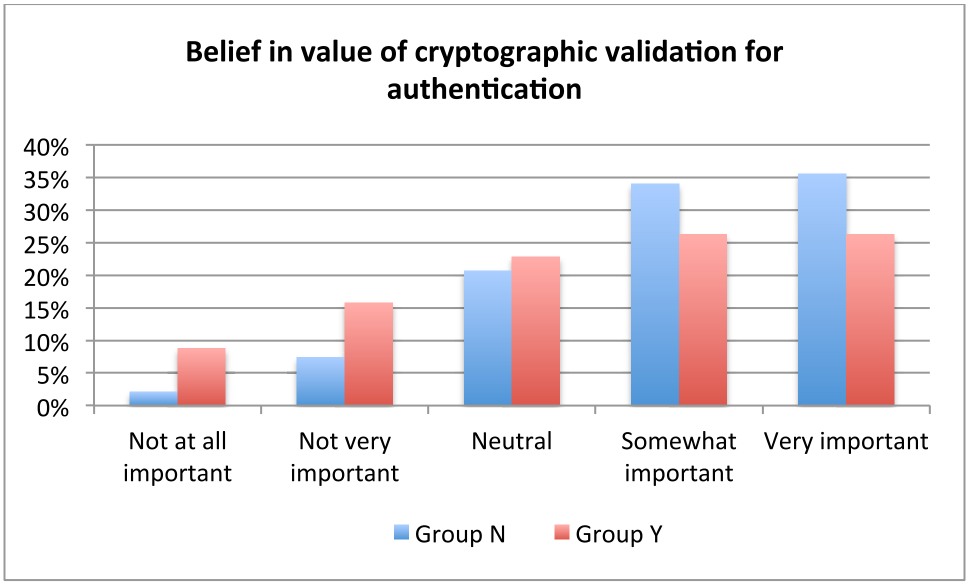


Figure 25 - Belief in importance of cryptographic validation techniques

Secure transmission was cited as the most frequently used, followed by checksums. Digital signatures were relied upon least, with only 11.32% of respondents citing use of digital signature technology to assure authenticity, while 61.13% never used this technology (see Figure 26). However, as further analysis showed, cryptographic validation techniques were not extensively relied on when making an attestation of authenticity.

4.2.2.8 Metadata

Respondents were asked to indicate if they routinely used or managed any of a standard metadata schema or guideline (e.g. Dublin Core, PREMIS, MoReq), a modification of a standard schema, a custom-built schema designed in-house for their purposes, system-generated metadata, or non of the above/not sure. The greatest percentage of respondents identified system-generated metadata. The most commonly used schemas were Dublin Core, followed by PREMIS. This question was explored further in the interviews.

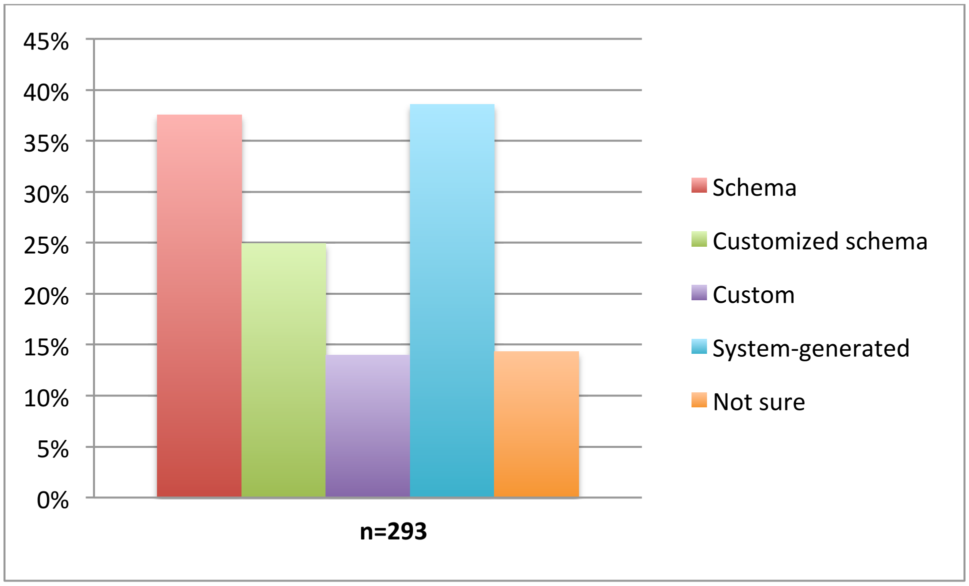


Figure 26 - Frequency of metadata

4.2.2.9 Storage

As expected, respondents had little confidence in various storage options for digital records. Removable media (e.g. USB drive) was the least trusted, followed by 3rd party cloud providers. Surprisingly, network drives were considered more trustworthy than 3rd party storage of traditional records. Archives retain their authority, being most highly trusted for analog and digital records (Table 14, Figure 27).

Confidence level				
Storage type	**Little**	**Neither**	**Considerable**	**Total**
Removable media (n=251)	55.38%	26.69%	17.93%	100.00%
3rd party cloud (n=251)	34.26%	37.45%	28.29%	100.00%
Stand-alone computers (n=250)	39.60%	30.40%	30.00%	100.00%
Traditional stored by 3rd party (n=251)	23.11%	43.43%	33.47%	100.00%
Network drives (n=250)	20.80%	36.00%	43.20%	100.00%
Traditional stored by creator (n=249)	6.83%	22.49%	70.68%	100.00%
Digital stored by archives (n=250)	2.80%	12.80%	84.40%	100.00%
Traditional stored by archives (n=251)	1.20%	4.78%	94.02%	100.00%

Table 14 - Confidence levels in storage options

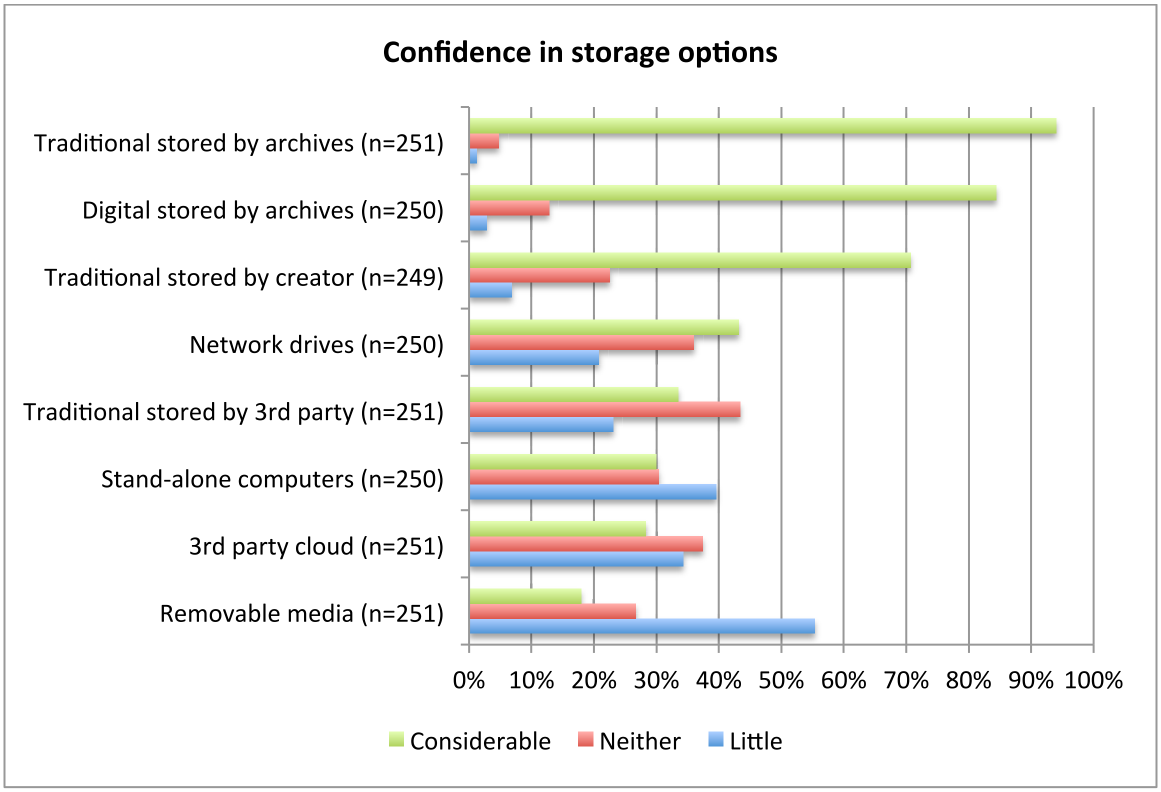


Figure 27 - Confidence levels in storage options

4.2.2.10 Organizational Recognition of Authenticity in Policy Instruments

Two hundred and thirty-five respondents reported on the explicit mention of authenticity requirements in policy instruments governing digital records. The majority reported that their organization did not define authenticity of digital material (127 respondents; 54%) or that they did not know if authenticity was mandated in policy (40 respondents; 17%). This was explored further in semi-structured interviews.

4.2.2.11 Narrative Questions Exploring Respondents' Views

Respondents were asked two narrative questions:

1. What is your definition of authenticity of digital records, and
2. What do you consider essential to proving the authenticity of digital records.

Responses were coded in TAMSAnalyzer in an iterative process using constant comparative analysis and reflection. Constant comparative analysis, originally developed for use in grounded theory, allows for categories to emerge from examination of the raw data, without *a priori* expectations, although "it is inevitable that prior research will have identified some of the salient issues" (Pickard 2013, 269). Coding of the present data began with a close reading, and identification of concepts and practices about authenticity and authentication. From that point, a set of codes emerged until a point of theoretical saturation. The final code book is found in Appendix 5.

Analysis of these narrative responses reveal that authenticity is still generally assessed according to traditional social heuristics. In defining authenticity (n=175), eighty respondents (46%) defined authenticity in terms of integrity, and several stated that bitwise integrity was necessary after the moment a record was "fixed" – that is, chosen to be kept as evidence of the action represented in the record, or preserved for long-term reference in an archives. This concept was frequently expressed as originality and proof of chain of custody, evident in these three opinions:

> *That we are able to prove chain of custody in court for a digital record. that means who had access to it, was the record ever altered, migrated or had its format changed, is it the same record as the original stored.*
>
> *Records are from certified originals with no change and approved by the system.*
>
> *Records which reflect, in content and composition, the original material as manifested by the creator.*

The second most frequently occurring concept in respondent definitions was that of identity (61) followed by provenance (18). In eleven instances standards were cited: the definition for ISO 15489 was cited eight times. Eight respondents (three archivists, five records managers) explicitly noted that records produced in the usual and ordinary course of business could be presumed authentic, thus reflecting statute and precedent law governing business records in common law traditions. Conceptual ideas were balanced with specific practices. The most common practices cited within definitions of authenticity focused on security: creation and availability of audit logs was cited 16 times, and secure storage 12 times. Of the 175 responses, 14 were coded as “concept>none, meaning that the respondent took the time to write and answer, but had no definition of authenticity. Responses included, “Don’t know,” “Having trouble coming up with one, even though I have taken all the required courses for the SAA DAS certificate,” “I believe that we are still debating the finer definition of this,” “I do not currently have a valid one, though my impulse is to consult legal requirements,” “The ability to be able to prove that your records are authentic,” and “I don’t know.” One respondent rejected authenticity entirely, although admitting the possibility of integrity:

> *Digital records can, in my opinion, never, I repeat: never, be considered 'authentic'; You could say that 'the expression 'authentic digital record' is a contradictio in adiectis, like 'black snow' or 'white coal'; An entirely other thing is the 'integrity of digital records'; Integrity of a digital record is in principle a possibility; It does however require a fabulously complex and thorough monitoring and recording of all the - inevitable - changes made to a record during its 'lifetime', or rather: during the many 'reincarnations' a digital record goes through.*

Another respondent posed a challenge:

> *Well, what is *your* definition? This is not a term I use; I am familiar with it only (e.g.) in diplomatic, when it refers to a charter or other legal document being what it pretends to be, i.e. originating in the circumstances, and at the date, indicated on its face; issued at the behest of the signatories, who are real people, etc. Applies to electronic transcriptions only rarely (e.g. if a file was transcribed from a source other than that indicated in its header); Tends to have relevance only in matters of*

> *intellectual property / digital rights / copyright; Much more pressing are version control, and sheer accuracy.*

Finally, the contextual and relative nature of authenticity was most clearly and succinctly expressed thusly: "That would depend entirely on who is asking, and why."

The second narrative question explored respondents' beliefs about essential indicators of authenticity (n=191). Use and presence of metadata was mentioned most frequently (37 times) followed by audit logs (31), written procedures (26), standard policies (23), access and security controls 23), presence of a systematic recordkeeping system (21), use of checksums (16), presence of documentation (13), and secure storage (12). Concepts reflected in this question were, in decreasing order of frequency, integrity, chain of custody, provenance, identity, and reliability. Several respondents noted the importance of cryptographic validation techniques, and several specifically stated that security and access controls were paramount (although one respondent noted the importance of these controls in the context of using public cloud-based email and document sharing.)

One of the more comprehensive answers identifies both controls necessary and the challenges of implementing them that again, supports the idea of a pragmatic approach:

> *Ensuring the system managing the records can preserve its identity, integrity, and accuracy over the long term; This includes having proper system documentation, controls, defined processes, procedures, and policies; The challenge, however, is that some controls required to ensure the authenticity of a digital record have severe impacts on system performance and platform stability; Within the context of*

> *a business that requires well-performing systems to support efficient and reliable operations, sacrifices must be made, including disabling some functionality that would support the presumption of authenticity.*

Analogies were drawn to paper:

> *In a lot of ways, I see the base questions to be the same as with paper - whatever happened to and with this digital record, was it done in the usual and routine course of business - does it comply with the same precedent that has been set for other records -- ideally those precedents are based on sound metadata and access controls.*

As with the definitions, many answers were very general in nature, for example:

> *Creation information, storage information, actions taken on the record.*

Or:

> *The complete control of the cycle of life of the digital record and the archive.*

Three observations may be drawn from these responses. First, definitions as well as essential indicators tend to be very general and broad, without offering any real guidance or reflecting any positive model. Second, answers supported the traditional models of authenticity, the social heuristics based on establishing identity, and showing chain of custody and provenance. The heavy reliance on integrity, however, overshadows other parts of the traditional archival definition, and demonstration of integrity is made through technological mechanisms of control. Finally, responses generally reflected a pragmatic approach to authenticity, for example, one respondent answered:

> *Is [the record] sufficient for the purposes it may be used for? Would it satisfy a judge or adjudicator? Whatever I can claim about it, can I back that up with facts? / The basic definition of an authentic record is "Can it be used as an authentic record in a situation where an authentic record would be needed?" This is not a yes/no answer (though the question is), but rather a range. I want the records as authentic as they need to be for future uses. They needn't be the MOST authentic - just authentic enough.*

4.3 From Survey to Interviews

The survey data suggest a considerable disconnect between practice and belief. In practice, respondents rely on social indicators of authenticity over technical indicators, but their expectations are reversed. This indicates the need for further research to explore in greater depth the importance of social versus technical indicators of authenticity, and how these are used when authenticity is questioned. Eighty-nine respondents indicated a willingness to be interviewed, thus providing the research frame of the qualitative strand of this study. In the next chapter, the interview methodology and results will be presented.

5 Interviews

5.1 Introduction

Semi-structured interviews were conducted as the second strand of the mixed methods study. The purpose was to explore in greater depth practitioner behavior and beliefs.

5.2 The Research Sample

The interview sample was chosen from the subset of survey respondents who indicated their willingness to be interviewed (n=89). Purposive sampling allowed a strategic approach intended to establish correspondence between the research questions and the sample (Bryman 2004, 334). The goal was to develop a sample that showed a balance of professional practice, sector, region, and legal system. The initial thought was to interview only respondents who had been required to make an attestation of authenticity in their professional capacity, however it was later decided that interviewees who had not performed this function would allow greater understanding of the role that such experience plays in ensuring and assessing authenticity. The expectation was to interview 20-25 individuals, beginning with a smaller selection and continuing until theoretical saturation was achieved (Bryman 2004, 334).

In total, twenty-one invitations were issued over the course of seven weeks, and seventeen interviews were conducted. All interviews were recorded with the permission of the interviewees, and then transcribed. Three interviews were conducted in person, and the remaining interviews were conducted by telephone. Each interview lasted from 45 minutes to 1.5 hours. The profile of interviewees is found in Table 15.

ID	Profession	Sector	Notes	Country	Legal system
D025	Records or Information Manager	Government or industry	Government/ national archives	Portugal	Civil
D026	Records or Information Manager	Government or industry	Municipal government	Spain	Civil
D043	Other	Cultural industries	Institutional repositories	US	Common
D073	Archivist	Cultural industries	Educator, consultant	Italy	Civil
D087	Archivist	Cultural industries	State archives	US	Common
D097	Records or Information Manager	Government or industry	Government/ national archives	UK	Common
D103	Other	Cultural industries	Scientific data	UK	Common
D112	Records or Information Manager	Cultural industries	Bio-pharmaceutical industry	US	Common
D123	Archivist	Government or industry	State government	US	Common
D129	Records or Information Manager	Government or industry	Government	Brazil	Civil
D131	Records or Information Manager	Cultural industries	Museum	Canada	Common
D148	Archivist	Cultural industries	University archives	Canada	Common
D187	Records or Information Manager	Government or industry	Government	Spain	Civil
D206	Archivist	Government or industry	Government/ open data	UK	Common
D334	Other	Government or industry	Municipal government	Canada	Common
D429	Archivist	Cultural industries	University archives	Canada	Common
D441	Archivist	Cultural industries	International organization	Netherlands	Civil

Table 15 - Interview profile

5.3 Emerging Themes

As discussed in the last chapter, the survey results suggest that records professionals adopt a pragmatic approach to authenticity based on resources, sensitivity of the records, and organizational or legislative framework. There is a greater reliance on social indicators of authenticity in the daily practice of ensuring authenticity, but a higher belief in the value of technical indicators when assessing or attesting to authenticity. Furthermore, professionals who have never been required to make an attestation of authenticity placed higher emphasis on the value of technical mechanisms than did their colleagues who have made attestations.

Several themes emerged from the survey questionnaire that supported the research questions, and warranted further investigation through interviews. These were:

- the difference between practice and belief of records professionals regarding different indicators for ensuring or assessing authenticity,
- the difference between employment types (archivist, records manager, other) in establishing or assessing authenticity,
- the relative weight of technical versus social factors in ensuring or assessing authenticity either in practice or belief, and
- the role of experience in making attestations of authenticity in practice and belief.

Furthermore, the interviews provided an opportunity to ask specific questions about the role of the system of law (civil or common law), and the extent to which traditional archival models of authenticity suffice in the digital environment.

5.4 Interview Protocol

The interview protocol followed the outline of the second half of the survey (work practice and belief about authenticity), and questions focused on specific responses (see Appendix 4). Interviewees were asked to elaborate on what consideration of authenticity was involved in the tasks they conducted most frequently. Subsequent questions explored use of and reliance on metadata, on cryptographic validation techniques, attitudes toward moving records to the cloud, the role and purpose of specific indicators of authenticity (as outlined in the survey) for establishing or proving identity, integrity, or context, and interviewees' opinion of the validity of the classification of indicators into technical or social categories. All interviewees were asked whether they felt that traditional models of authenticity sufficed in the digital environment, or required expansion or revision. Issues of terminology and comparability of concepts such as provenance, lineage, quality and authenticity were also explored. Questions were tailored to each interviewee's survey responses, and a copy of the interviewee's survey responses and the interview questions was provided to the interviewee prior to the interview. However, this protocol was used strictly as a guide to the ensuing conversation, allowing the researcher to freely follow themes as they emerged, and incorporate new ideas into subsequent interviews.

5.5 Analysis

All interviews were recorded and then transcribed for analysis. For the purpose of reporting, transcripts were anonymized, with names of individuals and institutions removed. Throughout this chapter, interviewees will be referred to by a unique code applied to each survey respondent and carried over into the interviews.

Coding was done differently from the narrative questions in the survey. Where the survey questions were approached from an open coding perspective in order to reveal concepts and practices (although with an awareness of the author's predisposition to certain concepts based on archival theory), coding of the interview transcripts began with an *a priori* set of codes, derived from the survey questionnaire and reflecting the research questions and emerging themes. Interview transcripts were coded using this initial set, with more codes reflecting greater specificity being added in an iterative process as the coding progressed. This required coding to be reviewed and updated iteratively throughout the process. Interview codes were developed in the following categories:

- assessment – these were the indicators used in ensuring or assessing authenticity, and interviewees' opinions about these indicators,
- purpose – these reflected the specific purpose or function of indicators in establishing identity, proving integrity, or describing context,
- cloud services – these gathered opinions and practice regarding cloud services
- critical incidents and legal issues – these identified the drivers behind requirements for authenticity,
- archival models – these assessed the continuing validity of and sufficiency of traditional archival models of authenticity.

5.6 Findings

The findings that follow are organized around several themes:

Terminology

- Terminology is an ongoing challenge in interdisciplinary work and is of great concern in information fields. The publication of glossaries by national professional associations (c.f. Society of American Archivists 2005), and the extensive preparation of multi-lingual as well as interdisciplinary terminology databases by research projects such as InterPARES (InterPARES 2012) and InterPARES Trust (www.interparestrust.org), and international associations (ICA 2015) are evidence of the concern for this topic.

Experience attesting to authenticity

- Several interviewees talked about their experience authenticating material. Much of what they have done is still rooted in paper-based methods using social indicators. When authenticating material for law enforcement or court, there is still little call for technical authentication by information professionals. When technical measures are required, forensic experts provide technical expertise and present their findings as expert witnesses (c.f. Carrier and Spafford 2003; Garfinkel 2009).

Classification of indicators as Social or Technical

- In general, interviewees agreed with the classification of indicators as social or technical.

Metadata

- Metadata were identified as one indicator that could encompass either social or technical aspects.

Social and Technical Indicators of Authenticity

- Interviewees talked about the frequency with which they used different indicators and the relative weight they assigned to them. Their comments suggest that the categorization of indicators into social and technical can be interpreted with greater nuance.
- They also discussed the purpose of using specific indicators. This addresses theoretical concepts of authenticity generally, and integrity, identity, provenance, and legal purposes.

Redundancy

- Redundancy was identified as an essential means of establishing authenticity by one interviewee.

Trust and control

- The theme of trust versus control emerged implicitly and explicitly throughout the interviews.

Cloud computing

- Moving certain records for certain purposes to cloud platforms is inevitable, and almost all interviewees are experiencing this shift to some degree in their organizations. Response to the cloud depends on the risk, real or perceived, that the cloud poses, and the level of involvement of the information professional.

Legislative frameworks

- Distinctions between work practice and belief were not revealed through the survey. This was a topic of interest in the interviews.

Critical incidents

- A concept issue that arose several times was that of the "critical incident" and the role of crises in effecting change.

Traditional archival model of authenticity

- The traditional models of authenticity that archivists and records managers have used for analogue materials are generally still appreciated, although several respondents suggested that changes were required.

Summary

5.6.1 Terminology

One of the challenges in understanding concepts across communities of practice is the use of terminology. When discussing authenticity, the challenge is particularly acute due to the ambiguous nature of the term and other terms that are used in conjunction or synonymously with it. One such term is provenance. Provenance information is understood as a necessary support to an assessment of authenticity, but is it alone sufficient? One interviewee (D073) spoke about the relationship between the findings of InterPARES and the OAIS model. In the OAIS there are four categories of preservation descriptive information: reference, provenance, context, integrity. Do these map to concepts in InterPARES? It would be useful to put these vocabularies together to achieve clarity and greater understanding. Another interviewee (D131) spoke about creating a provenance package to accompany audio recordings of oral histories (who made the recording, who was interviewed, the date, time, and other contextual and identifying information), and the challenges of human error, for example if the date is not set properly.

When discussing scientific datasets, the concepts of authenticity, lineage, quality, and provenance overlap and their interpretation depends on the scientific discipline concerned. Provenance includes time stamps, information about who collected the data, where they were collected, how the data were put together, how they were cared for over time, and what algorithms or aggregations may have transformed then. All this information is necessary in order to trust the quality of the data and to support reproducibility. Where the data come from will determine their fitness for use, and so provenance or lineage provides evidence to support a presumption of authenticity and quality (D103). Scientific data researchers might be wary of data from an open data portal, primarily because these portals don’t have enough information, captured in metadata, about the methodological rigor with which those data were collected, nor are they likely to have a contact person to answer questions. Provenance information is critical when searching and retrieving scientific datasets, and the scientific researcher would not use a dataset that didn't have that information available (D103). For the public, however, provenance information may be of no interest:

> *Researcher: Are the users [of open datasets] making assumptions about authenticity based on it[s source]?*
>
> *D206: I would say not - all we get from data users is "give us the data" and they will worry about authenticity and quality down the road - it doesn't matter where it comes from, they are just not interested.*

5.6.2 Experience Attesting to Authenticity

The survey measured the extent to which records professionals are called to authenticate digital material and found it to be an infrequent occurrence. This was supported by comments of the interviewees – even those who had answered in the affirmative in the survey. Rather than discussing specifics of how these individuals authenticated material, the conversations tended to focus on the value or weight of indicators of authenticity. Among interviewees who had made attestations, indicators mentioned most frequently were checksums and metadata.

> *D043: The only times I've been involved in authenticating materials almost always involves attorneys and a judge, when we have done things in the process of creating paper, and then microfilming – people say should we keep the*

> *original, we say if you keep it, you are going to have to produce it… if we scan the microfilm and put it on a write-once material with a checksum, and destroy the microfilm, then that's the best [evidence]… when it's on write-once media with a checksum, it's hard for people to say I changed this.*

For many, attesting to the authenticity of digital records has not yet happened. D026 reported that "this year [2014] for the first time, the police required some information about some [digital] files… I tried to explain the system to them but they only wanted paper copies and screen shots - that's the level that is required for these cases." Another interviewee (D334) concurs that in his experience, courts seem happy with screenshots of digital information, and that the number of times authenticity is questioned is "very tiny." A challenge would only arise where something out of the ordinary has happened – without some alarm, there will be no question. "It's the relationship between what's supposed to happen and what's actually in front of me." These two interviewees both work in municipal governments but in different countries and legal systems, yet their experiences are similar.

5.6.3 Classification of Indicators as Social or Technical

All interviewees agreed in principle with the categorization of indicators of authenticity as social or technical. As previously stated, this categorization defines as social indicators instruments developed by an organization to support the creation, management, or preservation of records (e.g. classification schemes, retention and disposition plans, policies and procedures documents). These instruments may also be imposed externally in the form of mandatory or voluntary standards or regulatory or legislative requirements. Technical indicators are non-discretionary, that is, they are the result of a work process or state change in the records. Technological indicators may be used to control or protect the records (e.g. checksums or digital signatures), but are often focused on controlling, protecting, or monitoring the system in which the records reside (e.g. access and security measures and audit logs.)

Three interviewees suggested that standardized metadata could be considered either a technical indicator or a social indicator (D026, D131, D334). This bears further consideration – if the metadata in question is generated automatically by the technological system in which the records are created or maintained and is not describing the content of the records, then in keeping with the proposed definition of technical indicators, this author would classify them as technical. If the metadata is descriptive, whether generated automatically through a workflow process or added manually, then it serves as part of the social foundation of authenticity and is a social indicator.

D097 offered an alternative view, suggesting that the issue could be viewed as a distinction between the "policy process platform and people components," each of which generate attributes in different ways. The policy or process may generate an attribute automatically, while people might apply choice.

Another interviewee (D087) suggested explaining the categories as subjective versus objective, or human created versus automated, but added: "those maybe are too specific, yes, I would agree with this general distinction [of social versus technical]."

D123 summed up the distinction by saying that in his opinion, the social indicators form the core foundation of establishing authenticity, but the methods used to actually effect authentication of a body of records would be the technical mechanisms.

5.6.4 Metadata

The topic addressed most frequently by interviewees was that of metadata. The presence of metadata guidance in legislation, the issue of trusting metadata, the intent or purpose of their application and use, the question of what metadata are needed and how they support authenticity (among other functions), and their importance were all discussed.

In certain countries or domains of practice, very specific guidance about metadata exists. In Italy, the attention of records professionals when designing systems for management and storage of records is focused on the implementation of national legislation which dictates a rich set of metadata establishing identity and integrity, supported by documentation about policies and procedures outlining clear responsibilities. Italian legislators have recently provided detailed guidelines to verify the quality of metadata and applications for records creation and (as a consequence) for their transfer to any kind of custodial environment (including cloud storage). Specific attention is paid to the documentation and metadata provided for transferring records from the creator environment to any other system. However, even with this stringent legislative framework, these metadata were not considered by the interviewee to be sufficient to support an assessment of authenticity, but only some areas of management and interoperability. Additional metadata remain to be identified by guidelines, like descriptive metadata and structural metadata (D073).

Several ISO standards provide guidance for metadata required to support preservation, and by extension authenticity. Geographic data are supported by ISO 19115, which has 215 recommended metadata fields. These include most of the elements outlined by InterPARES in the Benchmark and Baseline requirements (D103). The OAIS Recommendation is published as ISO 16363, and complies with the general standard ISO 23081, Information and Documentation – Managing Metadata for Records. The latter is useful as a guideline, but it must be supported by other specific standards (D073).

What metadata are needed, how much, and for what purpose is still an object of debate (D148, 429, 334). The answer may differ if viewed from the perspectives of a creator, a preserver, or a user. For some users, almost no metadata would be enough, depending on what it is they want from the data. Metadata that are insufficient for one purpose will be very satisfactory for another. Approaching the issue from potential reuse,

> *A record is a product of a function - the data producers capture the metadata they need for their purpose, that's what goes into the archive, and that's either enough for the reuser or it's not, but the producer doesn't need to think about it (D206).*

Organizations are concerned with identifying what metadata are needed, whether generated by design or by the software and operating systems – systems generate a lot of data, much of which is not used or preserved. Different metadata are associated with how information is recorded or data are collected, e.g. GPS data (D103). The focus is on collecting information about actions and context: the metadata describing what, who, when and how (D026, D103).

One municipal organization has “a home grown ad hoc metadata set that has been used for physical records” but this has not been translated to digital records, which at this point are not maintained in an EDRMS. This organization developed a set of digital records management metadata some years ago, based on current national and international standards and guidelines, but it was not developed by the records professionals managing the records, and has not been implemented. This interviewee noted that “we are at a stage where we have very little knowledge of what it is we need to manage – we’re just starting to get an idea of how we are going to do that.” Much of the knowledge about the physical records, beyond what documentation exists, is held in the corporate memory of the people who work there, and unless there are huge flags raised about a set of records, people will assume things are fine. With respect to digital records, so far the metadata that exists hasn’t been mistrusted because it is the content that people are looking for, and authenticity is often judged by contextual information outside of the records themselves. For example, if a word processing application attaches information, such as author, that is incorrect because user profiles have not been kept up to date, that information will be ignored in favor of contextual information such as dates of employment (D334).

> *I want as little as possible – no one has been able to explain to me satisfactorily why we need it… no one knows how it’s going to be used through time – is anyone acting on it (D148).*

Working in a university archives, D148 is also more concerned with discovering the authoritative version of a record than its authenticity. When transfers to the archives are from organizational departments the transfer may contain as many as 17 versions of a record – this is a more urgent problem. Absent legal or regulatory requirements, it is very difficult to impose changes to recordkeeping behavior on departments, and adding requirements for metadata is likely to fail.

> *I'm skeptical of anything that's going to require the records creator to change their behavior… when we receive a transfer of records we're going to demand a date range and mapping to a [retention] schedule as a minimum, beyond that… a requirement? we are working that out (D148).*

Repository software being developed for this organization incorporates the Dublin Core metadata schema, but it is not yet being used. PREMIS is slowly being introduced, and national descriptive standards are adhered to and relied on heavily for contextual information. As the repository is built, an in-house schema will likely be developed. The approach taken to digital records management, through advice to departments, and preservation, as departments begin to transfer records to a new digital repository, is pragmatic and strategic, a mix of knowledge about digital recordkeeping and preservation, and human nature/organizational culture.

> *This comes back to what we can control and can't. I'm more open to documenting and using metadata in our context than requiring creators to do so, so we are taking checksums, doing fixity checking, maintaining a log of all the technical*

> *components of our system, software, etc. to prove over time that we are a TDR – we can control that so it's easy to apply (D148).*

The use of metadata in establishing authenticity is more important to some interviewees than others. For D334, describing management of active records from a policy development perspective, metadata are less critical for establishing authenticity than policy. Another interviewee talked about the process of designing a new archival repository system that will incorporate PREMIS and give more assurance of authenticity, but currently is relying for analysis on metadata generated by the software at the point of acquisition for analysis, which are kept, although not in a standardized system, and are not considered sufficient for authenticity (D087). By contrast, D123 reported that the in-house schema used in his repository does provide sufficient support for authenticity through integrity information: checksums and information about the transfer process.

Most interviewees, however, agree that existing schemas, whether standardized like Dublin Core or PREMIS, or custom built are insufficient as they currently exist for establishing a foundation for authenticity assessment (D073, D026). And, finally, in all cases, "metadata without controls cannot be relied on for authenticity" (D043).

5.6.5 Social Indicators of Authenticity

5.6.5.1 Policies and Procedures

The role of policy and procedures, whether governing the records systems or the records themselves, was the most frequently discussed of the social indicators across sectors and employment types. Many interviewees considered the presence of a strong policy framework to be the most important factor in ensuring and assessing authenticity.

Most interviewees advocate for a strong policy framework as the foundation of authenticity. Policy and procedures lay the foundation – some of the technical mechanisms are important but many organizations don't have the capacity to implement them yet. Policies (social indicators) are deemed more important than the "in the trenches" work (application of technical indicators) (D112). A policy framework is the common thread that applies regardless of community of practice or context of production (D025).

A particularly strong endorsement was put forward by D334, who began the interview by saying "I don't use indicators – the authenticity [of records] is absorbed or addressed in terms of the accountabilities and responsibilities outlined in policy, driven by business priorities and requirements, service delivery, administrative process etc. as opposed to specific components elements aspects of recorded information, so metadata [for example] is less critical to me for establishing authenticity than policy." Policy and procedure also provide the keys to accessing records. "Policy tells where the mandate is – that's the management of the records system writ large - not the techie who runs the EDRMS" (D334).

There are also organizations that are required by law to implement certain technical mechanisms, such as digital signatures and hash digests, but are in the process of developing policies to guide practice. Policy development is addressing increasing use of digital technologies: one municipal government has started to approve some policies and procedures at a high level of specificity for critical records about political decisions, that guide the context of creation and the sequence of actions for validation and integration into the recordkeeping system (D026). This presents the opposite view of social controls following technical implementations.

Policy and procedures can assist in convincing reluctant records creators or stakeholders of the importance of authenticity and good recordkeeping. Interviewees from one organization reported that they are working on providing a high level and more detailed policy and procedures framework, as part of a process of convincing people of the benefits to them from using a structured records system with classification plan and retention and disposition schedules, naming conventions and version control. “If we could accomplish that it would be a huge step forward for any organization – a huge process” (D148 and D429). Some people see policies and procedures as a way to control workflow and foster communication between departments. This has been extremely helpful to the producer of the records, in the experience of D123, and “gets their IT people talking to our IT people.”

There are detractors as well. For example, policies are identified as useless in any larger framework of societal or governmental corruption (D187). Procedures may be put in place originally that do not carry much weight over time. Written policy and procedures may allow future replacement personnel some understanding of how things should work, but personnel who remain in their jobs for a long time actually become the system. If they do not train new personnel, or if they have become lax in adhering to the policies or procedures, then the system falls apart (D043). Says D043, “written policies – we use them sometimes, but when you look at what is actually being done, you realize they aren't being followed…. written policy is not much help unless you can see a trail of implementation of the policy. Policy can be important if well adhered to."

Even the staunchest supporters have no illusions about the impact of the human factor. In response to the researcher's question, "Are you always confident that policy is being followed?" D334 replied, "Definitely not! - the [organization] is a large complex organization – policy is an imperfect creation… sometimes it is being ignored (sometimes with justification), not known, or [it is] being followed – a mix."

5.6.5.2 Documentation About the Record System

Opinions about the value of documentation about the record system were mixed.

D334 stated that "if it's there… it's usually not available or cryptic," and observed that "We don't look for documentation, we ask." Documentation about the record system is not important because staff is available to "go to the right space to find the records." When this interviewee was working with the vendor who developed the organization's EDRMS system, he requested regular reports indicating that the system was operating properly as part of good practice. The vendor replied that he didn't know if the system produced reports like that, and indicated that no one else had ever asked for them.

In the same vein, another interviewee noted that original documentation is often either so poor as to be useless, or it is a "that's how it should have been done" guideline, but has not been observed. In one case, the documentation about the record system was considered to be very important but unavailable because the IT department did not retain it (D026).

5.6.5.3 Classification, Retention and Disposition

The presence of a classification code measured as the most important indicator of authenticity used in daily work by survey respondents. Interviewees were less enthusiastic. According to interviewees in one organization (D148, D429), the use of classification codes is voluntary – although the archives/records management department has developed a model scheme for paper records, uptake is the exception rather than the norm, with implementation by only half a dozen out of a possible 180 departments. The classification code, the expression of the archival bond in traditional archival theory, is strictly voluntary, supplanted by two key pieces of information (in this organization), the records retention schedule and date range of the records. The issue is one of control. This archives/records management department must adopt a pragmatic approach to introducing records management because without legal requirements, departments will not change their behavior if the procedures are too onerous, and will only cooperate because they want to gain control of their records.

Retention and disposition schedules were discussed by two interviewees, who held opposite opinions. One considered them unimportant except to reveal the absence of records that should not have been destroyed (D112). This interviewee, a records manager in the cultural sector, noted that users make copies of records and keep them forever, regardless of the presence of a disposition schedule, and that convenience copies are not covered by retention and disposition. These schedules do not address authenticity. The other, an archivist in the cultural sector (D148), used the retention and disposition as a key indicator, with date range, in a presumption of authenticity.

5.6.6 Technical Indicators of Authenticity

The most frequently mentioned indicator of authenticity was metadata. Initially proposed as a technical indicator, on the basis of the interviews metadata will be considered on its own, as the variety of metadata may form part of a technical and social foundation for authenticity. Checksums (or hash digests) were the next most frequently discussed indicators. The following sections will examine comments about cryptographic validation techniques in general, and then proceed from the most frequently to least frequently discussed cryptographic mechanisms, followed by access and security, preservation actions, information about changes to records over time, information about the software used to create records, and audit logs.

5.6.6.1 Cryptographic Validation Techniques

With the exception of the use of checksums and hash digests, interviewees agreed that cryptographic mechanisms were only useful for establishing authenticity through an authentication function in the short term, at a point in time. These techniques have relevance for temporary transmission but cannot be used for long-term preservation. Their use in transmission, when documented and supported by policy, can be considered an element for assessment, but they cannot be used for assessing integrity and provenance when the records have been transferred to a trusted digital repository for final custody

(D073). Not all interviewees used cryptographic mechanisms. One remarked that they might be appropriate for “the top end user at that extreme end where authenticity is hugely important, yes, but for most potential use and reuse [they] won’t be necessary – and then not for the long term” (D206).

5.6.6.1.1 Checksums and hash digests

The most frequently mentioned indicators of authenticity were fixity checks – checksums and hash digests – to prove integrity. Of all the cryptographic mechanisms of control, checksums, with their ability to prove the bitwise integrity of digital material over time, seem to be the most used. “In my 30 years of experience, [checksums and redundancy are] the only thing[s] that work” stated D043. Archivists use fixity checks at point of acquisition, and rely on checksums when moving objects: “when you move a digital object you aren’t actually moving a digital object but creating a new object that purports to be the same thing.” This is integrity at the bit level – if the checksums between two objects match, then there has been no change of any kind between the two objects. Again, the issue of control is raised – “checksums are about control – it all comes down to what we can control” (D148).

Maintaining authenticity is complex, particularly when many hands are involved in the process of record transfer to an archival repository. In the transfer process, it is easy for the authenticity of the records to be violated. Archivists are preserving the original bit stream and limiting the number of people who are involved in the transfer process. In one case, this means there is one person responsible for records coming in, and one for running checksums. “Checksums on the electronic records is what we go by; if they are correct in relation to the bit stream we consider them to be authentic" (D123). In this case, the originality of a digital record is equated with an unchanged checksum.

Over the long term, fixity checks need to be maintained just like any record. Long time periods for preservation pose a problem, e.g. a migration from 16-bit Windows 3.1 up to a 32-bit operating system may seem trivial, but will have changed the hash of documents in the system, so the document, while still authentic, has been modified, and therefore the hash is not reliable (D112). Preserving a chain of hash values would alleviate the problem, but only if checksums or hashes were generated at every migration. Even if fixity checks are generated over time, how is the authenticity of your hash established? If the original object is migrated and the hash regenerated and stored alongside the object, it is necessary to know the system access controls and how the hashes were generated and preserved – were they checked into some vault to prevent modification: "If you can describe to me the system in which the hash is maintained separate from the digital object, then I might be able to trust that…" (D112). It still comes down to controls around procedure and the trustworthiness of the people maintaining the system ("talk to system administrators and they'll say, yeah I've changed records").

5.6.6.1.2 Digital signatures

> *Algorithmic? never use them. (D334 – Canada).*
>
> *Checksums are required; we don't really have a way of dealing with digital signatures. (D112 – United States).*

> *By law we have to use digital signatures. (D187 – Spain)*

> *Digital signatures? We hope we never have to deal with them! D148 – Canada)*

Few cryptographic mechanisms have created as much controversy for records professionals as digital signatures. A digital signature offers non-repudiation and confidence in the identity and integrity of the record to which it is attached, but presents challenges to long-term preservation of that record (Blanchette 2006). Whether required by law for legal transactions as in some countries (e.g. Spain) or their authority recognized by various electronic transaction acts (e.g. United States), records professionals view them with suspicion and consider them useful only for immediate or short-term validation. One interviewee in Spain reported that the digital signature is "only for the moment", and it is the fact of existence of the signature that is integrated into the record system. Another commented that they may compromise the preservation of the record (D187). While the use of digital signatures may be required by law in public administrations, governments still try to use other systems of validation because citizens don't use digital signatures.

> *Inside the administration we use attached signatures, but for communication with citizens we have to use embedded signatures. (D026).*

While European jurisdictions have considered digital signatures in some depth, one interviewee from North America noted that the law doesn't recognize digital signatures or doesn't understand them, and the legislation hasn't kept up with the technology (D123). In Canada, digital signatures have become less important since provincial and federal governments have introduced Electronic Transactions Acts in their jurisdictions which, if you are creating, receiving, and handling transactional records in an electronic environment, allow for a scanned signature to be satisfactory proof of the transaction (D148 and D429).

5.6.6.1.3 Secure transmission

Like digital signatures, secure transmission provides short-term confidence about the integrity of digital material. Secure transmission ensures integrity through space and time, and is used in conjunction with other cryptographic methods, such as checksums to prove chain of custody (D123). Secure transmission may also be used to address concerns for privacy and confidentiality more than authenticity (D148 and D429). It is also used because of the ability to transmit large files, irrespective of concerns for their security (D334).

As with other technological measures of security, implementation is under development in many organizations. As one public administration moves to an increasingly online presence, secure transmission will be relied on much more in the coming year (D026). Another is experimenting with receiving records in various ways, with SFTP (secure file transfer protocol) being the most common.

> *[There are] six different transfer methods identified right now, working out the metadata (D148); It's all very fluid – it won't be any one way in the foreseeable future (D429).*

5.6.6.1.4 Time stamps

Information about time – time of creation, time of modification – can be important in establishing a chain of events, and proving chain of custody. Interviewees who mentioned time stamps, however, were not talking about capturing trusted time stamps, whose fixity would afford them a measure of trustworthiness, but time logging attached by native software. These time stamps may change in any number of circumstances (D131).

> *I opened a TIFF file and made changes, and reran the checksum – it didn't change the time stamp – they are not fixed (D043).*

Time stamps logged by software in a native file can help establish version history, particularly if files are on removable media that are not themselves reliably labeled, for example if several disks have all been labeled "final draft" (D087). Forensic tools such as Fiwalk and Bitcurator reporting, in increasing use in archives to help identify, process, and manage digital acquisitions, can help establish the trustworthiness of such version history through file-level data.

5.6.6.2 Information About the Software

The degree of importance placed on the availability of information about the software that created records or manipulated data depends on the community of practice. In the scientific data community, this information is important. One interviewee gave as an example the testing of a genetic sample. When the sample is run through three different software applications, it may return three different results. Interpretation depends in part on understanding the role of the software in the results, hence the need for information or documentation about that software (D103).

Other records professionals want format information but care little about the software that created it (D148 and D429). Feedback on a draft transfer form for archival deposit developed for digital materials was that information about the software is not important or relevant, “we don’t need to know,” while format information and related metadata were considered important (D043). However another interviewee felt that information about the software that created a file was good to know and cited PDF (many versions of the Adobe software as well as many third party sources of PDF) as an example (D043).

5.6.6.3 Access and Security Measures:

> *It’s not the content that is important, but the security of the record – I don’t care about the author or date, but how [the record] has been protected over time (D043).*

Interviewees identified access and security measures as very important in establishing integrity, although with the caveat that one only knows the current state. Evidence of strong security today does not imply proof of strong security yesterday or tomorrow. Tracking changes made to security measures therefore is important. However, the same interviewee who raised this caveat also stated that it is mostly security and access controls that protect the records ((D112). All interviewees stressed the fluid and emerging nature of digital records management and preservation activities. Most were dealing with short or intermediate term issues, and both social and technical measures of control were constantly developing. For example, in one situation concerning active records in a public administration, access was strictly controlled to protect the records, but audit logs had been fully implemented only in critical processes, specifically in databases containing personal data (D026). Another organization considered access controls and security measures very important and was in the process of trying to make PREMIS rights actionable. The development of the program was driven in part by what it was possible to control, and this was being worked out iteratively (D429).

5.6.6.4 Audit Logs

Opinion was mixed among the three interviewees who spoke about audit logs. This offered a clear example of the difference between practice and belief. Two spoke about their belief in the importance of audit logs, but described their secondary use at this time in their organizations (D026 and D112). As one put it, "audit logs take a back seat – they're still important, but policies and procedures lay the foundation to manage records – I think these are more important than the 'in the trenches' work" (D112). This also raised the issue of high level planning versus implementation – concept versus practice. To the third respondent audit logs were vitally important, and had been used by this respondent in helping solve a murder with time information from the logs, validated through a check of multiple systems (D043).

5.6.6.5 Forensic Applications

The survey did not question the degree of use of forensic tools in the practice of records professionals. In the interviews, one interviewee mentioned using these tools.

> *We run fiwalk and other forensic applications – that generates a lot of data; we don't yet have the capacity to extract from that output and make it either human or otherwise readable, so if we could get to that stage, we could take that object level metadata and put it in something that could be a system that a researcher or user could see and make sense of (D087).*

Forensic tools provide records professionals with valuable methods of data extraction and analysis, and are in common use in many repositories (John et al. 2010; John 2012; Rogers and John 2013).

5.6.7 Redundancy

> *Even with checksum validation you have to seal it and store in multiple places – it's always "easy" to say a digital file is original - the only way to prove is through checksums and redundancy. In my 30 years of experience that's the ONLY thing that works (D043).*

Only one interviewee talked about redundancy as a means to assess authenticity of digital material. However, he was unequivocal. When the authenticity of an object is called into question, assessment is normally done by having more than one copy and comparing them – through redundancy. Redundancy may be accomplished by copying a file to several different media, for example, paper, microfilm, and scanned digital, or by distributing it in many copies across a network (although if it is in the cloud, it needs to be distributed among several cloud service providers and not just a single vendor in different places).

5.6.8 Trust and Control

Ensuring and assessing authenticity depends on a balance of technical mechanisms and social measures. The social measures (indicators and concepts) form the foundation of authenticity and the technical mechanisms allow for their implementation. Trust and control combine to create a framework for authenticity. Creators/preservers and users depend on the trustworthiness of people maintaining the systems (D112).

There are also cultural contexts within different communities of practice (different epistemological communities) that encourage trust in certain indicators and practices. One example is found in the relationship between government open data projects and scientific data portals. The current driver for open data in some jurisdictions is, as far as possible, to push datasets out and make them available. Questions about authenticity and accuracy have a lower priority than they once would have had, when public sector staff would worry about the quality of the data. Now the goal is to get the data out and let the users worry about data quality (D206). Some scientific communities will not use open data for this reason unless the source is known and the metadata robust, although this may be changing and open data is normalizing (D103), at least in some areas. To D103, when assessing data quality (authenticity) it is important to know the producing institution, and to know that the scientific community to which that dataset belongs has a culture of describing its data well, adding caveats, and telling all the issues, such as the instrumentation by which the data are collected. This cultural context of producing and sharing provides the user with a greater sense of trust in those datasets, as compared to many or most open data portals, in which the metadata are so limited that decisions on quality are impossible to make. Lineage doesn't apply to the data only, but to the producing institution and trusted institutions have a long lineage of good practice D103).

This sounds like transparency, which is considered an important goal in open government and open data movements, but is dismissed by D103:

> *We wouldn't have ever called it transparency – we would have just called it good practice – that's what we call it! (D103)*

One interviewee conflates authenticity with trust, and suggests that in order to trust a recordkeeping system, in particular an EDRMS, one needs a complex system of controls. For his organization, the scope of an EDRMS would be to serve 22,000 people, while the current records management database currently serves a mere 800.

> *You can rely on a certain amount of local knowledge [with 800 people], but 22,000 requires a managed or controlled process - baked into the tools and systems that a business uses - that's where I'll be looking to have less of a human connection and more of a structured system (D334).*

However this same interviewee also noted that given the volume of information, the number of people involved in creating and maintaining information, and the frequency of technological change, "if you don't want to trust the organization then there is nothing that will establish authenticity – if you can't trust your trusted users, then you've got nothing."

It comes down to controls around procedure, and the trustworthiness of the people maintaining the system (D112). In the organization of one interviewee, the system was designed intentionally from the perspective of work practice – the people need to work *in* the system, not *outside* the system. It was designed so that it is not possible to create official records outside the system; this has been done through controls implemented in critical points in the process, mainly the points of political decision.

> *[It is] very important that all essential metadata is managed in the RM system where we have control – always, always control! (D026)*

In another institution the interviewees are choosing strategically the things they can control and those they cannot, for example to log through time any changes made to file formats. They are more open to documenting and using metadata in their context (of the institution's archives) than requiring creators to do so. They are taking checksums, doing fixity checking, maintaining a log of all the technical components of the system, software, etc. to prove over time that the archives is a trusted digital repository (D148 and D429).

5.6.9 Cloud Computing

There is little trust among interviewees in any assurance of authenticity when records are maintained by cloud service providers. Several problems emerged from the interviews: ownership, security/redundancy, control and the ability to know if changes have occurred.

Storing records with a single vendor presents the same need for redundancy as with digital records maintained locally. If one uses a single vendor, then one is vulnerable. No single system is faultless, so cloud storage can only approach trustworthiness if material is duplicated across vendors (D043). Some organizations mandate that everything is held within the facility – by regulatory requirement (D112). Neither would this interviewee entrust his own records to the cloud: his is such an extensive collection that the ability to perceive problems would be lost if the records were out of his control (in a RAID system with backups). In the cloud, inadvertent modification and inadvertent data loss would be undetectable and subject to trust that the service provider would disclose the event.

Organizations may use cloud services for non-essential records when the risk, measured or perceived, is low. Trust has to be balanced against consequence (D334). Communication through social media, or presentation of resources like photographs that are considered to be of low consequence, is viewed with a degree of confidence. Are tweets from this organization's twitter feed managed as records? Yes, the archives maintains a twitter feed and logs it – it is deemed to be low risk. Similarly a Flickr site offering low resolution images is enormously popular. This interviewee's organization has implemented a crowd-sourcing tool for people to access the public database and mark up the contents ("we expect the risk to the [institution] is low"). This institution is identified as not an early adopter of cloud services. The biggest challenge is around intellectual property.

Much depends on the organization's ability to negotiate terms with the service provider. D097 has been involved in writing cloud service provider risk assessment matrices and policies. If it is a reputable third party providing a service and contractual terms can be negotiated, then this interviewee believes that the organization can be quite confident that the information is being held in a way that is trustworthy; this becomes unclear when the third party sub-contracts. This has not changed from third party storage of analogue records – it is not so much the service provided but how it's being controlled (D097).

Part of the negotiation should articulate what the chain of custody would be, identifying through a risk assessment where weak points are, and then testing.

> *One of my personal bugbears – the ownership and custody issue – organizations should always maintain ownership of their material and grant custody to the third party. Ownership should never be relinquished – that [bothers] me because so many organizations don't do that – they have to protect themselves (D097).*

In the case of some organizations, like universities, certain exploratory initiatives in cloud service adoption may be funded by grants. If base funding is not ongoing and sustained that is a serious underlying weakness in the long term. This is something CIOs are talking about and trying to plan for, at regional and federal levels, and is a concern when granting agencies start to impose requirements that research data be preserved and accessible but the infrastructure isn't there to make that happen in a sustained way – how can one say that research data will remain accessible when the repositories are funded by grants themselves? This is another reason to take care when deciding to use cloud storage regardless of laws governing where data is to be stored. Another problem in educational institutions is the increasing use of service providers for classroom and learning applications (learning management systems). The external service providers of learning technologies and services are storing student and instructor data on their servers; in these cases, who is the owner of the information, how long is it being retained, what are the agreements about disposal, and what are the security requirements in place for authenticity? (D429)

> *In our point of view, we want to have some sort of control. For example, using Amazon? apart from the fact that it is illegal [in our jurisdiction], it's about trust, what can we control, when it comes to storage I don't ever see us using 3rd party cloud storage (D148).*

5.6.10 Legislative Frameworks

Interviewees were asked whether there was a difference in the framework supporting authenticity of records between common law and civil law countries. This author expected a clear answer, but this was not the case. The legislative framework in Italy has already been introduced. General legislation clarifies the minimum amount of metadata required for identity and integrity, including classification, reference code obligations, and specific rules for records management systems. Information related to provenance, identity, and context must be well defined and must be written and exchanged according to specific rules (D073). The interviewee noted strong cooperation in Italy between archivists, records managers, and legislators. However, Italy appears to be the exception rather than the rule in civil law countries. In another European civil law jurisdiction interviewees reported that there has been a change over the last six or so years, and today a common trend is to ignore the law and take the path of least resistance (D025). An interviewee from a third jurisdiction noted that the regulations are at a very high level and do not provide any specific guidance.

Some interviewees found solid legislative support for recordkeeping, while others noted a failure to keep up with technological advances. In New Zealand, there is a high-level legislative regulatory apparatus for designing systems in the public sector, and guidance at a more granular level for recordkeeping systems in ISO 1617 Part 2. A template for writing records policy, and a metadata standard with technical requirements and a schema provide additional support. This provides a good basic framework to work from and test the systems against minimum requirements. In Australia, the UK, the US, and New Zealand standards assess and accredit recordkeeping systems from vendors, but throughout the 2000s this became increasingly untenable (D097).The situation is more complicated in the UK and North America, where legislation regulating records generally does not address authenticity specifically (D206, D087, D148, D429).

5.6.11 Critical Incidents

> *Legal crises are very positive to help solve these problems, and to get more money [to help solve them] (D026).*

> *Consider [Organization X] – they have the standard because they've been to court! ...It costs millions if you make a mistake (D046).*

As stated previously, trust is considered a corporate asset, to be protected and nurtured. Risk management measures the effect of breach of trust on reputation and resources, and change is frequently driven by critical incidents. Unfortunately, record authenticity is infrequently questioned (D334, D046), and so establishing authenticity is often not a priority. It becomes a byproduct of broader goals, for example security or preservation.

One interviewee remarked that more than one critical incident is required to promote change.

> *You need multiple critical incidents because I've worked now for close on 30 years… and I've seen some critical incidents happen, but in isolation it has no lasting impact; I would say the only time it can have an impact is if it is going to have a significant impact on the public reputation of the institution, assuming that's important to them, and a significant financial impact, especially if those two are combined, and let's not forget that in many organizations there's a lot that's flying under the radar – [things are] not on peoples' radar, or are being managed in a way that makes [the problem] go away (D429).*

D429 explains that the vast majority of records used each day is routine and administrative and will not attract that necessary level of scrutiny that would be brought to bear in a legal proceeding. So, in the absence of jurisprudence, and without being taken to court to discover that records are not admissible because they have not been created or managed sufficiently well, internal legal counsel tends to take a very realistic approach and opinion on most records and implementing complex measures to establish record authenticity will remain a low priority.

5.6.12 Traditional Archival Model of Authenticity

The traditional archival model of record authenticity, dependent on establishing identity and demonstrating integrity, may still be considered to form the foundation of record authenticity in the digital environment, but requires extension or adaptation to accommodate the challenges presented by digital technology. It is, of course, no longer possible to evaluate the authenticity of a record by visual inspection.

> *We can't immediately apprehend the authenticity without mediating the object, having it represented to us – so we have to choose something else – this is something named receipt.pdf, open it up – we make this evaluation – we are trying to establish identity at that point, rather than is this really the object I care about – we don't have the ability, easily, if this is an authentic receipt – it always looks the same, for one thing – no fading, no folding – so we have to come up with other ways of giving ourselves some assurance that it is a real thing (D112).*

Traditional models are still immensely helpful, but they require significant transposition to apply in the digital environment. Furthermore, some believe that authenticity cannot be considered a binary state – some things are "authentic enough" (D334). The same issues arise but manifest themselves in a radically different way. With enough money and time it is straightforward to test the authenticity of a physical document, but it is much more contentious in the digital world because of so much duplication, surrogacy, and different processes (D097). The issue of identity also arises: when the same document (content) exists on two different devices, are they the same? They are physically separated and the contextual information needs to be identified and assessed (ibid.).

However, this is a problem that exists at a point in time and is ever changing. Twenty years ago we might have been talking about thermal paper – longevity issues are still the problem (D097). The digital environment has introduced new activities that did not exist for paper – it is more about expanding the model rather than coming up with a new paradigm entirely (D087).

Other interviewees see the difference between traditional (paper-based) models and digital realities to be greater. D206 believes the traditional model is declining, partly because we are "not talking about paper records any more, we are talking about digital, and it is just a different beast." We need technical instruments to ensure some aspects of the records that are not necessary in the paper world, so the liability of the digital records is more complicated, particularly for active records (D026).

Our assumptions about records and systems design from the analogue world influence our approach to the digital environment. One interviewee wants to believe that the three elements of record authenticity, integrity, identity and context, still form the foundation. However, this interviewee thinks it is unclear to what extent we can be successful at applying this model, for reasons that include human nature and personal behavior, organizational culture, and technological capacity (of people and systems) (D429).

The level of detail recommended in documents like the Creators' and Preservers' Guidelines of InterPARES 2 is considered prohibitive by to another interviewee (D148).

If that is the level required, in his opinion, “then we have to stop archiving, because no one is going to live up to those standards.” Furthermore, when authenticity is considered from a legal perspective, it is surprising how little legal counsel seems to be concerned – there is a disconnect in trying to equate archivists’ concerns about authenticity with lawyers’ concerns. If records are made in the course of business, then according to legal counsel they are fine (D148).

5.6.13 Summary

The interviews confirmed that records professionals take an approach to authenticity of digital records that is strategic, realistic, and pragmatic with respect to managing expectations of users and stakeholders and behavior of creators. For many, the social indicators form the core foundation for authenticity, but the methods used to effect the authentication of a body of records would be the technical tools, most commonly checksums or hash digests. Regarding the reliance on technical mechanism, D429 states that “for as long as we have had the capability for creating digital records, people have fallen back on the belief that the solution lies in technology without considering that all the solutions may not be provided by the technology – these social factors need to be incorporated to provide a holistic approach… generally people are resistant to the social factors being incorporated in the work environment because they have become so disconnected from what we know of this structure we knew existed in the analog world which broke down very quickly 30 years ago when digital technology was introduced into the workplace.” Trained clerical staff in the past handled much of the records management functions, but today’s knowledge workers do not want this task. D148 suggests that what we need are electronic filing clerks who, for example, will take the time to classify records (D148).

> *[There is] a disconnect going on that is intensifying with time – people work at the desktop in silos without taking a more corporate view at whatever level – program, office, department, institutions; factor in publicly funded institutions with prolonged resource constraints, the people who are doing the work to deliver services and programs have less time today than they may have had in the past; factor in as well that the majority of organizations have not had any critical incidents or disasters that have acted as an object lesson for why these things need to be in place (D429).*

6 Analysis and Discussion

6.1 Introduction

This chapter presents a discussion of the research findings as related directly to the questions that have guided this study. To begin discussion of the results of the survey and interviews viewed in the framework of the traditional archival concept of record authenticity, it is worth reviewing the starting premise of the study. The findings are then discussed in the context of each research question and sub-question.

6.2 Review

The ability to assess and protect the authenticity of records and data is axiomatic to archival science, and enjoys a centuries' long theoretical foundation. The object of archival science is records, defined as documents (recorded information) made or received in the course of activity as their instruments and byproducts, and set aside for future action or reference. These records may be in active or semi-active use in the custody of their creator, or they may be in the custody of an archival program or institution, maintained and preserved by a third party.

Authenticity is defined as the quality of a record that is what it purports to be, free of tampering, corruption, or other unauthorized change or manipulation. An authentic record is as reliable as when first created (reliability being the responsibility of the creator, ensured through controls over creating persons and procedures). In the digital environment, authenticity must be actively protected, as records are subject to corruption through time (because of bit rot, media failure, obsolescence, etc.) and across space (because of transmission errors).

In archival science, records that are relied upon by their creator in the course and for the purposes of its activity are presumed authentic. This is codified in evidence laws through the business records exception to the rule against hearsay evidence (cf. *Canada Evidence Act, RSC 1985, c C-5* 2015, s. 30). Standards such as ISO 15489 "provide "guidance on managing records… of originating organizations, public or private, for internal and external clients" so that "appropriate attention and protection is given to all records, and that the evidence and information they contain can be retrieved more efficiently and effectively, using standard practices and procedures" (International Standards Organization 2001, Part 1, vi). However, digital technology challenges this presumption by putting records' authenticity at risk whenever the records' status of transmission (i.e. degree of perfection of a record's version) changes through transmission across space or through time (MacNeil 2000a, 28; Bearman 2006, 24–34).

Regardless of medium, the determination and assessment of authenticity depends on the circumstances of record creation, management, and use, and on the framework of subsequent preservation. According to InterPARES, to assess the authenticity of a digital record, one must be able to establish its identity and demonstrate its integrity. The identity of a digital record is established by the attributes of the object that, together, uniquely distinguish it from other records, while integrity refers to its wholeness and soundness, that is, the degree to which it is complete and uncorrupted. Based on archival diplomatics, InterPARES developed a set of benchmark and baseline requirements for supporting the authenticity of records and producing authentic copies, and guidelines for records creators and preservers. InterPARES recognized and articulated the difference between the form in which a document is viewed by a person reading it, and that in which it is stored in the electronic system. The layers of abstraction introduced by the technology between the physical and logical record have implications for the assessment of authenticity. This significant difference between paper and digital records is at the root of the challenges of ensuring and protecting the authenticity of digital records, and respecting the legal system's conception of documentary evidence.

Archival diplomatics has been shown to offer archivists a powerful methodology for analyzing digital records but it is not without its limitations. As MacNeil has noted, archival diplomatics assumes that the object of study is reducible to a set of well-defined elements, but InterPARES 1 research found this to be rarely the case. The complexity of electronic systems is too great, record boundaries too amorphous, and too contextually determined (MacNeil 2004, 219). Duranti confirmed that digital diplomatics alone may not be sufficient to understand the challenges posed by information inscribed by increasingly complex digital systems (2009, pp.42-43). This was seen in the issues faced by the Domain 2 Task Force of InterPARES 2, and becomes even more evident as technology continues to develop. Traditional concepts of provenance and identity are severely challenged by the default of anonymity on the Internet. The identity of creator, author, writer or originator may be obscured and separated from the message by virtue of the layers of technology that mediate between physical person and transmitted document. Integrity, once presumed from the controls on the procedures dictated by the creator, now must be assessed separately from provenance and identity, and at both the physical (bits) and logical (meaning) layers of the record.

If the requirements for ensuring and assessing the authenticity of records were available through archival science, why would records professionals and researchers still be grappling with implementation?

6.3 Problematizing Authenticity

Digital preservation research continues to investigate the nature of digital objects, including records and data, and the attributes that may support the presumption of their authenticity. There is a desire to find systemic answers that can be standardized in practice as financial, governmental, or health critical infrastructure, while social networks systems increasingly rely on complex integrated, interdependent (although not necessarily interoperable), distributed networked systems (cf. Factor et al. 2009; InterPARES Trust 2015, Preservation as a Service for Trust). Still, current means of assessing authenticity do not offer quantifiable measures, and generalizable models that can reduce the problem to concrete, atomistic elements are elusive.

The hypothesis of this study was that, despite clear guidance from archival science on the means of ensuring, managing, and assessing record authenticity, a guidance reflected in the products of several large-scale, significant and influential research projects into the topic of authenticity in the context of long-term preservation, the theoretical recommendations of these projects are not being consistently applied in practice, and records professionals are often unclear about how to define authenticity, how to protect it, and how to assess it (i.e. how to authenticate records and data). The survey and interview findings support this. They show that records professionals rely heavily on traditional heuristics in the practice of ensuring authenticity of records but believe in the greater value of technological mechanisms for assessing authenticity and authenticating records. Records professionals, traditionally the trusted agents of record control (trustees), have frequently become the trustors, placing their trust in technology of which they may have little understanding and even less control.

This is a complexity problem. The issue of record authenticity in the digital environment is a complex problem, a ‘messy’ problem. Attempts to reduce it to the identification of concrete attributes of records and data that can then be systematically analyzed and controlled minimizes the impact of the human factor. Actor-network theory argues that knowledge is a social product, rather than something that can be described and guaranteed through the operation of scientific methods (Law 1992, 380). The intuition of actor-network theory is that the world is a messy place, and that dominant methodological approaches try to suppress the mess. According to Law, simplicity will not help us to understand mess (Law 2006, 2). What would method and its politics look like, he asks, “if it were not caught in an obsession with clarity, with specificity, and with the definite”? (ibid.)

Soft systems methodology (SSM) teaches us that we need to think less in terms of “real-world systems in need of repair or improvement” and more in terms of human activity systems, human situations in which people attempt to take meaningful purposive action. Human activity systems are modeled as sets of linked activities which together can exhibit the emergent property of ‘purposefulness’. This leads to the understanding that there are many interpretations of any declared ‘purpose’ and that it is necessary to decide for each selected purposeful activity the perspective or viewpoint taken. In the context of this study, this idea might be expressed as viewing the purpose of managing record authenticity as part of a system to manage risk, or preservation, or evidence. SSM moves away from the idea of an ‘obvious’ problem which requires a solution, to that of working with the idea of a *problem situation* (Checkland and Scholes 1999, A5–A9).

In this study, the problem situation concerns how records professionals view, ensure, manage, and protect authenticity. Thus the focus is on work practice, consistent with a pragmatic worldview. Schatzki “conceives of practices as embodied, materially mediated arrays of human activity centrally organized around shared practical understanding” (Schatzki, Knorr-Cetina, and Savigny 2001, 11). Practice research develops knowledge empirically, by studying the actors, their actions, and related material, linguistic, and institutional elements (Goldkuhl 2004; Goldkuhl and Lagsten 2012). “Positioning a practice lens” involves an empirical focus on how records professionals act, the relation between their actions and organizational structures, and the constitutive role of their practices in producing social reality (Feldman and Orlikowski 2011, 2). In this study, such focus involves examining how and why records professionals use and apply indicators of record authenticity, and the context of doing so, which is conditioned by experience, professional identity, and organizational or domain expectations.

6.4 Findings from Survey and Interviews

It is the mark of the educated man and a proof of his culture that in every subject he looks for only so much precision as its nature permits. Aristotle[12]

One of the premises of the UBC-MAS Project was that “policies, standards, and requirements for the management and preservation of trustworthy electronic records cannot be properly designed if the entities concerned are not clearly determined and recognizable” (Duranti, Eastwood, and MacNeil 2003, 5). This assumes that the record is the unit of analysis, but in contemporary records management practice, the entity “record” is understood in a multiplicity of ways, many of which do not conform to the definition from archival science. In fact, the policies, standards, and requirements that records professionals have to work with do not regulate entities that are clearly determined and recognizable. This fact is an embodiment of the messiness of the problem mentioned earlier. This study adopts a different unit of analysis, that of the indicator of authenticity, and questions its use and application.

Applying a practice lens, then, means that the survey and interview results are considered in the context of the following statements:

- simplifying the problem situation does not help us to understand it;
- record authenticity is to be considered as part of human activity systems comprised of actors, actions, and material, terminological, and organizational elements;

[12] In *Ethics of Aristotle: The Nichomachean Ethics* Translated (Hammondsworth, U.K.:Penguin, 1955), 27-28; quoted in (Buckland 1994, 347).

- the determination, maintenance, and continuing assessment of record authenticity can be seen as actions in work practice; and
- these actions can be traced to their practical consequences.

Furthermore, a pragmatic focus on practice means an acknowledgement of the full dialectics between knowledge and action:

- proper action is knowledgeable action; and
- proper knowledge is actable knowledge (Goldkuhl 2004, 24)

We will now examine the findings related to research question 1 and sub-questions 2 through 4, leaving sub-question 1 for later discussion.

Research Question 1: What elements of the context, content, and structure of digital records and data do records professionals use and rely on in order to determine and manage authenticity?

> *Sub-question 2: How do records professionals approach the issue of authenticity of digital records and data in their work?*
>
> *Sub-question 3: What elements or indicators of authenticity do records professionals rely on or believe are important?*
>
> *Sub-question 4: Does experience with authenticating records affect what indicators records professionals rely on or believe are important?*

These questions address the broad range of practice conducted by records professionals working in different professional roles, across sectors, and with different record types at different stages of the record life cycle. Human agency that affects authenticity is not only that of the records professionals, but also that of other stakeholders with an interest in records. These stakeholders may be record creators or users of the records within and outside the organization. This requires an overview, or a macro view, of the 'human activity system' in which record authenticity is considered in a network that includes agents, practices, and relationships. Agents may be the record creators, records professionals, or records users. Practices include the responsibilities, accountability, expectations, and work practices of the agents. Authority relationships exist between agents, each relationship affected and conditioned by the actions and expectations of the agents.

The consequences are seen in interviewees' comments about records only needing to be "authentic enough." When interviewees talk about a continuum of authenticity, or being authentic enough, they are really talking about the archival diplomatics concept of reliability. This is antithetical to the view of archival diplomatics, which recognizes authenticity as binary. A record is either authentic or it is not, although a presumption of authenticity is drawn on the basis of circumstantial evidence, which "can be nonexistent, sufficient, strong or absolute. Either something is what it claims to be or it is not, regardless of how close it is to its first instantiation."[13] A determination of authenticity is based on intention (intended purpose or use by the creator or another person), according to one interviewee. This begs the question of who is judging authenticity. Of course

[13] Email between the author and Luciana Duranti.

authenticity relates to the *record*, however, the *requirement* that a record is known to be authentic (as defined by archival science) is dependent on who is concerned.

This leads to linguistic difficulties that could be avoided if we recognize authenticity to be a term of art,[14] and instead talk about trustworthiness, which is so often used synonymously with authenticity. Duncan discusses this linguistic morass when using the term to describe records, analogue or digital: “the application of authenticity is hindered on several fronts: its subjectivity as a term; the subjectivity of its application; and the existence of intermediary and opposite states” (Duncan 2009, 101). When authenticity is strictly defined, as it is by archival diplomatics, then assessing authenticity is a matter of measuring the record in question against the ideal model. However, the empirical evidence shows that many records professionals do not approach record authenticity in this way. Some do not use the term authenticity at all. When we look at the practical consequences of their actions, taken for the purpose of ensuring, maintaining, or assessing authenticity (which they may refer to by another term), we see that integrity is much more frequently what their actions are intended to protect. If authenticity was the concern of the diplomatists of the 17th and 18th centuries, we might postulate that today that for records professionals, integrity as guaranteed by security is the new authenticity.

One model of trustworthiness records professionals might consider proposes that trust, as colloquially understood, is equivalent to confidence, defined as the reasonable expectation of a trustor (person or device) that the future behavior of a trustee (person or device) will be beneficial to the trustor. If the trustor has no means to influence the

[14] A term of art is a term that has a particular meaning in a domain of practice (Merriam-Webster 2015).

behavior of the trustee but relies on intention, then the trustor must *trust*, but if the trustor has influence over the trustee, then the trustor has *control* (Cofta 2007a, 173; Cofta 2007b, 28). Confidence, therefore, is composed of a balance between trust and control, although control itself must be subject to an assessment of confidence. Cofta's model is designed to analyze agent-agent trust. It serves as a tool for a better understanding of the human process of the decision-making regarding confidence, and to bridge the gap between the sociological and psychological perception of trust (confidence) and the perception of confidence as a measurable property of digital systems (Cofta 2007b, 27).

This model is transferable to the concept of agent-record trust. An agent can be confident of the identity and integrity of a record by measuring the mechanisms by which its identity is determined and its integrity proven. This is the archival diplomatic definition of authenticity, and empirically one that records professionals agreed to, even though their use of the term 'authenticity' does not match the archival diplomatics one. The indicators of authenticity presented to the study participants as social or technical mechanisms fit this model as well, with some adjustment. Technical indicators of authenticity are mechanisms of control, while social indicators are less so and may be subject only to trust in the record system (including the human agents involved in the system), due to their existence external to the record (such as policies and procedures) or subject to human application or interpretation.

We can now discuss sub-question 1.

> *Sub-question 1: What are the domain definitions of, and relationships between, terms such as authenticity, identity, integrity, authentication, provenance, lineage,*

traceability, originality as they relate to how records professionals view authenticity?

The premise for the UBC-MAS Project was that the identification of the criteria, techniques, and methods needed to solve the problems posed by the use of electronic information systems for carrying out business "cannot derive from purely pragmatic or ad hoc decisions but must be rooted in principles and concepts that can be applied in different situations and various contexts" (Duranti and Eastwood 1995, 214).

"The first step in the study was to refine more clearly the concept of *integrity*, which was internally broken down in two concepts, *reliability* and *authenticity*. This amounts to saying that preserving the integrity of records means ensuring that they are created reliable and maintained authentic" (1995, 215, italics mine).

The ideal model of an authentic record developed from archival diplomatics and expressed in the results of InterPARES offers clear and precise definitions of these terms and their relationships (with the exception of lineage and traceability, which have specific meaning to the science data communities). The term used most consistently by study participants is 'integrity' in connection with technical mechanisms. When talking about actively protecting authenticity and about authentication, most participants mentioned technical indicators exclusively. In contrast, in the course of daily work practice, authenticity was more frequently presumed from social indicators.

However, the empirical data suggest that in fact most practices are driven by pragmatic concerns, and many decisions are indeed ad hoc. This suggests that a closer examination of the interpretive nature of work practice may be necessary to reconcile theory and practice.

Many participants equated authenticity with originality, focusing on record content. If a record remained identical in terms of its content through time and space, they would presume it to be original and therefore authentic. However as Duranti notes: "If each is what it claims to be and they claim to be the same, they are the same, except that they cannot claim to be the same because at least one metadata would different, that related to location. So it will be the same but in a different status of transmission, a different copy or version. So the issue is not which is authentic but which is the most authoritative."[15]

The remaining research questions were about the use and effectiveness of the traditional model of authenticity:

Research question 2: Is the traditional model of authenticity of records used in the digital environment and to what degree?

Research question 3: Is the traditional model of authenticity sufficient to support a presumption of authenticity in the digital environment over time and across technological change?

The application of a traditional model assumes one of two things: 1) digital records can be understood as analogous to their paper forebears and are thus subject to the same logic of analysis, or 2) digital records can be analyzed and decomposed into their component

[15] Email between the author and Luciana Duranti.

parts in such a way that the foundational theoretical concepts can be applied or adapted. There is a subtle difference between these two approaches. There are some seductive arguments in support of the first assumption: treating digital records analogically. For example, medieval chanceries controlled the production of knowledge through technology (ink and parchment), procedures, and skills (i.e. literacy). Just as confinement in the most remote and inaccessible part of the ancient Rome archives amounted to automatic authentication, security technologies may confer the same presumption of authenticity on digital material. However, the complexity of digital technology and the unintended consequences that may arise from treating digital records as analogous to paper records suggests caution. In the interviews, the argument of analogy was raised with respect to cloud storage as well. D097 claimed that nothing really has changed with respect to 3rd party storage. The owner of the records is still at the mercy of a third party. However, the complexity and lack of transparency has changed to such a significant degree that this argument may easily be challenged.

The second assumption is also challenging because of the complexity of digital technology. New models of authentication are being tested, arising from the affordances of the technology itself; consider for example the Wikileaks model, which renegotiates the boundaries of knowledge and power that exist between citizens and the state. Some scholars argue that the function of traditional archives is weakening due to the volume of paper-based and digital records, the lack of adequate technologies to identify and capture records of significance, the separation of the process of recordkeeping and records management from records preservation, and the laws and practices increasing favoring secrecy and privacy (Findlay 2013, 15–18).

In summary, empirical evidence shows that authenticity is not defined in practice with the rigor of archival science in general, and archival diplomatics in particular, and that the use of social and technical indicators of authenticity is a matter of work practice that are conditioned by the larger context of human activity systems.

One question remains: why is authenticity not more seriously considered by practitioners, when it is foundational to archival science and the subject of much research, past and ongoing? The answer emerging from the interviews is again a pragmatic one: there has not been a sufficient number of critical incidents to galvanize action broadly outside the research community. Until crises of record authenticity have a big enough impact on people's lives and organizations' financial interests, the disconnect between theory and practice is likely to continue. The courts are the crucible in which authenticity will be tested when that happens.

6.5 Contribution of This Study

This study contributes to the discussion of record authenticity by its empirical design, linking theory and practice. The importance of creating, maintaining, and preserving records and data that are authentic has never been in question. However, though the means of achieving this goal has been the subject of much research, there has been little structured investigation of the implementation of its recommendations. This study charts the ideal model of record authenticity from the perspective of archival science to specific implementations and beliefs of practitioners, and finds that there are myriad perceptions about the nature of authenticity and its management. Some of these conform to the engineering problem-solving approach of abstraction that is the basis of much records research and several standards, while others are oriented more closely towards soft systems methodology, which is based in pragmatism and accounts for the human factor. The idea that authenticity is reducible to a concrete set of indicators assumes the possibility of reduction of the complexity of authenticity as a component of trust to simple, generalizable assertions. This study finds that such an idea appears to be problematic.

The study introduces the possible differences between broad juridical systems, namely common law, civil law, and pluralistic/religious systems, in relation to issues of record authenticity. Although the results are inconclusive, this is an avenue for further investigation.

6.6 Next Steps for Further Research

Records may be viewed as action, communications, evidence, or memory. Each particular point of view brings with it its own assumptions about truth value and authenticity. What these points of view share is the desire to stabilize those elements of the record that allow us to trust them in their particular context. As MacNeil says, "Authenticity emphasizes a return to the essential, the finding of centres, the fixing of reference points, the certification of truth, and the privileging of the singular and definitive over the multiple and indeterminate (MacNeil 2001, 42)." If 'integrity through security is the new authenticity' when we enter the digital realm, then we must accept that our desire for order and stability is still strong.

Why, if authenticity is something to be protected and valued, do we not see the risks of its loss? According to actor-network theory, something simple – a working TV, a well-managed bank, a healthy body –masks the complex networks that produce it (Law 1992, 385). Therefore a record system that is responsible for maintaining records and is assumed to maintain them authentic is functionally invisible until there is a crisis that demands attention.

More research is needed that delves into the details of individual situations. Where this study was an empirical exploration, ethnographic studies may offer the possibility of developing thick descriptions of the issues identified in this study in specific contexts. To paraphrase Foscarini (writing about functional classification systems in banks), records professionals are well aware of the deep transformations in society, but when they act to ensure the authenticity of records for which they are responsible, they tend to forget that the supposed stability of authenticity indicators is only apparent, and that *actual* world practices may be quite different from the way in which laws, regulations, internal manuals of procedures, and the people themselves who are in charge of given activities articulate how work gets done (Foscarini 2010, 391). This study also indicates that in many cases they may not have forgotten about authenticity, but are restricted in what they can do or feel required to do.

There is another driver that should be considered and provides impetus for further research. This is the absence of crises of authenticity. These crises, when they come, will surely manifest themselves in litigation. In the legal literature, George Paul has decried the “authenticity crisis”, and yet the courts, by and large, continue to deal with the issue of authenticity by ignoring it or relying on traditional methods of assessment. The next phase of research should conduct a structured search of legal and administrative proceedings in which authenticity of records posed a serious problem. Authenticity crises lend themselves to the application of critical incident theory, which is an exploratory research design aimed at theory development (Münscher and Kühlmann 2012, 164).

6.7 Final Thoughts

This study began from a proposition, developed from research and experience, that authenticity of records is desired and its importance taken for granted, but its presentation and implementation are often varied, vague, messy, and even ignored. What emerged is the outline of a wicked problem (Kolko 2012; Rittel and Webber 1973), a messy problem (Law 2006), that does not respond well to traditional models and frameworks. The issue of authenticity in the digital environment is as much a social as a technical issue. Solutions to the problem of establishing and protecting authenticity are equally varied and messy. In some cases it may be expedient to use technical and procedural controls, in others it may be sufficient, or only possible, to use human and social means. The contribution of pragmatic theories such as actor-network theory, practice theory, and soft systems methodology to the discussion of authenticity of records and data is the ability to view authenticity as not a quality that adheres to such entities alone, but is the practical consequence, or the effect, of the heterogeneous network of materials and relationships that surround records and data.

Bibliography

ACA. 1999. "Code of Ethics | The Association of Canadian Archivists." http://archivists.ca/content/code-ethics (accessed 15 November 2014).

Alliance for Permanent Access. 2012. "About APARSEN." *Alliance for Permanent Access*. www.alliancepermanentaccess.org/ (accessed May 28. 2014).

Bearman, David. 1993. "The Implications of *Armstrong v. Executive of the President* for the Archival Management of Electronic Records." *American Archivist* 56 (Fall): 674–90.

———. 2006. "Moments of Risk: Identifying Threats to Electronic Records." *Archivaria* 62: 15–46.

Bearman, David, and Ken Sochats. 1996. "Functional Requirements for Electronic Evidence: The Pittsburgh Project: Recovered Web Site: Metadata Requirements for Evidence." http://www.archimuse.com/papers/nhprc/BACartic.html (accessed 12 December, 2014).

Bearman, David, and Jennifer Trant. 1998. "Authenticity of Digital Resources: Towards a Statement of Requirements in the Research Process." *D-Lib Magazine*, no. June (June). http://www.dlib.org/dlib/june98/06bearman.html (accessed 12 December, 2014).

Black's Law Dictionary. 2012. "Black's Law Dictionary Free OnLine 2nd Ed." *The Law Dictionary*. Accessed May 28. http://blackslawdictionary.org/ (accessed 6 December, 2014).

Blanchette, Jean-François. 2006. "The Digital Signature Dilemma." *Annales Des Télécommunications* Mai/Juin: 908–23.

———. 2012. *Burdens of Proof : Cryptographic Culture and Evidence Law in the Age of Electronic Documents*. Cambridge, Mass.: MIT Press.

Brannen, Julia. 2005. *Mixed Methods Research: A Discussion Paper*. Working Paper. NCRM Methods Review Papers, NCRM/005. Unpublished. http://eprints.ncrm.ac.uk/89/ (accessed 18 December, 2014).

Bryman, A. 2004. *Social Research Methods*. 2nd ed. Oxford ; New York: Oxford University Press.

———. 2006. "Integrating Quantitative and Qualitative Research: How Is It Done?" *Qualitative Research* 6 (1): 97–113. doi:10.1177/1468794106058877 (accessed 4 January 2015).

Buckland, Michael. 1994. "On the Nature of Records Management Theory." *American Archivist* 57 (2): 346–51.

Canada Evidence Act, RSC 1985, c C-5. 2015. Accessed February 15. http://canlii.ca/t/52cjv.

Carrier, Brian. 2003a. "Defining Digital Forensic Examination and Analysis Tools Using Abstraction Layers." *International Journal of Digital Evidence* 1 (4): 1–12.

———. 2003b. "Open Source Digital Forensics Tools: The Legal Argument." www.digital-evidence.org/papers/opensrc_legal.pdf (accessed 2 February, 2014).

Carrier, Brian, and Eugene Spafford. 2003. "Getting Physical with the Digital Investigation Process." *International Journal of Digital Evidence* 2 (2): 1–20.

Casey, Eoghan. 2007. "What Does 'forensically Sound' Really Mean?." *Digital Investigation* 4 (2): 49–50. http://www.sciencedirect.com/science/article/B7CW4-4NWNCSP-1/2/36717bc8a1dc225cfec6a4c835866999 (accessed 14 March, 2014).

CCSDS. 2012. "Reference Model for an Open Archival Information System (OAIS): Recommended Practice Issue 2." Consultative Committee for Space Data Systems. http://public.ccsds.org/publications/archive/650x0m2.pdf (accessed 2 December, 2014).

Chasse, Kenneth L. 2007. "Electronic Records as Documentary Evidence." *Canadian Journal of Law and Technology*, November, 141–62.

Checkland, Peter. 1999. *Systems Thinking, Systems Practice: Includes a 30-Year Retrospective*. Chichester ; New York: Wiley.

Checkland, Peter, and John Poulter. 2007. *Learning For Action: A Short Definitive Account of Soft Systems Methodology, and Its Use for Practitioners, Teachers and Students*. Hoboken, NJ: Wiley.

Checkland, Peter, and Jim Scholes. 1999. *Soft Systems Methodology in Action*. New Sub edition. Chichester, Eng. ; New York: Wiley.

CLIR. 2000. *Authenticity in a Digital Environment.* Washington (D.C.): Council on Library and Information Resources. http://www.clir.org/pubs/reports/pub92/pub92.pdf (accessed 10 May, 2013).

Cofta, Piotr. 2007a. "Confidence, Trust and Identity." *BT Technology Journal* 25 (2): 173–78.

———. 2007b. *Trust, Complexity and Control: Confidence in a Convergent World*. 1st ed. Wiley.

Cofta, Piotr. 2013. *The Foundations of a Trustworthy Web*. Boston, Delft: Now Publishers.

Committee on Electronic Records. 1997. *Guide for Managing Electronic Records from an Archival Perspective*. Paris, France: International Council on Archives.

Cook, Michael. 1986. *The Management of Information from Archives*. Aldershot, Hants, England ; Brookfield, Vt., U.S.A: Gower.

Cook, Terry. 1997. "What's Past Is Prologue: A History of Archival Ideas since 1898 and the Future Paradigm Shift." *Archivaria* 43: 17–63.

———. 2001. "Archival Science and Postmodernism: New Formulations for Old Concepts." *Archival Science* 1: 3–24.

———. 2012. "Evidence, Memory, Identity, and Community: Four Shifting Archival Paradigms." *Archival Science* 13: 95-120.

Creswell, John W. 2009. *Research Design : Qualitative, Quantitative, and Mixed Method Approaches*. Thousand Oaks: Sage Publications.

Creswell, John W. 2011. *Designing and Conducting Mixed Methods Research*. 2nd ed. Los Angeles: SAGE Publications.

Crotty, Michael. 1998. *The Foundations of Social Research: Meaning and Perspective in the Research Process*. London ; Thousand Oaks, Calif: Sage Publications.

Crusco, Peter. 2014. "Authenticating Digital Evidence." *New York Law Journal*. February 25. http://www.newyorklawjournal.com/id=1202644221423/Authenticating-Digital-Evidence (accessed 14 October, 2014).

Daughtrey, William H., Jr. 2000. "Adapting Contract Law to Accommodate Electronic Contracts: Overview and Suggestions." *Rutgers Computer & Technology Law* 26 (2): 215–76. http://go.galegroup.com/ (accessed 6 June, 2013).

Davies, Alysia. 2008. "The Development of Laws on Electronic Documents and E-Commerce Transactions (PRB 00-12E)." *Library of Parliament - Parliamentary Information and Research Service*. December 20. http://www2.parl.gc.ca/Content/LOP/ResearchPublications/prb0012-e.htm (accessed 10 June, 2013).

Diamond, Elizabeth. 1994. "The Archivist as Forensic Scientist – Seeing Ourselves in a Different Way." *Archivaria* 38 (Fall): 139–54.

Dollar, Charles. 1978. "Appraising Machine-Readable Records." *The American Archivist* 41 (4): 423–30.

———. 1993. "Archivists and Records Managers in the Information Age." *Archivaria* 36 (Autumn): 37–52.

Doueihi, Milad. 2011. *Digital Cultures*. Cambridge, Massachusetts: Harvard University Press.

Duff, Wendy M. 1996. “Ensuring the Preservation of Reliable Evidence: A Research Project Funded by the NHPRC.” *Archivaria*, no. 42 (Fall): 28–45.

Duncan, Chris. 2009. “Authenticity or Bust.” *Archivaria* 68 (Fall): 97–118.

Duranti, Luciana. 1989a. “Diplomatics: New Uses for an Old Science.” *Archivaria* 28 (Summer): 7–27.

———. 1989b. “Diplomatics: New Uses for an Old Science (Part II).” *Archivaria* 29 (Winter): 4–17.

———. 1990a. “Diplomatics: New Uses for an Old Science (Part III).” *Archivaria* 30 (Summer): 4–20.

———. 1990b. “Diplomatics: New Uses for an Old Science (Part IV).” *Archivaria* 31 (Winter): 10–25.

———. 1991a. “Diplomatics: New Uses for an Old Science (Part V).” *Archivaria* 32 (Summer): 6–24.

———. 1991b. “Diplomatics: New Uses for an Old Science (Part VI).” *Archivaria* 33 (Winter): 6–24.

———. 1993. “The Archival Body of Knowledge: Archival Theory, Method, and Practice and Graduate and Continuing Education.” *The Journal of Education for Library and Information Science* 34 (Winter): 8–24.

———. 1996a. “Archival Science.” In *Encyclopedia of Library and Information Science*, 59:1–19. New York, Basel, Hong Kong: Marcel Dekker.

———. 1996b. “The Thinking on Appraisal of Electronic Records: Its Evolution, Focuses, and Future Directions.” *Archivi and Computer* 6: 493–518.

———. 1997. “The Archival Bond.” *Archives and Museum Informatics* 11 (3-4): 213–18.

———. 1998a. *Diplomatics: New Uses for an Old Science*. Lanham: Scarecrow Press.

———. 1998b. “The Odyssey of Records Management, Part 1.” *Records Management Quarterly* 23 (3).

———. 1998c. “The Odyssey of Records Management, Part 2.” *Records Management Quarterly* 23 (3).

———. 2001. “The Impact of Digital Technology on Archival Science.” *Archival Science* 1 (1): 39–55.

———. 2002. “Authenticity and Appraisal: Appraisal Theory Confronted With Electronic Records.” In *Proceedings of the 3rd International Colloquium on Library and Information Science: “The Refined Art of Destruction: Records’ Appraisal and Disposal.”* Salamanca, Spain: University of Salamanca. www.interpares.org/display_file.cfm?doc=ip1_dissemination_cpr_duranti_clis_2002.pdf (accessed 28 November, 2014).

———. 2005a. *The Long-Term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*. San Miniato: Archilab.

———. 2005b. “The Long-Term Preservation of Accurate and Authentic Digital Data: The InterPARES Project.” *Data Science Journal* 4 (October): 106–18.

———. 2007. “Reflections on InterPARES - The InterPARES 2 Project (2002-2007): An Overview.” *Archivaria* 64 (Fall): 113–21.

———. 2009. “From Digital Diplomatics to Digital Records Forensics.” *Archivaria* 68 (Fall): 39–66.

Duranti, Luciana, and Terry Eastwood. 1995. “Protecting Electronic Evidence: A Progress Report on a Research Study and Its Methodology.” *Archivi and Computer*, no. 3: 213–50.

Duranti, Luciana, Terry Eastwood, and Heather MacNeil. 2003. *Preservation of the Integrity of Electronic Records*. Springer Science & Business Media.

Duranti, Luciana, and Barbara Endicott-Popovsky. 2010. “Digital Records Forensics: A New Science and Academic Program for Forensic Readiness.” *Journal of Digital Forensics, Security and Law* 5 (2): 1–12.

Duranti, Luciana, and Heather MacNeil. 1997. “The Preservation of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project.” *Archivaria* 42 (Spring): 46–67.

Duranti, Luciana, Heather MacNeil, and William Underwood. 1996. "Protecting Electronic Evidence: A Second Progress Report on a Research Study and Its Methodology." *Archivi and Computer* VI (1): 37–70.

Duranti, Luciana, and Giovanni Michetti. 2015. "The Archival Method: Rediscovering a Research Tradition." In *Research in the Archival Multiverse*, edited by Anne Gilliland, Sue McKemmish, and Andrew Lau. Melbourne: Monash Publishing.

Duranti, Luciana, and Randy Preston. 2008. *Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential. Interactive and Dynamic Records*. Padova: Associazione Nazionale Archivistica Italiana.

Duranti, Luciana, and Corinne Rogers. 2012. "Trust in Digital Records: An Increasingly Cloudy Legal Area." *Computer Law & Security Review* 28 (5): 522–31.

Duranti, Luciana, and Kenneth Thibodeau. 2006. "The Concept of Record in Interactive, Experiential and Dynamic Environments: The View of InterPARES." *Archival Science* 6 (1): 13–68.

Eastwood, Terry. 1994. "What Is Archival Theory and Why Is It Important?" *Archivaria* 37 (Spring): 122–30.

———. 2010. "A Contested Realm: The Nature of Archives and the Orientation of Archival Science." In *Currents of Archival Thinking*, edited by Terry Eastwood and Heather MacNeil, 3–22. Santa Barbara, Calif: Libraries Unlimited.

EMC. 2013. "Leaders Edge: Highlights from CIO Summit 2013. Atlanta, GA." In . Atlanta, GA: EMC. http://www.emc.com/microsites/cio/articles/cio-summit-2013/cio-summit-2013-atlanta.pdf (accessed 2 October, 2014).

Erlandsson, Alf. 1997. *Electronic Records Management: A Literature Review*. Paris: International Council on Archives.

Facciolo, Hon. John M. 2010. "Explosions and Eruptions: Some Thoughts on a Conceptual Approach to Law School and Post-Graduate Education." *EDDE Journal* 1 (1). http://www2.americanbar.org/sections/scitech/ST203001/PublicDocuments/EDDE%20Journal%20-%20volume%201%20issue%201.pdf (accessed 6 September, 2012).

Factor, Michael, Ealan Henis, Dalit Naor, Simona Rabinovici-Cohen, Petra Reshef, Shahar Ronen, Giovanni Michetti, and Maria Guercio. 2009. "Authenticity and Provenance in Long Term Digital Preservation: Modeling and Implementation in Preservation Aware Storage." http://www.research.ibm.com/haifa/projects/storage/datastores/papers/Auth_Prov_CamReady_sent.pdf 10 October, 2014).

"*Federal Rules of Evidence*." 2014. Accessed December 27. http://www.law.cornell.edu/rules/fre 18 January, 2015).

Feilzer, Yvonne. 2010. "Doing Mixed Methods Research Pragmatically: Implications for the Rediscovery of Pragmatism as a Research Paradigm." *Journal of Mixed Methods Research* 4 (1): 6–16.

Feldman, Martha S., and Wanda J. Orlikowski. 2011. "Theorizing Practice and Practicing Theory." *Organization Science* 22 (5): 1240–53. doi:10.1287/orsc.1100.0612.

Ferguson-Boucher, K. A., and Barbara Endicott-Popovsky. 2008. "Digital Forensics and Records Management: What We Can Learn from the Discipline of Archiving," 1–6. In conference proceedings *Where Information Technology, Law and Risk Management Converge*, University of Washington, Seattle. http://www.e-evidence.info/1008.ht (accessed 17 October, 2013).

Findlay, Cassie. 2013. "People, Records and Power: What Archives Can Learn from WikiLeaks." *Archives and Manuscripts* 41 (1): 7–22.

Fisher, Paul. 2004. "Electronic Records as Evidence: The Case for Canada's New Standard." *Information Management Journal*, no. Mar/Apr.: 39-45. http://www.arma.org/bookstore/files/fisher_0304.pdf (accessed 13 April 2015).

Foscarini, Fiorella. 2009. “Function-Based Records Classification Systems: An Exploratory Study of Records Management Practices in Central Banks.” University of British Columbia.

———. 2010. “Understanding the Context of Records Creation and Use: ‘Hard’ versus ‘Soft’ Approaches to Records Management.” *Archival Science* 10 (4): 389–407.

Fosmire, M. Sean. 2006. “Refining the Standard: Authenticating Computer-Based Evidence.” *Law and Technology Resources for Legal Professionals*. November 3. http://www.llrx.com/features/computerbasedevidence.htm (accessed 7 July, 2013).

Garfinkel, Simson L. 2009. “Providing Cryptographic Security and Evidentiary Chain-of-Custody with the Advanced Forensic Format, Library, and Tools.” *International Journal of Digital Crime and Forensics* 1 (1): 1–28. http://simson.net/clips/academic/2009.IJDCF.AFFLIB.pdf (accessed 10 September, 2012).

Giaretta, David. 2011. *Advanced Digital Preservation*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.

Giaretta, David, Brian Matthews, Bicarregui, Lambert, and Maria Guercio. 2009. “Significant Properties, Authenticity, Provenance, Representation Information and OAIS.” In *iPRES 2009: The Sixth International Conference on Preservation of Digital Objects*, 67–73. San Francisco, CA: California Digital Library. http://escholarship.org/uc/cdl_ipres09 (accessed 14 September, 2011).

Gilliland, Anne. 2008. “Investigating the Roles and Requirements, Manifestations and Management of Metadata in the Creation of Reliable and Preservation of Authentic Digital Entities - Description Cross-Domain Task Force Report.” In *International Research on Permanent Authentic Records in Electronic Systems, IP2*, edited by Luciana Duranti and Randy Preston, 305–59.

———. 2014. *Conceptualizing 21st-Century Archives*. Chicago: Society of American Archivists.

Gilliland, Anne, and Sue McKemmish. 2012. “Recordkeeping Metadata, the Archival Multivers, and Societal Grand Challenges.” In *Proceedings of the International Conference on Dublin Core and Metadata Applications 2012*, 106–13. Kuching, Sarawak, Malaysia. http://dcevents.dublincore.org/IntConf/dc-2012/paper/view/108/66 (accessed 25 September, 2014).

Golden, Bernard. 2014. “As IDC Sees It, Tech’s ‘Third Platform’ Disrupts Everyone.” *CIO*. March 27. http://www.cio.com/article/2377568/cloud-computing/as-idc-sees-it-tech-s-third-platform-disrupts-everyone.html (accessed 14 October, 2014).

Goldkuhl, Gören. 2004. *Meanings of Pragmatism: Ways to Conduct Information Systems Research*. Sweden: VITS Research Network, Linköping University. http://www.vits.org/konferenser/alois2004/html/6901.pdf (accessed 4 January, 2015).

Goldkuhl, Gören, and Jenny Lagsten. 2012. “The Many Prepositions of Practice Research: About, For, in with and from.” In *2nd International Conference on Practice Research, Helsinki, May 30-31 2012*. http://blogs.helsinki.fi/practice-research-conference-2012/files/2012/06/GGJL-PracticeResearch2012.pdf (accessed 13 April, 2015).

Greene, Jennifer C., Valerie J. Caracelli, and Wendy F. Graham. 1989. “Toward a Conceptual Framework for Mixed-Method Evaluation Designs.” *Educational Evaluation and Policy Analysis* 11 (3): 255–74.

Grimm, Hon. Paul W., Michael V. Ziccardi, and Alexander W. Major. 2009. “Back to the Future: Lorraine v. Markel American Insurance Co. and New Findings on the Admissibility of Electronically Stored Information.” *Akron Law Review* 42: 361–418.

Guba, Egon G., ed. 1990. *The Paradigm Dialog*. Newbury Park, Calif: SAGE Publications, Inc.

Guercio, Maria. 2005. “Records and Archival Education as Fundamental Requirement for Digital Society Citizenship: Challenges to Face, Risks to Govern, Knowledge to Re-Build and Implement.” DLM Forum, Budapest 5-7 October 2005.

http://ec.europa.eu/transparency/archival_policy/dlm_forum/doc/27_guercio_07-10-05am.pdf (accessed 13 April, 2015)

———. 2008. “Authenticity and OAIS: The CASPAR Model and the InterPARES Principles & Outputs.” presented at the Delos Summer School, Tirrenia, June 11. http://www.interpares.org/display_file.cfm?doc=ip1-2_dissemination_ws_guercio_delos-ss_tirrenia_2008.pdf (accessed 13 April, 2015).

Guercio, Maria, and Giovanni Michetti. 2009a. “Modeling Authenticity, Part 1.” January. http://www.alliancepermanentaccess.org/index.php/training/training-materials/lecture-3-modelling-authenticity-in-caspar/ (accessed 15 September, 2014).

———. 2009b. “Modeling Authenticity-Part 2.” September. http://www.alliancepermanentaccess.org/index.php/training/training-materials/lecture-3-modelling-authenticity-in-caspar/ (accessed 15 September, 2014).

Guercio, Maria, and Silvio Salza. 2013. “Managing Authenticity through the Digital Resource Lifecycle.” In *Digital Libraries and Archives*, edited by Maristella Agosti, Floriana Esposito, Stefano Ferilli, and Nicola Ferro, 249–60. Communications in Computer and Information Science 354. Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-642-35834-0_25 (accessed 21 September, 2014).

Gurría, Angel. 2009. “Responding to the Global Economic Crisis - OECD’s Role in Promoting Open Markets and Job Creation.” http://www.oecd.org/fr/echanges/respondingtotheglobaleconomiccrisisoecdsroleinpromotingopenmarketsandjobcreation.htm (accessed 22 September, 2014).

Hackett, Yvette, William Underwood, and Philip Eppard. n.d. “Part One - Case and General Studies in the Artistic, Scientific, and Governmental Sectors: Focus Task Force Report Yvette Hackett, Librar Y and Archives Canada William Underwood, Georgia Tech Research Institute Philip Eppard, University of Albany , State University of New York.” In *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records*, edited by Duranti and Randy Preston.

Hasan, Ragib, Radu Sion, and Marianne Winslett. 2007. “Introducing Secure Provenance.” ACM Press. http://digitalpiglet.org/research/sion2007storagess-provenance.pdf (accessed 13 April, 2015).

Hjørland, Birger. 2008. “The Epistemological Lifeboat.” http://www.iva.dk/jni/lifeboat/info.asp?subjectid=92 (accessed 11 January, 2015).

Hookway, Christopher. 2015. “Pragmatism.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2015, forthcoming. http://plato.stanford.edu/archives/spr2015/entries/pragmatism/ (accessed 12 January, 2015).

Hurley, Chris. 1995. “Ambient Functions: Abandoned Children to Zoos.” *Archivaria* 40 (Fall).

ICA. 2015. “Multilingual Archival Terminology.” *ICA: International Council on Archives*. http://icarchives.webbler.co.uk/14282/multilingual-archival-terminology/multilingual-archival-terminology.html (accessed January 31, 2015).

ICA Committee on Archival Legal Matters. 2002. *Authenticity of Electronic Records: A Report Prepared for UNESCO*. Study 13-1. Paris, France: International Council on Archives.

Imwinkelried, Edward J. 2005. *Evidentiary Foundations*. 6th ed. Lexis/Nexis.

International Standards Organization. 2001. *ISO-15489 Information and Documentation-Records Management, Part 1, 2*. Text. Geneva: ISO. http://www.iso.org/iso/catalogue_detail?csnumber=31908 (accessed 4 January, 2015).

InterPARES. 2012. “InterPARES 3 Project: Glossary.” http://www.interpares.org/ip3/ip3_terminology_db.cfm?letter=p&term=38 (accessed 21 December, 2014).

InterPARES Trust. 2015. “InterPARES Trust.”. www.interparestrust.org (accessed January 31, 2015).

Irons, Alastair. 2006. "Computer Forensics and Records Management – Compatible Disciplines." *Records Management Journal* 16 (2): 102–12.

Isaza, John. 2010. *Metadata in Court: What RIM, Legal and IT Need to Know*. Pittsburgh, PA: ARMA International Education Foundation. http://www.armaedfoundation.org/pdfs/Isaza_Metadata_Final.pdf (accessed 13 April, 2015)

ISO. 2001. *ISO-15489 (2001) Information and Documentation-Records Management*. ISO-15489 (2001). http://www.iso.org/iso/iso_catalogue.htm (accessed 9 January, 2015).

James, William. 1907. *Pragmatism: A New Name for Some Old Ways of Thinking*. Project Gutenberg EBook, 2004. New York: Longman, Green and Co. http://www.gutenberg.org/files/5116/5116-h/5116-h.htm (accessed 6 January, 2015).

Jenkinson, Hilary. 1937. *A Manual of Archive Administration*. New and Revised. London: Percy Lund, Humphries & Co. http://www.archive.org/details/manualofarchivea00iljenk (accessed 22 November, 2014).

John, Jeremy Leighton. 2012. *Digital Forensics and Preservation*. Digital Preservation Coalition. http://www.dpconline.org/component/docman/doc_download/810-dpctw12-03pdf (accessed 9 September, 2014).

John, Jeremy Leighton, I Rowlands, P Williams, and K Dean. 2010. *Digital Lives: Personal Digital Archives for the 21st Century - An Initial Synthesis*. Digital Lives Research Paper. http://britishlibrary.typepad.co.uk/files/digital-lives-synthesis02-1.pdf (accessed 9 September, 2014).

Johnson, R. Burke, Anthony J. Onwuegbuzie, and Lisa A. Turner. 2007. "Toward a Definition of Mixed Methods Research." *Journal of Mixed Methods Research* 1 (2): 112–33.

Kirschenbaum, Matthew G., Richard Ovenden, and Gabriela Redwine. 2010. *Digital Forensics in Born Digital Cultural Heritage Collections*. Washington, D.C.: Council on Library and Information resources.

Kolko, Jon. 2012. "Wicked Problems: Problems Worth Solving (SSIR)." March 6. http://www.ssireview.org/articles/entry/wicked_problems_problems_worth_solving (accessed 13 January, 2015).

Krementz, Steven. 2009. "Record Authenticity." *CA on Security Management*. October 12. community.ca.com/blogs/iam/archive/2009/10/12/record-authenticity.aspx (accessed 10 January, 2014).

Kuhn, Thomas S. 1970. "The Structure of Scientific Revolutions." *International Encyclopedia of Unified Science*. Foundations of Unified Science 2. Chicago: University of Chicago Press. http://projektintegracija.pravo.hr/_download/repository/Kuhn_Structure_of_Scientific_Revolutions.pdf (accessed 22 December, 2014).

Lamb, David. 2009. "CASPAR." *Digital Curation Centre*. http://www.dcc.ac.uk/resources/briefing-papers/technology-watch-papers/caspar (accessed 7 July, 2013).

Latour, Bruno. 2005. *Reassembling the Social an Introduction to Actor-Network-Theory*. Oxford; New York: Oxford University Press. http://site.ebrary.com/id/10233636 (accessed 14 January, 2015).

Lauriault, Tracey P, Barbara L Craig, D. R. Fraser Taylor, and Peter L Pulsifer. 2007. "Today's Data Are Part of Tomorrow's Research: Archival Issues in the Sciences." *Archivaria* 64 (Fall): 123–80.

Lavoie, Brian, and Lorcan Dempsey. 2004. "Thirteen Ways of Looking at...Digital Preservation." *D-Lib Magazine* 10 (7/8). http://www.dlib.org/dlib/july04/lavoie/07lavoie.html (accessed 13 April, 2015).

Law, John. 1992. "Notes on the Theory of the Actor-Network: Ordering, Strategy and Heterogeneity." *Systems Practice* 5: 379–92.

http://www.heterogeneities.net/publications/Law1992NotesOnTheTheoryOfTheActorNetwork.pdf (accessed 23 December, 2014).

———. 2006. “Making a Mess with Method (version of 19th January 2006).” http://www.heterogeneities.net/publications/Law2006Makinga MesswithMethod.pdf (accessed 23 December, 2014).

Lemieux, V. 2001. “Let the Ghosts Speak: An Empirical Exploration of the Nature of the Record.” *Archivaria* 51: 81–111.

———. 2014. “Toward a ‘Third Order’ Archival Interface: Research Notes on Some Theoretical and Practical Implications of Visual Explorations in the Canadian Context of Financial Electronic Records.” *Archivaria* 78 (0).

Lemieux, V., and Jason R. Baron. 2011. “Overcoming the Digital Tsunami in E-Discovery: Is Visual Analysis the Answer?” *Canadian Journal of Law and Technology* 9 (33): 1–15.

Lemieux, V., and L. Limonad. 2011. “What ‘Good’ Looks like: Understanding Records Ontologically in the Context of the Global Financial Crisis.” *Journal of Information Science* 37 (1): 29–39.

Lu, Rongxing, Xiaodong Lin, Xiaohui Liang, and Xuemin Shen. 2010. “Secure Data Provenance: The Essential of Bread and Butter of Data Forensics in Cloud Computing.” In *Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security*. Beijing, China. http://dl.acm.org/citation.cfm?id=1755688 (accessed 24 October, 2014).

Lynch, Clifford. 2000. “Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust.” In *Authenticity in a Digital Environment*. Washington, D.C.: Council on Library and Information Resources. http://www.clir.org/pubs/reports/pub92/lynch.html (accessed 2 January, 2015).

MacNeil, Heather. 2000a. “Providing Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Electronic Records.” *Archivaria*, no. 50 (Fall): 52–78.

———. 2000b. *Trusting Records: Legal, Historical, and Diplomatic Perspectives*. Dordrecht: Kluwer Academic.

———. 2001. “Trusting Records in a Postmodern World.” *Archivaria* 51 (Spring): 36–47.

———. 2004. “Contemporary Archival Diplomatics as a Method of Inquiry: Lessons Learned from Two Research Projects.” *Archival Science* 4 (3-4): 199–232.

MacNeil, Heather, and Anne Gilliland-Swetland. 2005. “Authenticity Task Force Report.” In *The Long-Term Preservation of Authentic Electronic Records: Finding of the InterPARES Project*, edited by Luciana Duranti. San Miniato, Italy: Archilab.

MacNeil, Heather, and Bonnie Mak. 2007. “Constructions of Authenticity.” *Library Trends* 56 (1): 26–52.

Mak, Bonnie. 2012. “On the Uses of Authenticity.” *Archivaria* 73 (Spring): 1-17.

Mason, Stephen. 2008. “Rethinking Concepts in Virtual Evidence.” *The Icfai Journal of Cyber Law* VII (1): 48–54.

———. , ed. 2010. *Electronic Evidence*. 2nd Edition. London, UK: Lexis Nexis.

———. 2012. *Electronic Evidence*. Third Edition. Croydon, UK: LexisNexis.

McKemmish, Sue. 2001. “Placing Records Continuum Theory and Practice.” *Archival Science* 1 (4): 333–59.

Meijer, Albert Jacob. 2003. “Trust This Document! ICTs, Authentic Records and Accountability.” *Archival Science* 3 (3): 275–90.

Merriam-Webster. 2015. “Term of Art.” http://www.merriam-webster.com/dictionary/term%20of%20art (accessed 7 January, 2015).

Millar, Laura. 2004. *Authenticity of Electronic Records: A Report Prepared for UNESCO and the International Council on Archives*. Study 13-2. Paris, France: International Council on Archives.

Mishler, Elliot G. 1986. *Research Interviewing: Context and Narrative*. Harvard University Press.

Münscher, Robert, and Torsten Kühlmann. 2012. “Using Critical Incident Technique in Trust Research.” In *Handbook of Research Methods on Trust*, edited by Fergus Lyon, Guido Möllering, and Mark N.K. Saunders, 161–74. Cheltenham, UK; Northampton, MA, USA: Edward Elgar Publishing.

NARA. 2015. “The History of the Electronic Records and ERA.” *National Archives: Electronic Records Archives*. http://www.archives.gov/era/about/history.html (accessed January 18, 2015).

Naugler, Harold. 1978. “Focus: The Machine Readable Archives Divison of the Public Archives of Canada.” *Archivaria* 6 (Summer): 176–80.

———. 1984. *The Archival Appraisal of Machine-Readable Records: A RAMP Study With Guidelines*. PGI-84/WS/27. Paris: UNESCO. http://unesdoc.unesco.org/images/0006/000635/063501eo.pdf accessed 15 January, 2015).

Nesmith, Tom. 2002. “Seeing Archives: Postmodernism and the Changing Intellectual Place of Archives.” *American Archivist* 65 (1): 24–41.

Oxford English Dictionary. 2014. “Authentic, Adj. and N.” *OED Online*. Oxford University Press. http://www.oed.com.ezproxy.library.ubc.ca/view/Entry/13314 (accessed 15 December, 2014).

Palmer, G. 2001. *A Road Map for Digital Forensic Research*. DFRWS Technical Report. http://www.dfrws.org/2001/dfrws-rm-final.pdf (accessed 18 November, 2012).

Park, Eun. 2001. “Understanding ‘Authenticity’ in Records and Information Management: Analyzing Practitioner Constructs.” *American Archivist* 64 (2): 270–91.

———. 2002a. *Developing a Framework for Authenticity Requirements in Student Records Systems: An Exploratory Study*. Dissertation. University of California, Los Angeles.

———. 2002b. “Morphological and Semantic Analysis of Language Uses and Concepts of Authenticity in Electronic Records Systems.” http://www.cais-acsi.ca/proceedings/2002/Park_2002.pdf (accessed 19 January, 2015).

Park, Eun G. 2001. “Understanding ‘Authenticity’ in Records and Information Management: Analyzing Practitioner Constructs.” *The American Archivist* 64 (2): 270–91.

Patton, Michael Quinn. 2014. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. SAGE Publications.

Patzakis, John. 2012. “Facebook Evidence Disallowed by Court Due to Lack of ‘Identifying Characteristics.’” *Next Gen eDiscovery Law & Tech Blog*. www.blog.x1discovery.com (accessed 16 September, 2014) .

Paul, George L. 2004. “The ‘Authenticity Crisis’ in Real Evidence.” *The Practical Litigator* 15 (6): 45–52.

———. 2008. *Foundations of Digital Evidence*. Chicago, IL: American Bar Association.

Paul, George L, and Jason R. Baron. 2007. “Information Inflation: Can the Legal System Adapt?” *Richmond Journal of Law & Technology* XIII (3): 1–41. http://law.richmond.edu/jolt/v13i3/article10.pdf (accessed 14 October, 2013).

Pearce-Moses, Richard. 2005. “A Glossary of Archival and Records Terminology.” *Society of American Archivists*. http://www.archivists.org/glossary/ (accessed 23 December, 2014).

Peter Checkland on the Origins of SSM. 2012. https://www.youtube.com/watch?v=XA2i1n-o9L0&feature=youtube_gdata_player (accessed 22 December, 2014).

Pickard, Alison Jane. 2013. *Research Methods in Information*. London: Facet.

Rittel, Horst, and Melvin Webber. 1973. “Dilemmas in a General Theory of Planning.” *Policy Sciences* 4: 155–69.

http://www.uctc.net/mwebber/Rittel+Webber+Dilemmas+General_Theory_of_Planning.pdf (accessed 14 January, 2015).

Roeder, John, Philip Eppard, William Underwood, and Tracey P Lauriault. 2008. "Authenticity, Reliability and Accuracy of Digital Records in the Artistic, Scientific and Governmental Sectors: Domain 2 Task Force Report." In *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records*, edited by Luciana Duranti and Randy Preston. Padova: Associazione Nazionale Archivistica Italiana.

Rogers, Corinne, and JL John. 2013. "Shared Perspectives, Common Challenges: A History of Digital Forensics & Ancestral Computing for Digital Heritage." In *The Memory of the World in the Digital Age: Digitization and Preservation*, 314–36. Vancouver, BC: UNESCO. http://www.unesco.org/webworld/download/mow/mow_vancouver_proceedings_en.pdf (accessed 15 November, 2014).

Rowlingson, Robert. 2004. "A Ten Step Process for Forensic Readiness." *International Journal of Digital Evidence* 2 (3): 1–28. www.ijde.org (accessed 4 October, 2012).

SAA. 2011. "SAA Core Values Statement and Code of Ethics." *Society of American Archivists*. http://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics (accessed 3 January, 2015).

Salza, Silvio, Maria Guercio, Monica Grossi, Stefan Pröll, Christos Stroumboulis, Yannis Tzatzikas, Martin Doerr, and Giorgios Flouris. 2012. *Report on Authenticity and Plan for Interoperable Authenticity Evaluation System*. http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2012/04/APARSEN-REP-D24_1-01-2_3.pdf (accessed 5 January 2015).

Scanlan, Daniel M. 2011. *Digital Evidence in Criminal Law*. Aurora, Ont.: Canada Law Book.

Schatzki, Theodore R, K Knorr-Cetina, and Eike von Savigny, eds. 2001. *The Practice Turn in Contemporary Theory*. London; New York: Routledge.

Shankar, Kalpana. 2004. "Recordkeeping in the Production of Scientific Knowledge: An Ethnographic Study." *Archival Science* 4 (3-4): 367–82.

Sheppard, A. F., and Luciana Duranti. 2010. *The Canadian Legal Framework for Evidence and the Digital Economy: A Disjunction?*. SSHRC Knowledge Synthesis and the Digital Economy. University of British Columbia.

Statistics Canada, and Standards Division. 2012. *North American Industry Classification System (NAICS) Canada*. Ottawa, ON: Statistics Canada. http://www.census.gov/eos/www/naics/ (accessed 14 September, 2014).

Strodl, Stephan, Petar Petrov, and Andreas Rauber. 2011. *Research on Digital Preservation Within Projects Co-Funded by the European Union in the ICT Programme*. SCAPE Project. http://www.scape-project.eu/wp-content/uploads/2014/08/SCAPE_digpres_research_ict.pdf (accessed 8 January, 2015).

Strutin, Ken. 2011. "Social Media and the Vanishing Points of Ethical and Constitutional Boundaries." *Pace Law Review* 31 (1): 228–90. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1789881 (accessed 18 October, 2014).

Sunlight Foundation. 2014a. "Open Data Policies at Work: A Bird's Eye View of Open Data Policies." *Sunlight Foundation*. http://sunlightfoundation.com/ (accessed 23 December, 2014).

———. 2014b. "Open Data Policy Guidelines." *Sunlight Foundation*. http://sunlightfoundation.com/opendataguidelines/ (accessed 23 December, 2014).

Sztompka, Piotr. 1999. *Trust a Sociological Theory*. Cambridge, UK; New York, NY: Cambridge University Press.

Takach, George. 2003. *Computer Law*. 2nd ed. Toronto, ON: Irwin Law Inc.

Tennis, Joseph T., and Corinne Rogers. 2012a. “Authenticity Metadata and the IPAM: Progress toward the InterPARES Application Profile.” In *Proceedings of the International Conference on Dublin Core and Metadata Applications*, 38–45. Kuching, Sarawak, Malaysia: DCMI. http://dcevents.dublincore.org/index.php/IntConf/dc-2012/schedConf/presentations (accessed 4 November, 2012).

———. 2012b. *General Study 15: Metadata Application Profiles for Authenticity*. University of British Columbia.

Thibodeau, Kenneth. 2013. “Wrestling with Shape-Shifters: Perspectives on Preserving Memory in the Digital Age.” In *The Memory of the World in the Digital Age: Digitization and Preservation*, 15–23. Vancouver, BC: UNESCO. http://www.unesco.org/webworld/download/mow/mow_vancouver_proceedings_en.pdf (accessed 15 December, 2014).

Thibodeau, Kenneth, and Daryll Prescott. 1996. “Reengineering Records Management: The U.S. Department of Defense, Records Management Task Force.” *Archivi and Computer* VI (1): 71–78.

Todd, Malcom. 2006. “Power, Identity, Integrity, Authenticity, and the Archives: A Comparative Study of the Application of Archival Methodologies to Contemporary Privacy.” *Archivaria* 61 (Spring): 181–233.

Tonkiss, Fran. 2009. “Trust, Confidence and Economic Crisis.” *Intereconomics* 44 (4): 196–202. http://www.ceps.eu/system/files/article/2009/09/196-202-Tonkiss.pdf (accessed 13 October, 2014).

Trace, C B. 2011. “Documenting Work and Working Documents: Perspectives from Workplace Studies, CSCW, and Genre Studies.” In *Proceedings of the 44th Hawaii International Conference on System Sciences*: 1–10. https://www.ischool.utexas.edu/~cbtrace/pubs/CBT_HICSS_2011.pdf (accessed 13 April, 2015).

Upward, Frank. 1996. “Structuring the Records Continuum Part One: Post-Custodial Principles and Properties.” *Archives and Manuscripts* 24 (2): 268–85.

———. 2005. “The Records Continuum.” In *Archives: Recordkeeping in Society*, 197–222. Topics in Australasian Library and Information Studies, No. 24. Wagga Wagga, N.S.W. : Centre for information studies: Charles Sturt University.

Wikipedia. 2014. “List of National Legal Systems.” *Wikipedia, the Free Encyclopedia*. http://en.wikipedia.org/w/index.php?title=List_of_national_legal_systems&oldid=613310870 (accessed 11 April, 2015).

Yakel, Elizabeth. 1997. “Recordkeeping in Radiology: The Relationships between Activities and Records in Radiological Processes.” University of Michigan.

Yeo, Geoffrey. 2007. “Concepts of Record (1): Evidence, Information, and Persistent Representations.” *American Archivist* 70 (2): 315–43.

———. 2008. “Concepts of Record (2): Prototypes and Boundary Objects.” *American Archivist* 71 (1): 118–43.

Appendix 1: Listservs Used for Survey Distribution

The survey was sent to the following listservs:

arcan-l@mailman.srv.ualberta.ca

Arcan-L is the national archival listserv of the Canadian archival community, encouraging discussion of archival issues and interests of particular relevance to Canadian Archives and archivists (http://www.mailman.srv.ualberta.ca/mailman/listinfo/arcan-l).

archives@forums.archivists.org

The Archives & Archivists (A&A) list is the forum of the Society of American Archivists in the United States. It is an open forum for all topics relating to archival theory and practice (http://www2.archivists.org/listservs/archives).

ica-l@mailman.srv.ualberta.ca

The International Council on Archives listserv provides a venue for sharing information on archival issues internationally, and a discussion forum for all aspects of archival theory and practice (http://www.ica.org/4715/ica-listserv/ica-listserv.html).

ARCHIVES-NRA@jiscmail.ac.uk

This is a discussion list for the UK education and research communities of archivists, conservators, and records managers (https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=ARCHIVES-NRA).

ERECS-L@listserv.albany.edu

ERECS-L is devoted to the management and preservation of electronic records (http://www.lsoft.com/scripts/wl.exe?SL1=ERECS-L&H=LISTSERV.ALBANY.EDU). (1,362 subscribers)

pasig-discuss@mail.asis.org

The Preservation and Archiving Special Interest Group (PASIG) is a vendor-independent community discussing and sharing open computing solutions and best practices in digital preservation (https://www.asis.org/Pasig/forum.html).

rda-all@lists.rd-alliance.org

This is a general list for the Research Data Alliance (http://lists.lists.rd-alliance.org/mailman/listinfo/rda-all). The mandate of the Research Data Alliance is to enable open sharing of data across technologies, disciplines, and countries (https://rd-alliance.org/about.html).

recmgmt-l@lists.ufl.edu

Records Management Program (http://lists.ufl.edu/cgi-bin/wa?INDEX). (2,103 subscribers, March 3, 2014)

Records-Management-UK@jiscmail.ac.uk

The Information and Records Management Society (IRMS) mailing list is used to share RM knowledge and expertise, in both the UK and internationally (https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=records-management-uk).

Shortly after the survey opened, an invitation was received to send the survey to the following listservs:

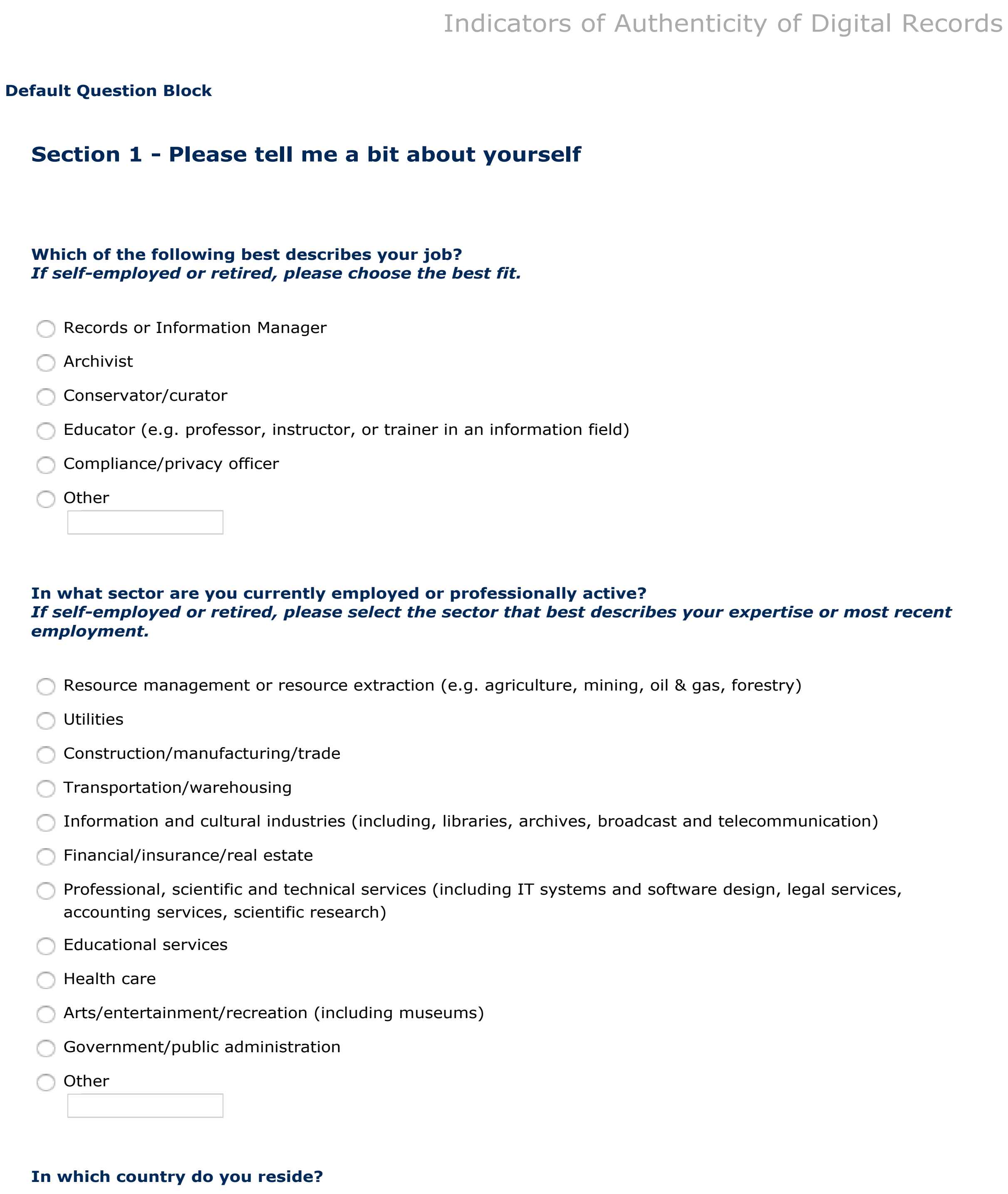
listserv@greynet.org

The GreyNet Literature Service exists to "facilitate dialog, research, and communication between persons and organisations in the field of grey literature." Its activities include the creation and maintenance of web-based resources, a moderated listserv, research, publication, open access, and education in the field of Grey Literature (www.greynet.org).

TEI-L@listserv.brown.edu

The Text Encoding Initiative (TEI) public discussion list (https://listserv.brown.edu/archives/cgi-bin/wa?INDEX=&p=16). (1104 subscribers)

Appendix 2: Survey Questionnaire



Indicators of Authenticity of Digital Records

Default Question Block

Section 1 - Please tell me a bit about yourself

Which of the following best describes your job?
If self-employed or retired, please choose the best fit.

○ Records or Information Manager

○ Archivist

○ Conservator/curator

○ Educator (e.g. professor, instructor, or trainer in an information field)

○ Compliance/privacy officer

○ Other

In what sector are you currently employed or professionally active?
If self-employed or retired, please select the sector that best describes your expertise or most recent employment.

○ Resource management or resource extraction (e.g. agriculture, mining, oil & gas, forestry)

○ Utilities

○ Construction/manufacturing/trade

○ Transportation/warehousing

○ Information and cultural industries (including, libraries, archives, broadcast and telecommunication)

○ Financial/insurance/real estate

○ Professional, scientific and technical services (including IT systems and software design, legal services, accounting services, scientific research)

○ Educational services

○ Health care

○ Arts/entertainment/recreation (including museums)

○ Government/public administration

○ Other

In which country do you reside?

What is your current age? (U.S. Census)

○ 20 to 24
○ 25 to 34
○ 35 to 44
○ 45 to 54
○ 55 to 64
○ 65 or over

What is the highest level of education you have completed?

○ High School / GED
○ Some College
○ 2-year College Degree
○ 4-year College Degree
○ Masters Degree
○ Doctoral Degree
○ Professional Degree (JD, MD)

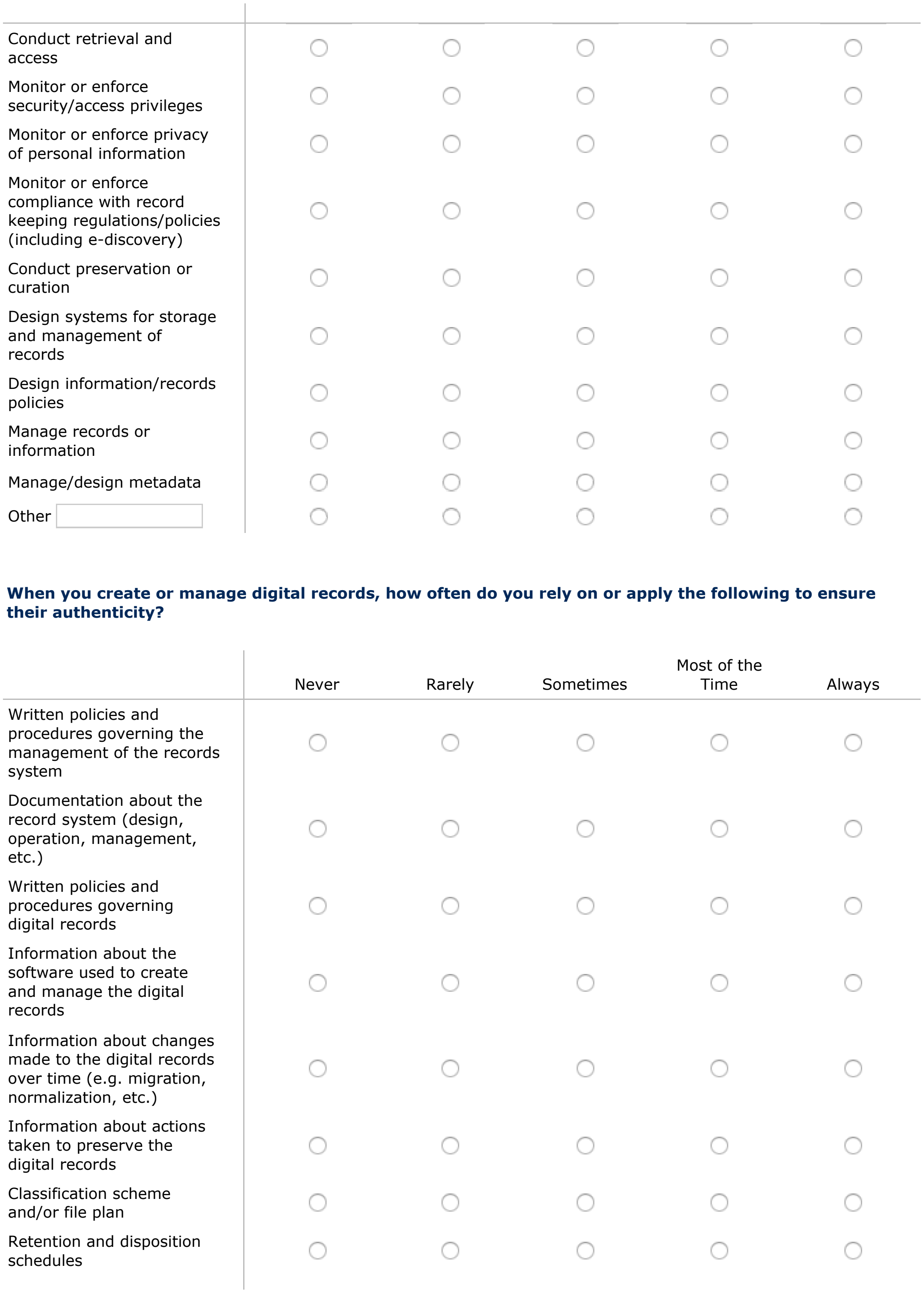
Do you have a degree in any of the following fields?
Please check all that apply.

□ Library and information science (LIS)
□ Archival science (or LIS with archival concentration)
□ Computer science
□ Law
□ History
□ Other

Section 2 - Please tell me a bit about your work

How often do you conduct the following tasks with respect to digital records (e.g. electronic documents, images, data, data sets, database records, electronically stored information [ESI], web pages, etc.)?
If self-employed or retired, please refer to the job or contract you feel is most relevant.

	Never	Rarely	Sometimes	Often	Very Often



Conduct retrieval and access	○	○	○	○	○
Monitor or enforce security/access privileges	○	○	○	○	○
Monitor or enforce privacy of personal information	○	○	○	○	○
Monitor or enforce compliance with record keeping regulations/policies (including e-discovery)	○	○	○	○	○
Conduct preservation or curation	○	○	○	○	○
Design systems for storage and management of records	○	○	○	○	○
Design information/records policies	○	○	○	○	○
Manage records or information	○	○	○	○	○
Manage/design metadata	○	○	○	○	○
Other □	○	○	○	○	○

When you create or manage digital records, how often do you rely on or apply the following to ensure their authenticity?

	Never	Rarely	Sometimes	Most of the Time	Always
Written policies and procedures governing the management of the records system	○	○	○	○	○
Documentation about the record system (design, operation, management, etc.)	○	○	○	○	○
Written policies and procedures governing digital records	○	○	○	○	○
Information about the software used to create and manage the digital records	○	○	○	○	○
Information about changes made to the digital records over time (e.g. migration, normalization, etc.)	○	○	○	○	○
Information about actions taken to preserve the digital records	○	○	○	○	○
Classification scheme and/or file plan	○	○	○	○	○
Retention and disposition schedules	○	○	○	○	○

Archival description	○	○	○	○	○
Access controls/security measures	○	○	○	○	○
Audit logs	○	○	○	○	○
Cryptographic validation techniques (e.g. digital signatures, hash digests, etc.)	○	○	○	○	○
Standardized metadata	○	○	○	○	○

How frequently do you use the following cryptographic validation techniques?

	Never	Occasionally	Very Often
Digital signatures	○	○	○
Trusted time stamps	○	○	○
Checksums	○	○	○
Hash digests	○	○	○
Secure transmission	○	○	○

What metadata do you routinely use or manage?
Please check all that apply.

□ A metadata schema or guideline (e.g. Dublin Core, PREMIS, MoReq, etc.) Please list:

□ A modification of a schema, customized for your organization. Please describe:

□ A custom-built metadata schema (designed without elements from existing schemas). Please describe:

□ Metadata generated by the software or record system in use only

□ Not sure

Have you ever been required to guarantee or attest to the authenticity of digital records in any of the following circumstances?

□ Providing testimony in court or administrative hearing

□ Pending litigation or administrative action (e-discovery process)

□ Authenticating copies of digital records for research or in response to reference requests

□ Other

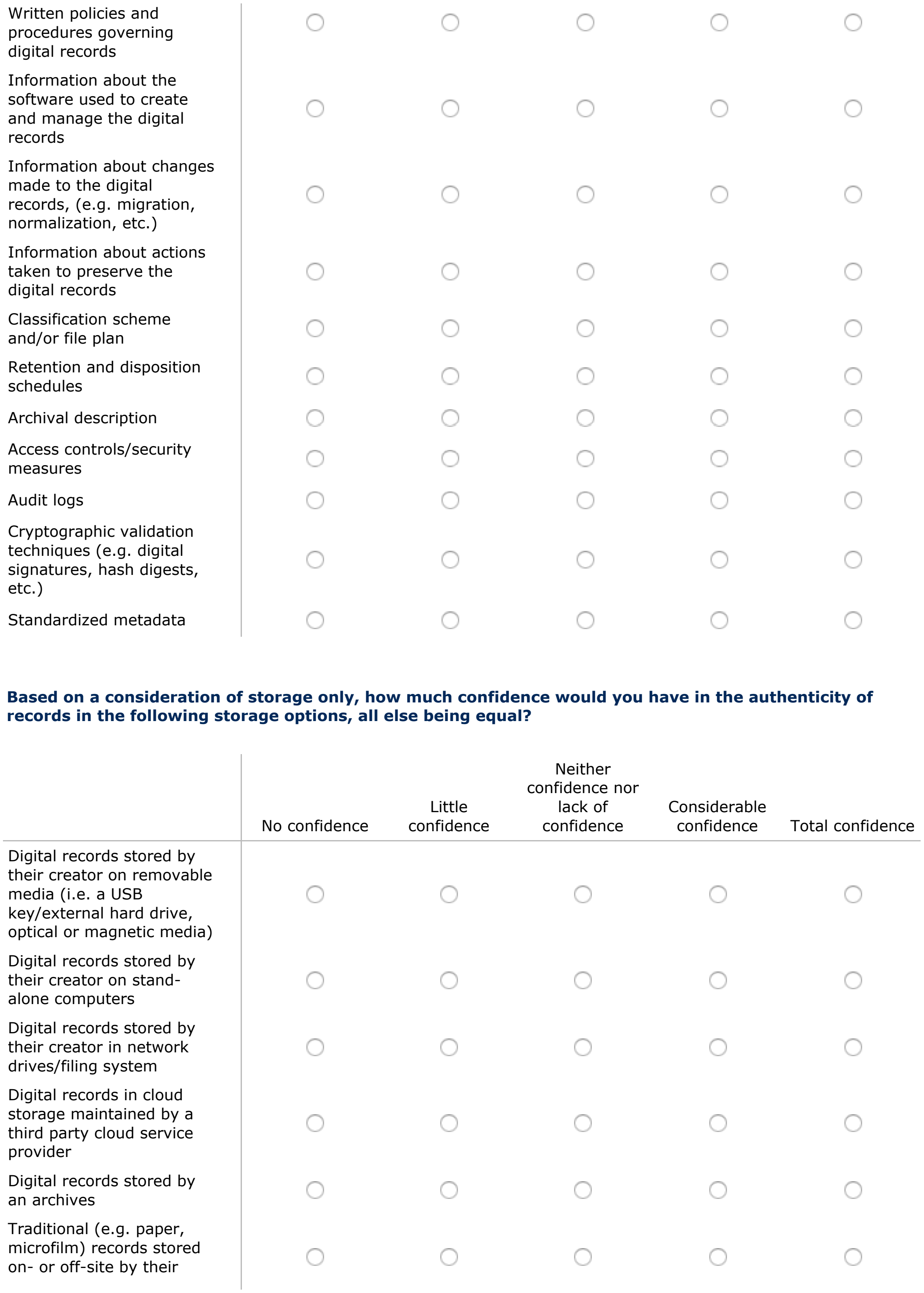
□ I have never been required to guarantee or attest to the authenticity of digital records

When you were required to guarantee or attest that digital records are authentic, how important were the following in making your assessment?

	Not at all Important	Very Unimportant	Neither Important nor Unimportant	Very Important	Extremely Important
Written policies and procedures governing the management of the records system	○	○	○	○	○
Documentation about the record system (design, operation, management, etc.)	○	○	○	○	○
Written policies and procedures governing digital records	○	○	○	○	○
Information about the software used to create and manage the digital records	○	○	○	○	○
Information about changes made to the digital records over time, (e.g. migration, normalization, etc.)	○	○	○	○	○
Information about actions taken to preserve the digital records	○	○	○	○	○
Classification scheme and/or file plan	○	○	○	○	○
Retention and disposition schedules	○	○	○	○	○
Archival description	○	○	○	○	○
Access controls/security measures	○	○	○	○	○
Audit logs	○	○	○	○	○
Cryptographic validation techniques (e.g. digital signatures, hash digests, etc.)	○	○	○	○	○
Standardized metadata	○	○	○	○	○

If you needed to assess that digital records are authentic, how important would the following be in making your assessment?

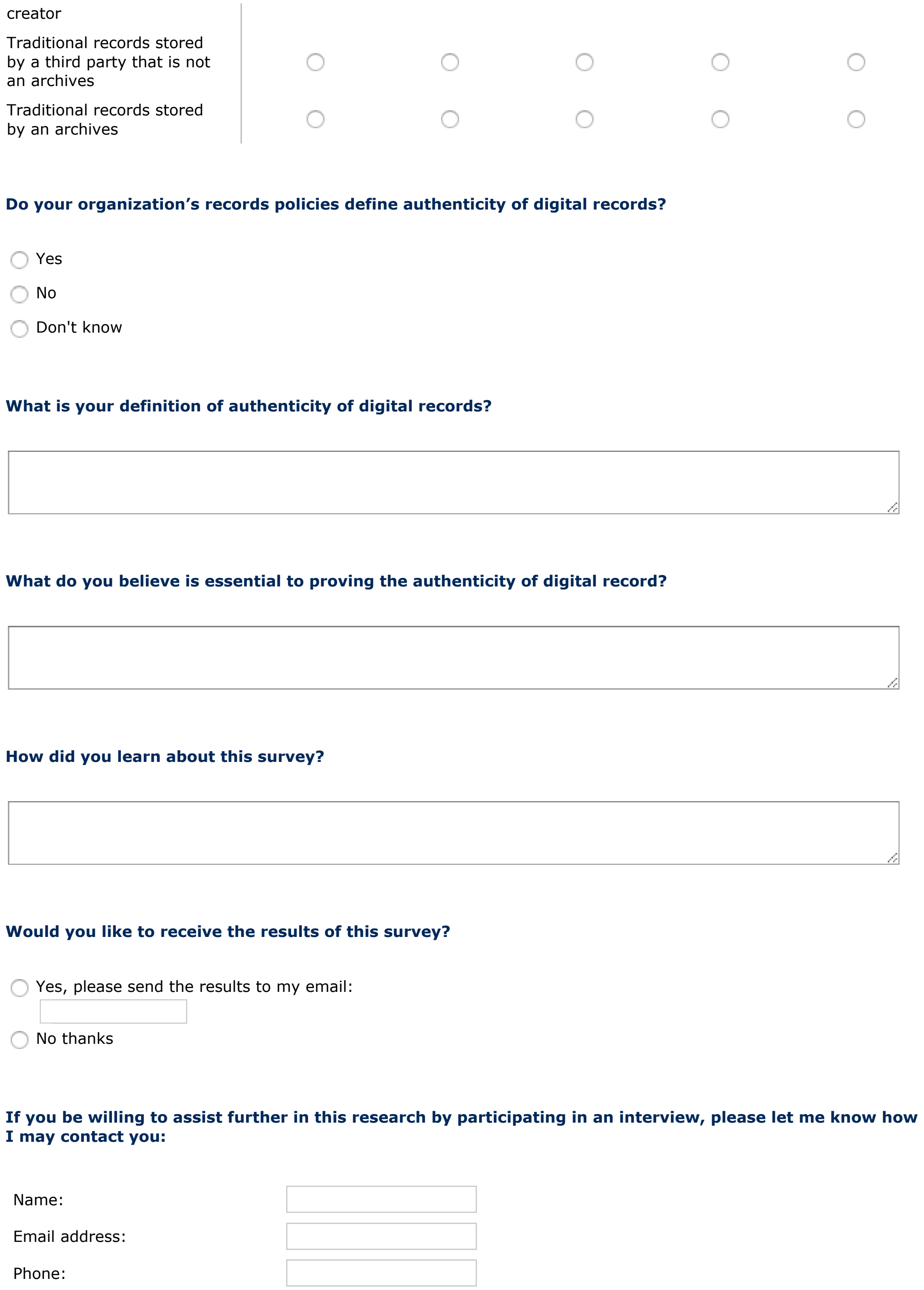
	Not at all Important	Very Unimportant	Neither Important nor Unimportant	Very Important	Extremely Important
Written policies and procedures governing the management of the records system	○	○	○	○	○
Documentation about the record system (design, operation, management, etc.)	○	○	○	○	○



Written policies and procedures governing digital records	○	○	○	○	○
Information about the software used to create and manage the digital records	○	○	○	○	○
Information about changes made to the digital records, (e.g. migration, normalization, etc.)	○	○	○	○	○
Information about actions taken to preserve the digital records	○	○	○	○	○
Classification scheme and/or file plan	○	○	○	○	○
Retention and disposition schedules	○	○	○	○	○
Archival description	○	○	○	○	○
Access controls/security measures	○	○	○	○	○
Audit logs	○	○	○	○	○
Cryptographic validation techniques (e.g. digital signatures, hash digests, etc.)	○	○	○	○	○
Standardized metadata	○	○	○	○	○

Based on a consideration of storage only, how much confidence would you have in the authenticity of records in the following storage options, all else being equal?

	No confidence	Little confidence	Neither confidence nor lack of confidence	Considerable confidence	Total confidence
Digital records stored by their creator on removable media (i.e. a USB key/external hard drive, optical or magnetic media)	○	○	○	○	○
Digital records stored by their creator on stand-alone computers	○	○	○	○	○
Digital records stored by their creator in network drives/filing system	○	○	○	○	○
Digital records in cloud storage maintained by a third party cloud service provider	○	○	○	○	○
Digital records stored by an archives	○	○	○	○	○
Traditional (e.g. paper, microfilm) records stored on- or off-site by their	○	○	○	○	○



creator					
Traditional records stored by a third party that is not an archives	○	○	○	○	○
Traditional records stored by an archives	○	○	○	○	○

Do your organization’s records policies define authenticity of digital records?

○ Yes

○ No

○ Don't know

What is your definition of authenticity of digital records?

What do you believe is essential to proving the authenticity of digital record?

How did you learn about this survey?

Would you like to receive the results of this survey?

○ Yes, please send the results to my email:

○ No thanks

If you be willing to assist further in this research by participating in an interview, please let me know how I may contact you:

Name:

Email address:

Phone:

Appendix 3: Distribution of Responses

Distribution by Country

Count of COUNTRY	
Row Labels	Total
Africa	11
Algeria	1
Kenya	1
South Africa	4
Botswana	2
Nigeria	1
Senegal	1
Uganda	1
Asia	12
Indonesia	1
Israel	2
Lebanon	1
Singapore	4
United Arab Emirates	1
Bangladesh	1
Korea	1
Saudi Arabia	1
Australasia	19
Australia	5
New Zealand	14
Europe	96
Austria	2
Belgium	2
Denmark	1
Finland	1
France	4
Germany	4
Vatican City State	1
Isle Of Man	1
Italy	16
Netherlands	7
Romania	2
Russian Federation	1
Slovenia	2
Switzerland	3
Turkey	1
Luxembourg	2
Norway	1
Portugal	1
Spain	4
Sweden	1
UK	39

Count of COUNTRY	
Row Labels	Total
North America	150
Canada	59
Saint Lucia	1
Barbados	1
Trinidad and Tobago	1
US	88
South America	4
Brazil	3
Colombia	1
Grand Total	**292**

Distribution by Legal system

Count of Legal system	
Row Labels	Total
Civil	64
Austria	2
Belgium	2
Brazil	3
Colombia	1
Denmark	1
Finland	1
France	4
Germany	4
Vatican City State	1
Isle Of Man	1
Italy	16
Korea	1
Lebanon	1
Luxembourg	2
Netherlands	7
Norway	1
Portugal	1
Romania	2
Russian Federation	1
Senegal	1
Slovenia	2
Spain	4
Sweden	1
Switzerland	3
Turkey	1

Count of Legal system	
Row Labels	Total
Common	215
Australia	5
Bangladesh	1
Barbados	1
Canada	59
Kenya	1
New Zealand	14
Nigeria	1
Singapore	4
Trinidad and Tobago	1
Uganda	1
UK	39
US	88
Pluralistic or religious	13
Algeria	1
Botswana	2
Indonesia	1
Israel	2
Saint Lucia	1
Saudi Arabia	1
South Africa	4
United Arab Emirates	1
Grand Total	**292**

Appendix 4: Interview Protocol

This is an anonymized example of the interview protocol sheet that was sent to each interviewee. Each sheet contained the survey answers given by that interviewee, with some answers highlighted in the protocol sheet to guide the interview questions. Each interview followed the same format, but questions were tailored to explore specific answers given in the survey. Questions identified with “Qx” and charts are from the survey, and the numbered questions that follow are specific to the interview and each interviewee.

Interviewee Profile Sheet (example)

ID	D000
Employment	[employment type]
Sector	[sector type]
Location	
Age	[range]
Education	[degrees]
Date of interview	[date and time, researcher and interviewee time zone]

Interview preface: This research explores the degree to which a traditional archival model of authenticity (that a record is what it claims to be and is complete and uncorrupted in all essential respects) is valid in the digital environment generally, and what further stresses may be introduced by new technologies such as cloud storage and processing. The web-based survey sent to professional listservs intentionally did not offer a definition of authenticity, seeking instead to gather respondents’ definitions based on practice.

Q9 How often do you conduct the following tasks with respect to digital records (e.g. electronic documents, images, data, data sets, database records, electronically stored information [ESI], web pages, etc.)? If self-employed or retired, please refer to the job or contract you feel is most relevant.

	Never	Rarely	Sometimes	Often	Very Often
Conduct retrieval and access	• X	•	•	•	•
Monitor or enforce security/access privileges	•	•	• X	•	•
Monitor or enforce privacy of personal information	•	• X	•	•	•
Monitor or enforce compliance with record keeping regulations/policies (including e-discovery)	•			X	•
Conduct preservation or curation	•	• X	•	•	•
Design systems for storage and management of records	•	•	•	•	• X
Design information/records policies	•	•	•	•	• X
Manage records or information	•	•	•	•	• X
Manage/design metadata	•	•	•	•	• X
Other	•	•	•	•	•

1. When you design systems for storage and management of records, how do you account for, or incorporate the collection of authenticity information?

2. When you design information/records policies, how do you account for, or incorporate the collection of authenticity information?

3. Please describe your work managing or designing metadata.

Q10 When you create or manage digital records, how often do you rely on or apply the following to ensure their authenticity?

	Never	Rarely	Sometimes	Most of the Time	Always
S Written policies and procedures governing the management of the records system	•	•		•	X
S Documentation about the record system (design, operation, management, etc.)	•	•	•	•	• X
S Written policies and procedures governing digital records	•	•	•	•	• X
T Information about the software used to create and manage the digital records	•	•	•	• X	•
T Information about changes made to the digital records over time (e.g. migration, normalization, etc.)	•	•	•	• X	•
T Information about actions taken to preserve the digital records	•	•	• X	•	•
S Classification scheme and/or file plan	•	•	•	•	• X
S Retention and disposition schedules	• X	•	•	•	•
S Archival description	•	•	•	•	•
T Access controls/security measures	•	•	•	• X	•
T Audit logs	•	•	•	• X	•
T Cryptographic validation techniques (e.g. digital signatures, hash digests, etc.)	•	•	• X	•	•
T Standardized metadata	•		•		• X

4. Which of these categories provide specifically for identity information? Integrity information? Contextual information?

5. Would you agree or disagree with a distinction between social (or traditional) indicators of authenticity and technical indicators, shown as “S” and “T” in the table above?

Q11 How frequently do you use the following cryptographic validation techniques?

	Never	Occasionally	Very Often
Digital signatures	• X	•	
Trusted time stamps	•	• X	•
Checksums	•	• X	•
Hash digests	• X	•	•
Secure transmission	•		• X

6. In what circumstances are these techniques beneficial or required for an assessment of authenticity?
7. are these useful for short-term or long-term assessments of authenticity?

Q12 What metadata do you routinely use or manage? Please check all that apply.

	DC	PREMIS	Other
A metadata schema or guideline	•	•	ISO 13081 & Mandatory NZ Recordkeeping Schema
A modification of a schema, customized for your organization	•	•	• Mandatory NZ Recordkeeping Schema
A custom-built metadata schema (designed without elements from existing schemas)	•	•	• Additional custom-built schema where necessary
Metadata generated by the software or record system in use only	•	•	• X
Not sure	•		•

7. Do you find that the metadata in these schemas is sufficient to support an assessment of authenticity?

Q13 Have you ever been required to guarantee or attest to the authenticity of digital records in any of the following circumstances?

Providing testimony in court or administrative hearing	•
Pending litigation or administrative action (e-discovery process)	• X
Authenticating copies of digital records for research or in response to reference requests	• X
Other	• For tax administration
I have never been required to guarantee or attest to the authenticity of digital records	•

8. Please describe how you have authenticated records in these situations.

Q14 When you were required to guarantee or attest that digital records are authentic, how important were the following in making your assessment?

	Not at all Important	Very Unimportant	Neither Important nor Unimportant	Very Important	Extremely Important
Written policies and procedures governing the management of the records system	•	•		•	X
Documentation about the record system (design, operation, management, etc.)	•	•	•	•	• X
Written policies and procedures governing digital records	•	•	•	•	• X
Information about the software used to create and manage the digital records	•	•	•	• X	•
Information about changes made to the digital records over time, (e.g. migration, normalization, etc.)	•	•	•	• X	•
Information about actions taken to preserve the digital records	•	•	• X	•	•
Classification scheme and/or file plan	•	•	•	• X	•
Retention and disposition schedules	•	•	•	•	• X
Archival description	•	• X	•	•	•
Access controls/security measures	•	•	•	•	• X
Audit logs	•	•	•	• X	•
Cryptographic validation techniques (e.g. digital signatures, hash digests, etc.)	•	•	•	• X	•
Standardized metadata	•		•	X	•

9. Which of these contributed to a determination of identity, integrity, or context of the content being assessed?

Q15 If you needed to assess that digital records are authentic, how important would the following be in making your assessment?

	Not at all Important	Very Unimportant	Neither Important nor Unimportant	Very Important	Extremely Important
Written policies and procedures governing the management of the records system	•	•	•		• X
Documentation about the record system (design, operation, management, etc.)	•	•	•	• X	•
Written policies and procedures governing digital records	•	•	•	•	• X
Information about the software used to create and manage the digital records	•	•	•	• X	•
Information about changes made to the digital records, (e.g. migration, normalization, etc.)	•	•	•	•	• X
Information about actions taken to preserve the digital records	•	•	• X	•	•
Classification scheme and/or file plan	•	•	•	• X	•
Retention and disposition schedules	•	•	•	•	• X
Archival description	•	• X	•	•	•
Access controls/security measures	•	•	•	• X	•
Audit logs	•	•	•	• X	•
Cryptographic validation techniques (e.g. digital signatures, hash digests, etc.)	•	•	• X	•	•
Standardized metadata	•	•	X	•	

10. Several of these items you have rated as having greater or lesser importance in this question (for a hypothetical situation) than in the previous – can you elaborate?

Q16 Based on a consideration of storage only, how much confidence would you have in the authenticity of records in the following storage options, all else being equal?

	No confidence	Little confidence	Neither confidence nor lack of confidence	Considerable confidence	Total confidence
Digital records stored by their creator on removable media (i.e. a USB key/external hard drive, optical or magnetic media)	•	• X		•	
Digital records stored by their creator on stand-alone computers	•	• X	•	•	•
Digital records stored by their creator in network drives/filing system	•	•	• X	•	•
Digital records in cloud storage maintained by a third party cloud service provider	•	•	• X	•	•
Digital records stored by an archives	•	•	•	• X	•
Traditional (e.g. paper, microfilm) records stored on- or off-site by their creator	•	•	•	• X	•
Traditional records stored by a third party that is not an archives	•	•	•	• X	•
Traditional records stored by an archives	•		•	X	•

11. In what situations would you trust content maintained by a cloud service provider?

12. Do you believe that chain of custody is/can be maintained when content is in the cloud?

Q17 Do your organization’s records policies define authenticity of digital records?
Y/N/Don’t know
YES
[discussion as relevant]

Q18 What is your definition of authenticity of digital records?
We use the definition from ISO 15489
[discussion as relevant]

Q19 What do you believe is essential to proving the authenticity of digital record?
Contextual metadata

13. Please elaborate

14. Is there legislation in your jurisdiction that supports or requires the authenticity of certain records?

15. Do you think that traditional models or rubrics of authenticity still suffice in the digital environment?

Thank you!

Appendix 5: Code Book

Code	Archivists- Q18	RIM-Q18	Other-Q18	Total
concept>integrity	35	27	18	80
concept>identity	24	26	11	61
concept>provenance	6	6	6	18
practice>audit	4	7	5	16
concept>reliability	6	6	1	13
concept>trustworthy	5	7	1	13
practice>secureStorage	0	9	3	12
concept>standard	1	6	4	11
practice>metadata	5	3	0	8
practice>usualOrdinaryCourseofBusiness	3	5	0	8
practice>procedures	5	2	1	8
concept>accuracy	2	3	2	7
concept>chain_of_custody	4	1	2	7
concept>usability	4	1	1	6
practice>policies	4	1	0	5
practice>pragmaticApproach	2	1	2	5
concept>originality	0	1	3	4
practice>standards_legislation	2	1	0	3
practice>documentation	1	1	1	3
practice>accessSecurityControls	0	1	2	3
practice>checksum	0	0	2	2
concept>security>system	0	2	0	2
concept>unknown	0	0	2	2
practice>digitalSignature	1	0	0	1
practice>persistentID	0	0	1	1
concept>content	0	1	0	1
concept>compliance	1	0	0	1
practice>redundancy	0	0	1	1
practice>secureTransmission	1	0	0	1
practice>testimony	1	0	0	1
concept>secureStorage	1	0	0	1
concept>authentication	0	0	1	1
practice>cryptographicValidation	1	0	0	1
concept>chain_of_preservation	1	0	0	1
Total	120	118	70	308