Building Foundations for Digital Records Forensics: A Comparative Study of the Concept of Reproduction in Digital Records Management and Digital Forensics

Sherry L. Xie

Abstract

The Digital Records Forensics project is a research collaboration among the fields of digital records management, law, and police investigation. It seeks to develop concepts and methods for determining the authenticity of digital records when they no longer exist in their originating environment. The project began with comparative studies of scholarly literature in each field to lay a conceptual foundation on which other research methodologies, such as analysis of case law, case study, and ethnography, can be designed and executed. The project expects that this conceptual foundation, along with findings from the other methodologies, will facilitate the proposal of a new discipline called digital records forensics, which will be beneficial to all relevant professions, with complementary strengths deriving from each participating field. This article reports on one of the comparative studies, which examined the concept of reproduction in the fields of digital records management and digital forensics. It presents findings of this examination as well as implications for both fields, with special emphasis on digital records management.

The author would like to express her sincere gratitude to the two anonymous reviewers and the editor for their critiques of and insights regarding this paper. Gratitude also goes to Dr. Luciana Duranti and colleague Corinne Rogers for their encouragement and comments on the first draft of the paper. All errors remain the sole responsibility of the author.

Digital records forensics (DRF) is a research project (2008–2011)[1] led by experts in the fields of digital records management, law, and police investigation that seeks to develop concepts and methods for determining the authenticity of digital records when they no longer exist in their originating environment.[2] Digital records are digital information utilized by humans to fulfill certain purposes, and they exist as memory and evidence of those activities. To function as memory and evidence, digital records must be assessed and maintained as authentic entities. Long a central responsibility of the records professions (i.e., records management and archival administration), maintaining the authenticity of records has become challenging due to complex and rapidly evolving digital technologies. With the pervasive presence of digital records, the possibility of digital records being used as evidence in legal proceedings has greatly increased. This new type of evidence poses challenges in its collection, processing, maintenance, and presentation in court. Each of these steps involves establishing and demonstrating authenticity of the potential digital evidence, typically handled outside the environment in which it originated. Digital forensics emerged as a response to these challenges and has evolved into an independent field over a fairly short time.[3] The reality of independent fields facing the same or similar challenges regarding authenticity provides the context in which the disciplines of digital records management and digital forensics work together to enable the usefulness of digital records and digital evidence whenever such needs arise.

The DRF project utilizes a number of research methodologies, including, among others, comparative study. This method is used to study scholarly literature in the participating fields first to understand the core knowledge of each discipline and then to distill relationships among them to build a conceptual foundation on which other research methodologies, such as analysis of case law, case study, and ethnography, can be effectively designed and executed. The project expects that this conceptual foundation, along with findings from the other methodologies, will facilitate the proposal of a new discipline called digital records forensics, which will harmonize strengths from each field and benefit all relevant professions.[4] This article reports on one such comparative study, which examined the concept of *reproduction* in the fields of digital records

[1] Funded by the Social Science and Humanities Research Council of Canada.

[2] See the Digital Records Forensics Project website at http://digitalrecordsforensics.org, accessed 22 December 2010, for more information.

[3] Digital forensics techniques were first developed forty years ago with data recovery. The field of digital forensics entered its golden age in 1999. Simson L. Garfinkel, "Digital Forensics Research: The Next 10 Years," *Digital Investigation* 7, Supplement 1 (2010): S65–S66. Consider also that the Digital Forensics Research Workshop (DFRWS) held its first workshop in 2001. Digital Forensics Research Conference, homepage, http://www.dfrws.org/index.shtml, accessed 22 December 2010.

[4] For an introduction to the emerging new discipline, see Luciana Duranti, "From Digital Diplomatics to Digital Records Forensics," *Archivaria* 68 (2009): 39–66.

management (DRM) and digital forensics. The research decision to focus on the concept of reproduction arose from the general analysis of the articles, websites, and case law relating to digital forensics that the DRF project gathered and annotated at its preparation stage. The term surfaced during the process and signaled the need for in-depth examination. This article examines the different usages and analyzes the implications for both fields, especially digital records management.

Method of Comparison

As foundation research, this study is limited to the comparison of literature. The criteria for selecting literature in the field of digital forensics are mainly based on relevance and sufficiency, that is, whether or not the literature contains explanations of the concept(s) to be compared and whether or not the explanations are sufficient for understanding. The selection process started with the articles, websites, and case laws gathered by the DRF project,[5] the analysis of which demonstrated the need for more systematic and targeted sources. Monographs in the databases to which the Libraries of the University of British Columbia subscribed were then sought. The database search process first queried the title field with keywords such as "computer AND forensics," "digital AND forensics," and "digital AND evidence," and then expanded the returned results through the subject headings assigned to these titles. Both relevance and sufficiency of these sources were established by assessing tables of contents and/or chapter abstracts. In addition, the selection limited sources to the United States, which is one of the two geographic areas the DRF project currently focuses on.[6]

The selection of literature in the digital records management field proved to be less straightforward. As an established and continuously evolving field, DRM is associated with a larger body of literature accumulated over its much longer time both as a discipline and as a profession. More significant, the records community worldwide neither unanimously accepts nor consistently uses any authoritative text on concepts and methodologies. More often, concepts and methods need to be understood within a particular context. For this reason, the study decided to rely on one representative body of literature in the records community—that produced by the InterPARES (International Research on Permanent Authentic Records in Electronic Systems) project. The InterPARES project maintains the development of concepts as one of its primary goals during all phases of its research, which yields a comprehensive terminology database

[5] See the DRF website at http://digitalrecordsforensics.org/drf_links.cfm.

[6] The other is Canada.

pertinent to DRM. Because terms in this database were developed in a methodologically consistent research environment, they possess the same theoretical roots and the same manner of adapting to new contexts, making illustrating or consolidating relationships among related concepts relatively straightforward.

Another reason for the decision to focus on the InterPARES vocabulary came from the consideration of adding value to InterPARES findings. Because of the large number of terms and the need to continuously adapt to new research findings, the project does not consider it practical or desirable to consolidate and/or synthesize related concepts in the absence of a particular purpose. The basic organization of the terms in the database is alphabetic, with cross references among conceptual relationships in a general manner (that is, the use of "see also" notes). This way, the database functions as a foundation for any type of DRM research that needs to consult terms and definitions. For the current study, the goal of examining the concept of reproduction in different disciplinary fields served a particular purpose, which determined the identification and selection of relevant concepts in the massive terminology database. The analysis of the selected concepts and the outcomes of it contribute to the dissemination and application of InterPARES findings.

These selection criteria are not intended in any way to exclude any other comparisons between the literature of the digital forensics field and other types of literature in the records field, such as glossaries developed by professional societies, other research projects, and international or national standards. In fact, the DRF project hopes that future studies can build on, or compare to, the current one.

Reproduction in Digital Records Management

Digital records management is part of records management,[7] but it entails managerial mechanisms different from many of those used for records in analog formats. One such area is digital records preservation. Preservation of analog records has traditionally been one of the central management functions of archival institutions, but organizational records management programs emphasize it less. Now, however, digital records preservation has become central to records management programs due to the instability of digital technologies, which typically become obsolete while records relying on them are still needed. Long-term preservation strategies and methods for those records that exist longer than the technologies that support them thus become integral components of digital records management. This long-term preservation need renders traditional preservation strategies and methods largely irrelevant to digital records

[7] "Records management" here refers to the management of current records in organizations, i.e., records under the control of their originating organizations.

preservation. As concluded by the InterPARES project, preservation of digital records demands managerial care at all stages during the records' existence: creation, primary use[8] (including maintenance[9]), disposition (i.e., either transfer to a designated custodian or destruction), and secondary use, contrary to traditional preservation strategies, which only takes place after records are transferred to an archives or they acquire a permanent status. Underlying this paradigm shift are fundamental building blocks resulting from the project's carefully crafted research domains, focuses, and questions.[10] The concept of reproduction is one of these fundamental building blocks.

Reproduce, Reproduction, and Reproducibility

As a verb, *reproduce* means "to produce . . . again," and it can be used in a variety of situations, both materially and intellectually. According to the *Oxford English Dictionary*, for example, it can refer to producing a chemical substance or a physical radiation, representing to the mind through a mental effort, replaying sound recorded on another occasion by electrical or mechanical means, and so on.[11] The noun, *reproduction*, and the adjective, *reproducible*, refer to the root meaning of bringing something into existence again. For the records professions, the term *reproduce* was traditionally associated with the *use* of records, in cases where originals could not, or should not, be presented to users for consultation.[12] The archival definition of *reproduce* is "to make a copy,"[13] an obvious adaption of the term's dictionary meaning, "to present again or replicate in writing or print." It refers to the generation of access copies[14] and is not meant

[8] In this paper, "primary use" refers to the use of records by the organization that created them for operational purposes; "secondary use," by contrast, refers to any other uses. The terms *primary* and *secondary* do not imply in any way the use's significance, but simply denote different types of usage. It is necessary to establish differences in usage because how records are used largely determines how records should be maintained or preserved.

[9] In this paper, "maintenance" refers to the situation where records' retention periods are shorter than the life span of the technologies used to support records' existence, while "preservation" refers to the situation where records' retention periods are longer than the life span of the technologies used to support records' existence.

[10] For the findings of the first two phases of the InterPARES project, see InterPARES Project, "Project Overview," http://www.interpares.org, accessed 20 December 2010.

[11] *Oxford English Dictionary*, 2nd ed., s.v. "reproduce."

[12] See, for example, "Terms Governing the Reproduction and Use of Material from the Collection of Library and Archives Canada," Library and Archives Canada, http://www.collectionscanada.gc.ca/the-public/005-6040-e.html, accessed 20 December 2010.

[13] InterPARES 2 Project, "Terminology Database," http://www.interpares.org/ip2/ip2_terminology_db.cfm, accessed 30 July 2010.

[14] In Richard Pearce-Moses, *A Glossary of Archival and Records Terminology*, an *access copy* is "a reproduction of a document created for use by patrons, protecting the original from wear or theft; a use copy." Society of American Archivists, http://www.archivists.org/glossary/term_details.asp?DefinitionKey=1584, accessed 20 December 2010.

to be used in relation to preservation. Preservation of records in the analog world seeks to prolong the life span of original documents for as long as possible.[15] Making a copy of an original is not considered a way of doing this, and it is usually not difficult to tell the copy from the original.

The term *reproduction* has acquired a completely new meaning for records in digital formats, one of direct relevance to preservation (while still also relevant to user consultation of records). As one of its fundamental findings, the InterPARES project concluded that empirically it is impossible to preserve digital records due to their innate construction; instead, it is only possible to preserve the ability to reproduce the records.[16] In other words, *reproduction* of digital records has become the only means that human users can rely on to re-access them after the first time they are saved to a storage medium, regardless of how long they will exist. Hence, preservation of the ability to reproduce—reproducibility—has become the cornerstone of digital records preservation.

The implications of this finding for the records field are profound, and both theories and methodologies reflect them. Theoretically, they ignite rethinking and reconstruction of archival concepts and methodologically; they extend traditional records preservation and other records management activities, such as identifying records and developing retention schedules, into a whole new territory.

Theoretical Implications

The theoretical implications of preserving reproducibility can be characterized by three groups of concepts: influenced concepts, concepts with newly acquired relevance, and invented concepts.

[15] In Pearce-Moses, *A Glossary of Archival and Records Terminology*, *preservation* refers to "the professional discipline of protecting materials by minimizing chemical and physical deterioration and damage to minimize the loss of information and to extend the life of cultural property." Society of American Archivists, http://www.archivists.org/glossary/term_details.asp?DefinitionKey=78, accessed 20 December 2010. Footnotes 15 and 16 cite SAA, not InterPARES definitions for better illustration purposes. The InterPARES project, while acknowledging the existence of analog records in today's organizations, has been continuously incorporating findings from digital records research into its terminology development. Therefore, the definitions of the terms in the terminology database attempt to capture only the essence of the concept that can be generalized without considering the types of records (i.e., analog or digital). This makes them less effective for illustrating aspects characteristic only of analog materials. For example, InterPARES defines *preservation* as "The whole of the principles, policies, rules and strategies aimed at prolonging the existence of an object by maintaining it in a condition suitable for use, either in its original format or in a more persistent format, while leaving intact the object's intellectual form," while the more specific SAA definition cites "chemical and physical deterioration and damage."

[16] "Preservation Task Force Report," InterPARES 1, http://www.interpares.org/book/interpares_book_f_part3.pdf, accessed 13 August 2010.

Influenced concepts

Examples in the first group include concepts of document, record, and original record—concepts fundamental to the archival discipline. To preserve reproducibility requires the concept of *document*[17] to adapt to the reality that a document in the digital world now only exists on the computer screen, where it is constructed. Once saved to a storage medium, the document dissolves itself into discrete bit streams, which then may be physically spread across the entire storage medium, with logical relationships discernable only to the operating system's file system. Therefore, to re-access the document, both hardware (typically the keyboard, mouse, storage medium, and monitor) and software (operating system and application[s] specific to the document) become central, for only with them can the document be re-assembled and re-presented—reproduced—as a unit interpretable by human users. Thus, the concept of *document* in the digital world now has two clearly distinguishable yet interrelated dimensions: a cognitive one that still treats a document as a coherent whole and an operational one that recognizes the composition of a document's preservable parts as required by reproduction. Both dimensions are necessary because the cognitive dimension maintains the human users' understanding of the content and context of digital documents, and the operational dimension provides guidance for the actual preservation activities.

The adaptation of the concept of document to preservation reproducibility impacts inevitably the concept of *records*, because, in archival science, records are a subset of documents. While fundamentally discernible from other types of documents by their role in practical activities, records have basic features of documents such as documentary form and affixation to a medium.[18] To exist and function properly, digital records need to be understood as possessing the same two dimensions as digital documents, which, however, does not change their different nature. Although both digital documents and digital records appear to be the same in regard to the challenge of preserving reproducibility technologically, they remain conceptually distinct.

[17] *Document* means "recorded information or object which can be treated as a unit." ISO 15489-1 *Information and Documentation—Records Management—Part 1: General* (Geneva: International Organization for Standardization, 2001).

[18] A *record* is "a document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference." InterPARES 2 Terminology Database.

The concept of *original record* in traditional archival diplomatics[19] is associated with a record's status of transmission. The status of transmission refers to a record's degree of perfection in terms of its completeness, primitiveness, and effectiveness. A record is complete if it is produced in accordance with the documentary form[20] intended by its author and/or required by the juridical system; it is primitive if it is the first to be produced in the complete form; and it is effective if it is capable of accomplishing the effects for which it was produced.[21] Among the three typical types of a record's status of transmission—draft, original, and copy—the original is the one that is complete, primitive, and effective. The establishment of a record's original status is important because it assigns the quality of authoritativeness to the record: an original is always more trustworthy than its drafts and copies.[22] The concept of *original*, however, is now challenged by the notion of preserving reproducibility. Because the only way to re-access digital records after the first time they are saved to a physical storage medium is to reproduce them, two of the three defining characteristics of an original, primitiveness and effectiveness, can no longer co-exist. To be primitive, a completed original has to exist on the screen of a computer and cannot be saved to a physical storage medium because the saving will dissolve the record into human-illegible pieces and the subsequent view of the record can only happen when a reproduction of the first screen manifestation is made. The reproduction is no longer an original but a copy. To be effective, on the other hand, the completed record needs to cross either time, space, or both to reach its intended recipients, which, however, cannot be accomplished without the record being saved. As a result, while the term *original record* is still used for convenience purposes, the concept of *original*—as defined for traditional records—ceases to exist. In the digital world, every instantiation of a record is now a copy.

[19] *Traditional archival diplomatics* refers to the body of knowledge consisting of concepts and principles that were obtained from the disciplines of archival science and diplomatics (general) and synthesized as a cohesive one. It first appeared in 1989 when Luciana Duranti started to publish articles on classic diplomatics (general) and its applicability to contemporary records. It subsequently served as the theoretical framework for research projects including the UBC project, the InterPARES project, and the Digital Records Forensics project. With inputs from the rich findings of these projects, traditional archival diplomatics has transformed into digital diplomatics, with both widened and deepened knowledge regarding digital records management (including long-term preservation).

[20] *Documentary form* refers to the rules of representation according to which the content of a record, its administrative and documentary context, and its authority are communicated. Documentary form possesses both extrinsic and intrinsic elements. InterPARES 2 Terminology Database.

[21] "Template 3 What is a Reliable Record in the Traditional Environment," The UBC project, http://www.interpares.org/UBCProject/tem3.htm, accessed 10 August 2010.

[22] "Ontology B: Concept of Status of Transmission of Record," InterPARES 2 Terminology Database.

Concepts with newly acquired relevance

The concept of *copy* is the exemplar in the second group—concepts with newly acquired relevance. As introduced above, within the field of archival administration, to reproduce records means to make a copy of a record, primarily for the purpose of facilitating use. Owing to the different types of usage, the concept of *copy* in fact includes a family of related concepts such as *copy in the form of original*, *conformed copy*, *imitative copy*, *simple copy*, and *authentic copy*.[23]

A *copy in the form of the original* appears identical to the original and has the same effects, but is generated subsequently.[24] A *conformed copy* transcribes some (usually major) portions of a record in an exact manner and replaces others that could not or were not transcribed with written explanations.[25] An *imitative copy* reproduces a record in exactness and completeness but in such a way that it is always possible to tell the copy from the record being copied.[26] A *simple copy* only reproduces the content of a record, either partially or completely, and an *authentic copy*[27] is certified by an authorized official so as to make it trustworthy. In light of these different types of copies, an all-encompassing definition in this context refers to a copy as a duplicate of a record[28] resulting from a reproduction process.[29] Because preservation of digital records means to preserve records' reproducibility, the concept of *copy* (including all its family members), while previously distant to records' preservation, has acquired in the digital environment a direct relevance to it.

Invented concepts

As demonstrated by the InterPARES case studies, preserving reproducibility requires more than just rethinking and revising existing concepts. If digital records preservation is to be properly done, new concepts are needed to describe

[23] Note that the concept of *backup copy* is not included here because it is not usually related to individual records.

[24] An example of this type of copy in the paper world can be a letter re-sent to its recipient who, for whatever reason, lost the original. The re-sent letter is identical to the original in completeness and effectiveness but different in primitiveness. The re-sent letter has a different medium and some new metadata (e.g., issue date, receiving date, an annotation of "re-sent," etc.), which essentially make the re-sent letter a copy.

[25] One example of untranscribable content could be a handwritten signature.

[26] A photocopy is an example of this type of copy.

[27] Note that the "copy" in the definition does not have any qualifiers to specify its type. This means that any and all kinds of copies can be authenticated, as long as the authentication process is carried out by a person who is entrusted with the power to do so. This has important implications for digital preservation.

[28] Note that the "record" in the definition can be in any of the record's three transmission statuses: draft, original, and copy.

[29] Adapted from InterPARES Terminology Database.

the new types and new functions of digital records, as well as the new ways of preserving and providing access to them.[30] Examples in this group of invented concepts include manifested digital record and stored digital record (new types of records), instructive record and enabling record (new functions of records), reproduced digital record and reproducible digital record (new ways of preservation), and digital component.

The concept of *digital component* was the InterPARES project's first response to the finding of reproducibility.[31] The Preservation Task Force of the first phase of the project (InterPARES 1) constructed the concept and used it to represent the technologically distinguishable constituent parts of a digital record. A digital record may have one or more digital components, which are determined by the way the bits are stored and by the software application that renders them. For a record to be preserved, each of its digital components must be preserved. Based on this concept and with new findings from the dynamic, interactive digital world, the second phase of the project (InterPARES 2) established the concepts of *manifested* and *stored records.* A manifested record is the visualization or materialization of its digital component(s) in a form suitable for presentation to a person or another system. A stored record is a digitally encoded object that is managed as a record.[32] The concept of *stored record* is an expansion of the concept of *digital component,* which was originally conceived as only a part of a record, not, by itself, a record. With reference to the dynamic, interactive digital world, the concept of *stored record* recognizes that although certain digital objects do not appear as parts of a manifested record (thus they are not digital components as defined by InterPARES 1), they are nevertheless necessary for bringing the record back to its manifested form. They therefore should be preserved, along with other manifestable digital components, as a part of the corresponding record's reproducibility. In addition, one such digital object, such as a software patch, may participate in processes that reproduce different records if they present the same technological requirements for the reassembling. The concept of *stored record* thus benefits the managerial considerations for long-term preservation in terms of identifying and maintaining the one-to-many relationship.

Representing two newly established records functions, *instructive records* refer to those containing instructions about executing an action or process, and

[30] The new concepts regarding use/access will not be introduced here but can be consulted in Luciana Duranti and Randy Preston, eds., *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records* (2008), InterPARES 2 Project, "InterPARES Book," http://www.interpares.org/ip2/book.cfm, accessed 20 December 2010, in particular, "Chain of Preservation," 195.

[31] "Preservation Task Force Report," InterPARES 1.

[32] Luciana Duranti and Kenneth Thibodeau, "The Concept of Record in Interactive, Experiential and Dynamic Environments: The View of InterPARES," *Archival Science* 6, no. 1 (2006): 13–68.

enabling records refer to those encoded in machine language and actively involved in carrying out an action or process. Contrary to records' traditional retrospective functions,[33] instructive and enabling records are mainly prospective in nature: They are created intentionally either to give instructions for actions to be carried out (by re-assembling digital components) or to play a technological supporting role in the action process (as stored-only records). The two concepts contribute to reproducibility by adding building blocks to a preservation system that must recognize the different functions of records.

Within a trusted preservation system, defined as "The whole of the rules that control the preservation and use of the records of the creator and provide a circumstantial probability of the authenticity of the records, and the tools and mechanisms used to implement those rules,"[34] a *reproduced digital record* is an authentic representation of a digital record reconstituted from its digital component(s), and a *reproducible digital record* refers to a unit that includes the digital component(s) of a record and the technical information or software necessary to reproduce and manifest it from the digital component(s). A reproduced digital record typically conforms to its reproducible digital record, which, however, may not represent the "original" record in its entirety. When the "original" is challenging for complete preservation, determination of the purposes of preservation for concrete scenarios and cost-effective analysis come into play. The formulation of reproducible digital records has emerged as a distinct and increasingly important activity in digital records preservation.

Methodological Implications

The methodological implications of preserving the ability to reproduce are indicated by the term *ability*. For digital preservation to be materially carried out, rigorous conceptual development is only the foundation. Based on this foundation, the records profession needs to acquire a variety of abilities, such as

- the ability to understand the digital environments in which records were created and preserved;
- the ability to identify records among other digital informational objects co-existing in the same environment;

[33] The retrospective nature here refers to the traditional records function of being either dispositive or probative. Luciana Duranti, *Diplomatics: New Uses for an Old Science* (Chicago: Society of American Archivists, Association of Canadian Archivists, and Scarecrow Press, 1998). Adapting to the contemporary business environment, a *dispositive record* documents the will or decision of commencing an action (or a series of actions), e.g., an enacted act. A *probative record* documents an action (or a series of actions) that was already completed, e.g., a diploma for a degree. These functions relate directly to the action that caused the generation of a record, not to the subsequent actions in which the record participates. In subsequent actions, the record is typically used as an instrument, thus functioning proactively instead of retrospectively.

[34] InterPARES 2 Terminology Database.

- the ability to decompose a manifested record into its digital components;
- the ability to identify stored records in relation to preservation requirements;
- the ability to establish retention schedules for both manifested and stored records;
- the ability to understand technologies that preserve particular digital components, such as format refreshing, migration, emulation, and any others developed in the future;
- the ability to make decisions on the type of copy to be reproduced based on usage purposes; and
- the ability to understand technical authentication measures and their relationships with archival authentication means.

Noticeably, these abilities signal a strong focus on digital technologies. They are listed here to emphasize their importance. Performing digital records preservation requires more than technologies.[35] It will be difficult, however, if not impossible, for the profession to carry out effective digital preservation without these technological abilities. It is worth emphasizing that such abilities are currently lacking.

Reproduction in Digital Forensics

Digital forensics is generally regarded as a branch of the forensic sciences, functionally similar to other forensic fields such as forensic entomology (study of bugs and insects), forensic anthropology (study of bones and skeletons), forensic linguistics (study of language), and the one most familiar to the general public, forensic biology (e.g., DNA analysis). The common goal of these different branches is to collect evidence to be used in a court of law through the application of scientific knowledge and the examination of the objects in question. The major differences among them are the objects they examine and the scientific knowledge and methods specific to each. As a field increasingly recognized as existing in its own right, digital forensics examines digital devices[36] and digital information created or stored by these devices for the purpose of finding evidence admissible in a court of law. Although digital forensics has grown rapidly over the past decade, it currently focuses heavily on methodological

[35] The definition of *preservation system* by InterPARES can illustrate this: "A set of rules governing the permanent intellectual and physical maintenance of acquired records and the tools and mechanisms used to implement these rules." InterPARES 2 Terminology Database.

[36] Refers to those designed to process or handle information encoded in discrete forms of "0"s and "1"s.

development as exemplified by manuals, monographs, and conference themes.[37] As a result, terms and definitions employed by the field are mostly borrowed from other professions, such as those associated with computer science and law enforcement. The concept of *reproduction* is one such example.

Reproduction, Duplicate, and Imaging

Within the context of the legal system in the United States, the association of the concept of reproduction/reproduce with digital forensics derives from the U.S. *Federal Rules of Evidence*,[38] in particular Article X, "Contents of Writings, Recordings, and Photographs." These rules acknowledge electronically compiled data as one type of "Writings and recordings,"[39] thus making them potential evidence admissible in courts, provided that they satisfy evidentiary requirements. One such requirement is Rule 1002, "Requirement of Original." While for data "stored in a computer or similar device," an accurate, readable output is considered an "original,"[40] the rules also allow duplicates with established evidentiary foundation to be accepted as evidence. With two exceptions,[41] Rule 1003 stipulates that "A duplicate is admissible to the same extent as an original." A *duplicate* in this context means "a counterpart[42] produced by the same impression as the original, . . . or by mechanical or electronic re-recording, . . . or by other equivalent techniques which accurately reproduce the original."[43] The rules here utilize the dictionary meaning of the term *reproduce* (i.e., to produce . . . again) and regard electronic/digital means of reproduction as acceptable for generating duplicates. Together, these rules establish two critical points of guidance for those concerned with evidence: first, electronic/digital data are recognized by law as potential evidence, and, second, reproduced data can be used as evidence as well. These points of guidance gave birth to both digital forensics as a field specializing in collecting, analyzing, and reporting on digital evidence and to one of its most central forensic techniques—the reproduction of digital data.

[37] For example, see Digital Forensics Research Workshop 2010, DFRWS, "DFRWS 2010 Agenda," http://dfrws.org/2010/program.shtml, accessed 20 December 2010.

[38] The *Federal Rules of Evidence* govern the introduction of evidence in proceedings, both civil and criminal, in the courts of the United States.

[39] *Federal Rules of Evidence,* Rule 1001. (1). Hence the growing number of publications on "electronic (or digital) evidence."

[40] *Federal Rules of Evidence,* Rule 1001. (3).

[41] The two exceptions are 1) a genuine question is raised as to the authenticity of the original, or 2) in the circumstances it would be unfair to admit the duplicate in lieu of the original.

[42] Note that because these rules are issued within the context of Article X, "counterpart" here refers to information/data-type of evidence (i.e., documentary evidence). Digital devices themselves (see below) can be used as physical or direct evidence if they bear fingerprints and/or DNA materials, and this is relevant to digital forensics in terms of the procedures of collecting evidence.

[43] *Federal Rules of Evidence,* Rule 1001. (4).

These rules are also attributable to the definition of the term *reproducibility* in the digital forensics field, which is different from that in DRM. In DRM, *reproducibility* refers to the ability to reproduce, and the results of the reproducing process are allowed to vary in certain aspects and to certain degree. Yet, in digital forensics, the term means repeatability of forensic processes and results (usually by different operators or tools), which are expected to be consistent.[44] This usage complies with the evidence rules, which, with reference to scientific knowledge as evidence, stipulate that the testimony of an expert is "the product of reliable principles and methods,"[45] and, in science, reliability refers to reproducibility.[46] As a result, the digital forensics literature uses the term *reproduction* only to refer to the action of duplication.

The reason that reproducing (or duplicating) digital data becomes central to the field is due less to the fact that duplicates are allowed by the evidence rules and more to the increasing impossibility of examining originals during a "live investigation" and to the notion of protecting originals seized during a "dead investigation."[47] A *live investigation* refers to a situation when digital devices cannot, or are not desired to, be removed from the incident scene to a forensic lab for examination, and thus data storage devices are duplicated for in-lab examination with information systems still operating. Because digital data are both complex and voluminous, on-site data inspection takes typically too long to be practical. As noted by the U.S. Department of Justice, in the vast majority of cases, forensic analysis of a hard drive takes too long to perform on-site during the initial execution of a search warrant, thus duplicating is necessary in nearly every computer search warrant case.[48] In addition, because some sets of digital data are indispensible for an organization's operation, removing digital data storage servers for investigation appears unreasonable in cases that do not absolutely require doing so.

In a *dead investigation*, data storage devices are removed from the incident scene to a forensic lab. Duplicating data on the removed digital devices is considered best practice because it allows forensic analyses to be performed on identical copies and the originals to be protected in a secure area, isolated from the

[44] See, for example, Brian Neil Levine and Marc Liberatore, "DEX: Digital Evidence Provenance Supporting Reproducibility and Comparison," *Digital Investigation* 6, Supplement 1 (2009): S48–S56.

[45] *Federal Rules of Evidence*, Rule 702. (2).

[46] Federal Judicial Center, "Reference Manual on Scientific Evidence" (2000), http://www.fjc.gov/public/pdf.nsf/lookup/sciman00.pdf/$file/sciman00.pdf, accessed 30 August 2010.

[47] The use of the two "originals" here and in other places in the following sections conforms to the definition by the evidence rules. See footnote 41. This definition is a less restrictive construct compared to the archival definition of *original.* In the digital forensics literature, seized digital devices such as hard drives are sometimes referred to as originals.

[48] Computer Crime and Intellectual Property Section, Department of Justice, "Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations," 78, 86, http://www.justice.gov/criminal/cybercrime/ssmanual/ssmanual2009.pdf, accessed 30 August 2010.

usually long process of evidence collection and examination. When needed, another identical copy can always be made from the original, either for repeating a particular analysis or for applying new forensic analytical tools. In fact, identical copies are made from the device duplicate produced during a live investigation to protect the first duplicate, now treated as the original. Making duplicates therefore becomes an integral part of digital forensics. Among the various methods of data duplication, imaging appears to be most common and basic.

Imaging in Digital Forensics

Generally speaking in the computer world, imaging is an action of copying. Yet it differs from ordinary computer copying in two aspects: first, it copies an entire storage device,[49] and, second, the copying takes place between different systems. Copying an entire storage device means that not only (computer) files displayed by the file system of the operating system (OS) are copied (i.e., logical file copying), but the hidden data—those not displayed by the OS file system—are also copied. The copying cannot happen within the system in which the data exist because copied data usually are further processed in a location different from the site where they were copied (e.g., a forensic lab), and they also need to be isolated from the system for data integrity purposes. Although imaging appears to be the most basic topic in digital forensics literature and is labeled as an essential skill for digital forensics professionals, no systematic deliberations on its definition, its relationship with storage devices,[50] or its relationship with the nature of data (i.e., stable or volatile[51]) seem to exist. For the purpose of this paper, hard drive imaging is used as a typical example due to its universal appearance in literature[52] and its seemingly higher degree of standardization.

[49] According to the U.S. National Institute of Justice, Department of Justice, *storage devices* can be "any medium that can be used to record information electronically." Examples include hard disks, compact discs, thumb drives, memory cards, digital cameras, mobile phones, etc. "Digital Forensics Glossary," http://www.nij.gov/topics/forensics/evidence/digital/digital-glossary.htm, accessed 20 December 2010. See also the Institute's *Electronic Crime Scene Investigation: A Guide for First Responders*, 2nd ed. (2008), 53, https://www.ncjrs.gov/pdffiles1/nij/219941.pdf, accessed 20 December 2010. This publication contains details on these different types of devices and the potential evidence that may be associated with them. See particularly pages 3–9.

[50] Imaging is sometimes used in relation to storage media generally, for example, in Albert J. Marcella and Doug Menendez, *Cyber Forensics: A Field Manual for Collecting, Examining, and Preserving Evidence of Computer Crimes,* 2nd ed. (Boca Raton, Fla. : Auerbach Publications, 2008). Sometimes it is also used with RAM, for example, in Anthony Reyes, *Cyber Crime Investigations—Bridging the Gaps Between Security Professionals, Law Enforcement, and Prosecutors* (Rockland, Mass.: Syngress Publishing, 2007).

[51] *Volatile data* refers to data existing on a live system that will be lost after a computer is powered off. Karen Kent et al., "Guide to Integrating Forensic Techniques into Incident Response SP800-86" (2006), National Institute of Standards and Technology (NIST), Department of Commerce, http://csrc.nist.gov/publications/nistpubs/800-86/SP800-86.pdf, accessed 3 August 2010.

[52] This is, of course, because currently a hard drive is the place where the majority of data are held. See chapter 3 in Michael Sheetz, *Computer Forensics: An Essential Guide for Accountants, Lawyers, and Managers* (Hoboken, N.J. : John Wiley and Sons, 2007).

As noted above, to make an image is to copy a storage device in its entirety. In the case of a hard drive, imaging produces an *image copy* of the entire hard drive, that is, a copy that "duplicates every bit and byte on the target drive including all files, the slack space, Master File Table, and metadata in exactly the order they appear on the original."[53] The key phrase in this definition is "every bit and byte" because only by copying every bit and byte can the entire hard drive be copied. This emphasis is derived from the way a computer's file system works. In tight relationship with the OS,[54] a file system stores and manages data in a computer system in the form of a hierarchy of directories, subdirectories, and files. It determines how computer data are organized and directs where data are written on the hard drive.[55] Specifically, the file system works in two dimensions. First, it stores and manages files in logical relationships determined by (ordinary) computer users who save, retrieve, and delete files at their discretion. Second, it maintains a physical structure irrelevant to those logical relationships in which certain portions of the hard drive are either unrecognizable to, or mistakenly recognized by, the operating system. One example of an unrecognizable portion is the space left from data overwriting when the size of the new file is smaller than the old file being overwritten.[56] When this happens, the remaining space may still contain data from the old file. The mistaken recognition by the OS is mainly due to the way it "deletes" files: the user instruction to delete a file only removes its logical relationship from the file system and does not perform any material action that makes the data disappear. In other words, while the OS will notify the user that the deletion has freed up the space, the system still holds the data in exactly the same places. The data continue to exist until they are completely overwritten by new data—the only action that makes the data disappear. Compared to the data recognized by the OS, the data in these kinds of spaces are hidden, invisible to the OS (and to users). They are, however, responsive to forensic tools (software applications), which are designed intentionally to bypass the OS and to copy the hard drive at bit-stream level. An image copy is therefore also called a *bit-stream image* or a *bit-stream copy*, deliberately emphasizing the particular way by which it was copied.[57] With the imaging method's growing standardization in digital

[53] *United States v. Vilar,* 2007 WL 1075041 (S.D.N.Y. 4 April 2007), quoting Orin S. Kerr, "Searches and Seizures in a Digital World," *Harvard Law Review* 531 (2005): 119, http://volokh.com/files/UNITED_STATES_v_VILAR.pdf, accessed 30 July 2010.

[54] Note that there are different types and versions of operating systems, which may have different types of file systems. The discussion here is generally about the Windows operating system.

[55] Steve Anson and Steve Bunting, *Mastering Windows Network Forensics and Investigation (*Indianapolis: Wiley Pub., 2007), chapter 7.

[56] There are technical terms for this kind of space such as *free space, slack space*, etc. See, for example, slack space in NIST, *Guide to Integrating Forensic Techniques into Incident Response SP800-86* and free space in Marcella and Menendez, *Cyber Forensics,* chapter 5.

[57] Reyes*, Cyber Crime Investigations.*

forensics, an image copy is now very often referred to as a *forensic copy* or simply an *image.*[58]

Imaging Process

In the process of digital forensics, imaging a hard drive takes place during the evidence collection or acquisition phase[59] and can be done either on-site or in a forensic lab, depending on the circumstances surrounding the specific case. Imaging a hard drive involves intensive decision making, and two sets of considerations are central to the process: legal and technical. When the purpose of producing a forensic image is to search for evidence, care must be taken to ensure that the discovered or recovered[60] information from such a process is legally indisputable. The considerations thus include understanding the scope and nature of the search, the order by which different types of evidence are to be collected, measures for guaranteeing evidence integrity and reliability, and so on.[61] From the viewpoint of procedure, legal considerations serve as prerequisites for the technical process of imaging, that is, the copying of the hard drive should begin only with a clear understanding of the legal implications.

The technical considerations for hard drive imaging center on knowledge about computers, networks, and devices associated with them exclusively. For example, knowledge is needed of a computer's various parts, the ways by which the computer connects to its parts and to a network, the computer's working status (i.e., on or off), and, last but not least, forensic tools for imaging and authentication. Without this kind of knowledge, relevant data may not be acquired and acquired data may not satisfy legal requirements. The digital forensics community has produced many manuals and guides precisely for this reason, extensively introducing computer-related knowledge and patiently describing procedural steps in great detail. For example, procedural steps for on-site imaging are:[62]

[58] For example, a forensic copy is "A precise bit-by-bit copy of a computer system's hard drive, including slack and unallocated space." Marcella and Menendez, *Cyber Forensics.* An image is "An accurate digital representation of all data contained on a digital storage device." National Institute of Justice, Department of Justice, *Forensic Examination of Digital Evidence: A Guide for Law Enforcement* (2004), http://www.ncjrs.gov/pdffiles1/nij/199408.pdf, accessed 1 August 2010.

[59] Within the scope of the works consulted in this paper, it appears that a standardized digital forensics process does not exist. Phases are similar in general but vary in order, specific tasks, or language (e.g., collection, acquisition, preservation used for the phase performing similar or same tasks).

[60] The term *discovery* in this paper refers to the locating of visible data, and the term *recovery* refers to the finding of hidden data.

[61] Marcella and Menendez, *Cyber Forensics,* chapter 11.

[62] Adapted from the works cited in this paper, in particular Sheetz, *Computer Forensics.* These steps are by no means complete or authoritative.

- determine if computer is on or off (without touching anything);
- determine live or dead acquisition (by assessing all decisive factors).

The following steps are for dead acquisition:

- remove the computer cover and document the location and type of system components;
- disconnect the hard drive(s) from the system board;
- start the computer and enter the BIOS mode to record all the information contained in the BIOS;
- change the boot-up sequence to instruct the system to look for an OS on either a floppy drive or a CD-ROM drive;
- insert the trusted boot floppy or CD that the digital forensics professionals carry in their "forensic toolkit" and conduct a second controlled boot of the system;
- turn the computer off and reconnect the hard drives;
- do a third controlled boot-up with the forensic CD in the CD-ROM drive, access the BIOS/CMOS setup menu, and collect system information;
- compare the physical information from the manufacturer with the information listed by the system;
- make the acquisition disk forensically clean (by wiping[63]);
- use a combination of software and hardware write-blocking;
- use one or more software solution to transfer the information from the hard drive to the forensic examination disk;
- check software reports for hash values;
- turn off the power to the write-block device and reverse the process to disconnect all cables;
- save the new image file;
- make a duplicate of the duplicate (a working copy for analysis);
- put the hard drive (now considered the original) in a secure storage area.

The digital forensics literature also notes that carrying out these detailed procedures needs to conform to other general digital forensics principles or best practices such as documenting every action taken and every result generated, and maintaining an unbroken chain of custody for both the seized digital devices and the copied images.

[63] Overwriting media or portions of media with random or constant values to hinder the collection of data. NIST, *Guide to Integrating Forensic Techniques into Incident Response.*

Conclusion

This analysis outlines the similarities and differences in the concept of *reproduction* in relation to digital records management and to digital forensics, and in so doing, yields insights that may be instructive for both fields. They both rely on the basic meaning of the concept of reproduction as making a copy, but DRM, as exemplified by the InterPARES project, derives a field-specific concept—*reproducibility*—from this basic meaning to refer to the ability to reproduce the digital components that consist of digital records. The field of digital forensics, on the other hand, maps the basic meaning to the term *imaging*, with added technical requirements for forensic purposes. For both fields, the application of these concepts requires knowledge of relevant legal conditions, for example, authenticity, though they differ in the major mechanisms for establishing authenticity. The certificate issued by a records custodian who has the legal authority to do so authenticates records, yet, for forensic images, hash function is typically used for authentication purposes.

Along with the introduction of the concept of reproducibility, the DRM field focuses strongly on concept building and has established a network of concepts covering almost every aspect of the relevant subject matter. This focus advances the discipline in theory construction, assuring its academic independence. By contrast, the field of digital forensics pays intensive attention to technical considerations and has produced a large number of step-by-step field manuals centering on digital device imaging. Understanding in-depth digital devices in relation to the information stored and processed by them for locating evidence characterizes the profession.

While it is absolutely necessary to emphasize understanding digital technologies for performing professional tasks, the digital forensics profession needs to recognize the equal necessity of establishing a theoretical framework for guiding its field work and setting foundations for future advancement. To build the framework, it needs to first identify the concepts essential to the profession and then establish definitions for them with relation to the profession's specific goals and objectives. The concepts should, at a basic level, include those specifying its professional status, such as *digital forensics* (computer forensics), *digital forensics professional*, and *digital evidence*; those needed for crafting technical manuals, such as *digital device*, *forensic tool*, and *digital forensics process*; and those relating to the admissibility and weight of digital evidence, such as *original*, *copy/duplicate*, *authenticity*, *authentication*, *reproducibility*, *chain of custody*, and *credibility* (as related to evidentiary weight). The literature consulted for this study illustrates confusion in the field's relationships with other fields, such as information security, incident response, and cyber investigation, signaling the insufficiency of its theoretical foundation as an independent discipline. Some of the concepts may be borrowed from disciplines with which the digital forensics profession interacts; however, borrowed concepts need to be understood first in their originating contexts and then harmonized to fit their field-specific functions. For example, the information technology field originally provided definitions of digital devices that need to be examined and refined for forensic imaging purposes because different types of devices may require different imaging techniques. The records community could be of great assistance in this area. The archival discipline has a long tradition of building concepts and analyzing conceptual relationships, and many of the developed concepts are of direct relevance to digital forensics work. Moreover, the records professions are also familiar with legal concepts of evidence because records—public or private—have long provided documentary evidence.

This study found that both fields require understanding of legal requirements of evidence. They rarely, however, interact with each other despite common ground and goals. One example of an archival concept applicable to the digital forensics field is *chain of custody*. This concept has long been established in relation to the assessment of the authenticity of records (i.e., when its chain of custody is proven unbroken, a record is assumed authentic), and it is useful for digital forensics work for the same purpose. In the digital forensics process, collected evidence moves through several sites,[64] and each movement needs to be documented by a trustable entity following established protocols.

Another example is the concept of *authentic copy*, which in archival science is established in relation to a neutral third party (as a trusted custodian) who is recognized by a legal system with official duties to safeguard records' integrity and therefore has the authority to certify (or testify in court) records' authenticity. As explained in the previous section, an authentic copy of a record is not necessarily a copy of the entire record, nor of an original. In other words, the official neutral third party has the capacity to issue a certificate of authenticity for a portion of a record or a copy of a record. The digital forensics profession would benefit from understanding this concept as it could be used to guide the decision for imaging. In *United States v. Hock Chee Koo*, a copy of a portion of all files on a laptop made by a computer analyst using a nonforensic tool was admitted as evidence despite the opposing party's objections that the copy was not an image of the entire hard drive and that the tool was not recognized by the forensics field. The admission was allowed because the court found that the computer analyst had no "desire or inclination to change the contents of the hard drive,"[65] a rationale that corresponds to the concept of neutral third party. It would also

[64] A rough example could be from the incident scene to the forensic lab and then to the court.

[65] *United States v. Hock Chee Koo*, No. 09-321-(2, 3)-KI, LEXIS 20905 (D. Ore., 2011). The admitted evidence is not the copy but an image of the copy made by the FBI. However, the rationale for admission is based more on the creation process of the copy than that of the image. The image made of a laptop by the FBI at the same time was not allowed for admission primarily because the laptop was in possession of the plaintiff, who could not be established as a neutral third party, for two days.

be beneficial for the digital forensics profession to take note of the concept of trusted custodian because it is related to the weight of evidence (i.e., its credibility or trustworthiness). Gaining weight for evidence for a jury usually requires a higher level of proof of authenticity, which would benefit from the testimony of an established trusted custodian who is capable of attesting to the custody history of the record in question and proving its authenticity accordingly.

Apart from building a theoretical framework, the digital forensics profession needs also to understand that the collected digital evidence as well as the documentation created in the course of performing forensic analysis, writing technical reports, and presenting evidence in court *are* records. Therefore, their management should subscribe to records management principles and practices to ensure their authenticity and preservability and to manage them effectively and efficiently. For example, the concept of reproduction in the DRM field and the field's general knowledge regarding long-term preservation of digital records are applicable to digital evidence. Digital evidence is usually needed for much longer than the technologies supporting its existence, and for cases that may be reopened, the need is permanent. How such evidence can be preserved in a way that ensures its authenticity and usability with reliably documented archival bond (i.e., its relationships with other records generated by the same case) is the primary reason that the Forensic Services Section of the Vancouver Police Department joined the DRF project.[66]

On the other hand, the DRM profession can learn much from the technical capabilities of the digital forensics profession. Both fields work with digital materials and both rely on digital technologies for handling the materials. A quick glance at the list of skills necessary for digital records professionals and the technical skills needed for imaging reveals much overlap. However, while the digital forensics profession considers a full and in-depth grasp of all the technological steps compulsory for its professional activities, records professionals rarely demonstrate an adequate and sufficient understanding of all the digital technologies relating to DRM tasks, or, indeed, a positive attitude toward acquiring such knowledge. According to a survey conducted by Cohasset in 2009,[67] 50 percent of the respondents did not know what storage device/media their organizations use for electronic archiving, and 71 percent did not know how many backup tapes are retained to meet the need of storing records.

[66] For a general description of the case study, please visit Digital Records Forensics Project, "Case Studies," http://digitalrecordsforensics.org/drf_case_studies.cfm, accessed 15 April 2011.

[67] Lori J. Ashley and Robert F. Williams, "2009 Electronic Records Management Survey: Call for Sustainable Capabilities," Cohasset Associates, http://www.cohasset.com/retrievePDF.php?id=10, accessed 10 July 2011.The survey invited members of ARMA International , all previous MER (Managing Electronic Records) conference registrants, members of the Records Management Listserv, and members of the Business Forms Management Association (BFMA) Listserv. It received 1,190 responses, a rate of 12 to 14 percent.

Moreover, only 54 percent of electronic records[68] in the responding organizations have retention schedules, and only 26 percent of the responding organizations give the records management programs "primary responsibility for the management of electronic records created and used in the normal course of business," while 41 percent reported that such responsibility rests with the information technology department. For 26 percent, individual business units manage electronic records. When compared to previous results, the survey found that the percentage reporting to the information technology department declined, yet the reporting shifted not to the records management function but to individual business units.[69] Given that records management programs do not control, or even know about, large numbers of digital records, it is not surprising that 65 percent of the respondents reported difficulty retrieving "information from archival storage media in response to legal discovery requests," and only 14 percent stated that when legally challenged, they are confident that their organizations "could successfully demonstrate that [their] electronic records are accurate, reliable and trustworthy—many years after they were created."[70]

This reality should raise serious concerns for the records professions as to the values they offer. In today's organizational settings, where digital technologies continue to predominate, lack of technological knowledge and skills makes it impossible for a records management program to achieve its professional goals of supporting its organization's operational effectiveness and legal compliance. As one of their traditionally valued functions, records custodians authenticate records by certifying authentic copies and providing testimonies, which, in the paper world, are not difficult tasks because of the limited skills required.[71] With digital records, however, more skills are required and technological knowledge has become indispensible. In *American Express Travel Related Services v. Vinhnee*, AMEX asked its records custodian to testify to the authenticity and accuracy of the digital records introduced as evidence. The trial court, however, disallowed these records from being admitted as evidence based on the grounds that the records custodian was not qualified to answer even basic questions about the computer equipment (i.e., hardware and software) by which the digital records were created and maintained. For the same reason, the appellate court ruled that the trial court judge did not abuse discretionary power in disallowing the

[68] Electronic records in the survey are operationalized as three categories: Communications (e.g., emails), Document Objects (e.g., Word documents), and Data Objects (e.g., application data).

[69] Ashley and Williams, "2009 Electronic Records Management Survey," 33.

[70] Ashley and Williams, "2009 Electronic Records Management Survey," 44, 37.

[71] See, for example, in *William Lewis Reece v. The State of Texas*, No. 14-98-00564-CR, Tex. App. LEXIS 4770 (14th Dist. 20 July 2000), where the records manager certified copies in the following way: "I have compared the foregoing and attached copies with their respective originals now on file in my office and each thereof contains and is a full, true, and correct copy from its said original."

evidence and affirmed the decision.[72] It is worth pointing out that the technical questions asked to establish witness qualifications in this case were basic ones, unlike those, for example, relating to the functionalities of an organization-wide digital records management system[73] or to a more complex business environment where the ideas of service-oriented architecture and cloud computing are applied. Digital information in settings like these may still be records. Therefore, their management, including the insurance of authenticity, should be the responsibility of the records management profession. The question now is whether the profession has the ability to act in alignment with the goal.

The goal of the Digital Records Forensics project is one step in this direction because the project studies the challenges of ensuring authenticity of records when they are moved outside their originating information system. When records reside in their originating information system and are used for business operations, their authenticity can be assumed based on the integrity of the system and a business's reliance upon them. This type of circumstantial support becomes invalid when records are moved outside that environment, and the establishment of the authenticity of records thus needs to rely on other types of circumstantial support, such as the traces uncovered by digital forensics.

To argue that it is necessary for the DRM profession to acquire more technological skills is not to suggest that DRM professionals must become technology experts. Expertise in digital technologies resides with the information technology profession; however, DRM professionals need to understand at least the functionalities of records-making and -managing technologies that organizations use to achieve their basic professional goals and tasks. In addition, the need is increasing for DRM professionals to communicate with potential partners outside their organizations such as those associated with the fields of e-discovery and digital forensics, whose work engages intensively digital technologies. An effective collaboration with these professions may substantially influence the outcome of any legal proceeding in which the DRM-sponsoring organization is involved. The DRF project is based on the premise that records custodians can only serve as expert witnesses by acquiring additional skills from the field of digital forensics.[74]

This study, by analyzing the concept of reproduction in two fields participating in the DRF project, demonstrates the need for interrelated fields to collaborate and the expected benefits for both academic development and practical advancement. The project facilitated other investigations into providing

[72] *American Express Travel Related Services Co., Inc. v. Vee Vinhnee,* 336 B.R. 437 (9th Cir. 16 December 2005).

[73] For example, the Records Management Application (RMA), as termed by the "DoD5015.2 Electronic Records Management Software Applications Design Criteria Standard," Department of Defense (2007), http://www.dtic.mil/whs/directives/corres/pdf/501502std.pdf, accessed 16 April 2010.

[74] Duranti, "From Digital Diplomatics to Digital Records Forensics."

an in-depth understanding of key concepts, which, in turn, facilitates effective communications among researchers and between researchers and respondents. With findings generated from other methodologies, the project is currently developing its main research products, a digital records forensics model and curricula for DRF trainings and educational programs. The project believes that it is necessary to propose a new discipline centering on digital records forensics, and it is imperative that the constituent groups in the records community work together to ensure the authenticity, and thus trustworthiness, of records for their entire existence.