Volume 28, Issue 5, October 2012 ISSN 0267-3649


# Computer Law & Security Review

The International Journal of Technology Law and Practice

Editor-in-Chief
Professor Steve Saxby

COMPUTER LAW & SECURITY REVIEW 28 (2012) 522–531

Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.compseconline.com/publications/prodclaw.htm

Computer Law & Security Review

# Trust in digital records: An increasingly cloudy legal area

*Luciana Duranti, Corinne Rogers*

*University of British Columbia, Canada*

*Keywords:*
Digital records
Digital forensics
Cloud computing
Law of evidence
Digital documentary evidence

ABSTRACT

Trust has been defined in many ways, but at its core it involves acting without the knowledge needed to act. Trust in records depends on four types of knowledge about the creator or custodian of the records: reputation, past performance, competence, and the assurance of confidence in future performance. For over half a century society has been developing and adopting new computer technologies for business and communications in both the public and private realm. Frameworks for establishing trust have developed as technology has progressed. Today, individuals and organizations are increasingly saving and accessing records in cloud computing infrastructures, where we cannot assess our trust in records solely on the four types of knowledge used in the past. Drawing on research conducted at the University of British Columbia into the nature of digital records and their trustworthiness, this article presents the conceptual archival and digital forensic frameworks of trust in records and data, and explores the common law legal framework within which questions of trust in documentary evidence are being tested. Issues and challenges specific to cloud computing are introduced.



## 1. Introduction

*Whatever matters to human beings, trust is the atmosphere in which it thrives* (Bok, 1999).

Trust has been defined in many ways but, at its core, it involves willingly acting without the full knowledge needed to act. It consists of substituting the information that one does not have with other information that supports confidence in the action. For example, a person who does not have the knowledge necessary to assess the authenticity of a record relies on the credentials of the expert who authenticates it. Traditionally, trust in records is based on four types of knowledge about their creator and/or their custodian: *reputation*, which results from an evaluation of the trustee's past actions and conduct; *performance*, which is the relationship between the trustee's present actions and the conduct required to fulfil his or her current responsibilities as specified by the truster; *competence*, which consists of having the knowledge, skills, talents, and traits required to be able to perform a task to any given standard; and *confidence*, which is an "assurance of expectation" of action and conduct the truster has in the trustee (Sztompka, 1999; Borland, 2009; Duranti and Rogers, 2011).

For over half a century, society has been developing and adopting new computer technologies for business and communications in both the public and private realm. From the early mainframe computers used by government, business, and academia, whose primary function was computation, to ever faster, smaller, cheaper, and more versatile digital devices used by individuals and organizations alike, whose functions now include automation, communication, commerce, entertainment, education, and citizen engagement, we increasingly rely on and live our lives in the digital realm. Our voracious appetite for technological innovation and engagement has raised a host of challenges to privacy, security and trust. Solutions are being debated in the blogosphere, developed in policy debates, and tested in the courts.

0267-3649/$ – see front matter 
http://dx.doi.org/10.1016/j.clsr.2012.07.009

The Internet spans the globe, erasing national boundaries for the transmission of data, information, documents and records. The interconnectedness of the Internet is forcing us into one global community without the benefit of gradually getting to know one another. When business is transacted over digital networks between people who do not know each other and likely will never meet, establishing trust becomes paramount. What does that mean in respect of policies and practices regarding the handling of digital records residing with Internet services and social media providers? When such policies exist and are sound, the speed with which digital technologies are changing far outpaces society's ability to adapt pre-existing structures and norms. In previous centuries technologies developed in the industrial economy over timeframes long enough to allow users to gain comfort with and trust in the new tools and their impact on society. Digital communications technologies have supplanted the industrial economy, based on the production of goods, with the information economy, based on production, ubiquity, and sharing of information. Trust in the digital environment relies on new methods of establishing and authenticating identity, and managing information in a way that supports security and privacy as well as sharing and access.

Today, individuals and organizations conduct their business in the highly networked, easily compromised environment of the Internet, in which cloud computing for record storage and access is becoming increasingly common. However, policies, practices and infrastructure in "the Cloud" do not currently support an assessment of the four types of knowledge used in the past to establish our trust in records. How can we make decisions related to trust in this new environment? Are there grounds for trusting the institutions and/or professionals who hold digital records about us to make the right decisions about keeping them safe and accessible only to those who have a right to see them, using them for good and in a transparent way, disposing of them when required, and selecting reliable Internet providers for storing and managing them? If yes, what are those grounds? Who has established them, and in the context of what values and purpose?

Issues of trust are difficult to isolate, and are often bound with more easily identified issues of privacy, security, and jurisdiction. Questions arise about the trustworthiness of digital records, or of organizations, service providers, and networked systems, or the juridical framework in which the organizations and systems operate, and in which or with whom the records are stored. As the United States (U.S.) developed the Internet, for example, its social, political, and economic views became evident in its use and rights policies, and this has rankled other countries. Can we trust our records to Web services without fully understanding what legal framework they will fall under? Several recent examples will serve to illustrate this: (1) in January 2012, U.S. federal prosecutors blocked access to the file-sharing site Megaupload.com on charges that the site violated piracy laws, and New Zealand police arrested Megaupload's founder based on the U.S. accusations. As a consequence the data of at least 50 million Megaupload users not implicated in the legal action has been seized and is in danger of being erased; (2) convinced that existing laws cannot deal with growing piracy concerns, the U.S. Congress introduced the *Stop Online Piracy Act (SOPA)*, which resulted in protests across the Internet that persuaded Congress to reject the bill (Maes, 2012); (3) Google established a blanket privacy policy for all materials on its cloud (Google, 2012), while Twitter chose to go in the opposite direction and to adopt the policy of the country of origin of the record (Twitter Blog, 2012). Whom can/should we trust?

Regardless of these warning signs, people trust (often blindly) all kinds of organizations (e.g. banks, phone companies, Internet service providers, social media sites, e-commerce sites) to keep and maintain their data/records/archives on their behalf. In effect they have shifted their trust from the filing cabinet or hard drive in their home office to distributed storage in the cloud, and handed over the stewardship of their personal information to others. Where their records actually reside, how well they are being managed, how long they will be available to them… they have no idea! Many organizations are recognizing this shift and becoming concerned about a liability they may not have thought they were assuming (especially as more and more clients abandon their own recordkeeping and place greater reliance and trust on the recordkeeping abilities of the organizations with which they interact).

'Data' and 'records' are very different in nature. Whereas records are information affixed to a medium in a fixed and stable form (i.e. documents) in the course of activity and kept for further action or reference, data are the smallest meaningful component of information. However, in the digital environment, the issues for records and data coincide. Can the data be trusted? Can the records from which the data are derived be trusted? Are these records complete? Are they authentic? How were they generated, by whom and under what conditions? Is there sufficient contextual information to enable them to be understood? These are some of the questions facing organizations, which are beginning to act on the realization that their data and records holdings are digital assets that need to be managed effectively if they are to be trusted by those making decisions and by clients, customers, citizens, etc. In 2009 the Information Commissioner of Canada wrote: "The poor performance shown by institutions is symptomatic of what has become a *major information management crisis*. A crisis that is only exacerbated with the pace of technological developments." (emphasis in original) (Office of the Information Commissioner of Canada, 2009). This is not an isolated occurrence.

The purpose of this article is to present the conceptual frameworks of trust in records and data developed in the context of archival science and digital forensics, and explore the common law legal framework within which questions of trust in documentary evidence is being tested. Of course, the problem of trust in the digital environment is much larger, encompassing authentication frameworks that allow parties to assess identity while protecting privacy, issues of organizational trust and responsibility, and the broader issues of trusted computing and trusted environments. Our intention is to begin the conversation about ways in which the integration of archival knowledge and digital forensics practices can lead to an understanding of the problems and unintended consequences of wholesale adoption of technology – and in particular the developing area of cloud computing. The article

focuses specifically on the North American legal tradition, as this is the context in which the authors have conducted their research, but the nature of digital material and the challenges posed by new technologies are applicable in any jurisdiction.

## 2. Trustworthiness of digital records: two conceptual frameworks

### 2.1. *Archival science*

In archival science, a record is defined as a document[1] created (i.e., made or received and set aside for further action or reference) by a physical or juridical person in the course of a practical activity as an instrument or by-product of such activity[2] (Duranti and Thibodeau, 2006). According to this definition, a digital record must have 1) an identifiable context; 2) identifiable persons[3] concurring in its creation; 3) an action, in which the record participates or which the record supports either procedurally or as part of the decision-making process; 4) explicit linkages to other records within or outside the digital system, through a classification code or other unique identifier; 5) a fixed form; and 6) a stable content.

We will explain each of these in turn. Records are evidence of actions and transactions – as such, the context of their creation is the framework of action in which they participate. Moving from the general to the specific, for a record, there are five identifiable contexts to consider: the juridical/administrative context, manifested in the laws, regulations, and norms that govern record creation and use; the provenancial context that situates the record within the creating organization or with the legal person, evident from, for example, organizational charts or functional competences; the procedural context, manifested in workflows, policies, or procedures; the documentary context, evident from instruments like classification schemes, indexes, or registers; and the technological context, manifested in documentation about hardware, software, and so on. For example, parliamentary workflow consists of work of government agencies arising form ministerial responsibilities, including to parliament,[4] and includes preparation of ministerial briefs to inform and educate ministers about matters upon which they must speak and act. Parliamentary workflow is performed within guidelines of various statements of government policy and administrative regulations, carried out in a legislative framework (juridical context). Briefs are created by the authority of the agency or ministry (provenancial context), as part of an established workflow (procedural context), captured in the electronic recordkeeping system and given a unique identifier within that system (documentary context), and created and maintained in identified and documented technological environment (technological context).

Persons concurring in records' creation include an author (the physical or juridical person(s) having the authority and capacity to issue the record, or in whose name or by whose command the record has been issued), a writer (the physical or juridical person(s) having the authority and capacity to articulate the content of the record), an originator (the physical or juridical person assigned the electronic address in which the record has been generated or sent),[5] an addressee (the physical or juridical person(s) to whom the record is directed or for whom the record is intended), and a creator (the physical or juridical person in whose archival fonds[6] the record exists). Returning to our example of a ministerial brief, the author would be the agency producing the brief; the writer is the individual within the agency whose responsibility it is to sign ministerial briefs; the originator would be the person identified by the IP or MAC address of the computer from which the brief is transmitted. The addressee of a ministerial brief is the Minister to whom the brief is directed, and the creator is its Ministry.

The next two attributes are straightforward. The action in which the record participates or supports is the subject matter of the record, in our example, the subject of the brief. The linkages will be the classification code or unique identifier that places the brief in question in sequence with related briefs, correspondence, or other records.

Fixed form and stable content are the most problematic characteristics of digital records. One of the great affordances of digital technology is the ease with which digital material can be generated, changed, combined, and shared. A digital record has a fixed form if its binary content is stored so that the message it conveys can be rendered with the same documentary presentation it had on the screen when first saved, even if its digital presentation has been changed, for example, from .doc to .pdf. A digital record has a fixed form as well if the same content can be presented on the screen in several different ways but in a limited series of predetermined possibilities; in such a case we would have different documentary presentations of the same stored record (e.g., statistical data viewed as a pie chart, a bar chart, or a table). Stable content means that the data or content of the record cannot be intentionally or accidentally altered, overwritten or deleted. The content is also considered stable when changes to what we visualize at any given time are limited and controlled by fixed rules, so that the same query or interaction always generates the same result, and we have different views of different subsets of content, due to the

[1] Document is defined as recorded information, where information is intelligence given, that is a message meant to be conveyed.

[2] Thus, while every record is a document, not all documents are records, as only their circumstances of creation, maintenance and use determine whether documents are records.

[3] In archival science, the 'persons' who participate in the creation and use of records are traditionally juridical persons made of one or more human beings, however, in the digital environment, many digital objects that can be identified as records are generated by the interaction of technologies without direct input from human actors, thus the persons involved in their creation are to be identified in terms of systems ownership and use.

[4] This example and the examples that follow are intended to be general in nature for purposes of illustration only, and are not intended to reflect any specific parliamentary system.

[5] Originator, author and writer may be the same physical person but they have different and necessary roles as legal persons.

[6] An archival fonds is the whole of the records made or received by a physical or legal person in the course of activity and kept for action or reference.

intention of the author or to different operating systems or applications, as when we wish to view specific sections of a large legal opinion (MacNeil, 2000; Duranti and Thibodeau, 2006; Duranti, 2009).

The electronic environment poses specific challenges to establishing trust in records. In archival science, records are considered trustworthy if they are reliable, accurate, and authentic. Reliability is defined as the trustworthiness of a record as a statement of fact, based on the competence of its author, its completeness, and the controls on its creation; accuracy is defined as the correctness and precision of a record's content, based on the above *and* on the controls on the recording of content and its transmission; and authenticity is defined as the trustworthiness of a record as a record, meaning that the records is what it purports to be, free from tampering or corruption, based on the competence of its keeper(s) through time (i.e. creator and/or preserver) and on the reliability of the records system(s) in which it resides. Authenticity is composed of both identity and integrity, where identity is the whole of the attributes of a record that characterize it as unique and distinguish it from other records (e.g. date, author, addressee, subject, classification code), and integrity is the quality of a record that is capable of transmitting exactly the message it is meant to communicate in order to achieve its purpose (e.g. fidelity of text and form, and absence of technical changes) (InterPARES, 2011). These attributes all reflect aspects of the reputation, performance, and competence of, and confidence in, the record's keeper(s) from creation to preservation.

If trustworthiness encompasses the qualities of reliability, accuracy, and authenticity (with its sub-qualities of identity and integrity), the process of authentication can only assess authenticity and infer from it the other two qualities, till proof to the contrary. To archivists, authentication is a means of declaring that a *record is what it purports to be* at one particular moment in time. In the digital environment, authentication is often entrusted to a digital signature. The digital signature is functionally equivalent to seals rather than to signatures (i.e. it is an attachment to a record rather than an integral and necessary part of it, like a signature) in that it verifies origin (identity), certifies intactness (integrity), and makes a record indisputable and incontestable (non-repudiation). However, seals are associated with a person while digital signatures are associated with a person and a record (MacNeil, 2000; Duranti, 2009).

### 2.2. *Digital forensics*

In digital forensics, when we wish to establish the trustworthiness of digital records we also must distinguish among the types of digital objects under consideration. This relatively new area of practice divides these objects in three groups: 1) Computer-Stored Documents, which contain human statements and, if created in the course of business, are records (e.g. email messages, word processing documents) and can be used in a court of law as substantive evidence; 2) Computer-Generated Documents, which do not contain human statements, but are the output of a computer program designed to process input following a defined algorithm (e.g. server log-in records from Internet service providers, ATM records, computer-generated animations or simulations) and in a court of law are generally considered demonstrative evidence[7]; and 3) Computer Stored & Generated Documents, which are a combination of the two (e.g. a spreadsheet that has received human input followed by computer processing, that is, by the mathematical operations of the spreadsheet program) and can be used in a court of law in either way.

According to digital forensics, reliability is the trustworthiness of a record as to its *source*, defined in a way that points to either a reliable person (for computer-stored documents) or a reliable software (for computer-generated documents), or both. If the source is a software application, trustworthiness will be more easily established if the application is open source, because the source code will be publicly available. This will allow the processes of records creation and maintenance to be forensically authenticated either by describing a process or system used to produce a result, or by showing that the process or system produces an accurate result. Open source software allows for both types of authentication (Kenneally, 2001; Carrier, 2003b).

Digital forensics considers accuracy to be a component of authenticity and, specifically, integrity. Digital entities are guaranteed accurate if they are repeatable, that is, if the same process carried out on them produces the same outcome. Repeatability, which is one of the fundamental precepts of digital forensics practice, is supported by the documentation of each and every action carried out on the digital evidence. Open source software is also the best choice for assessing accuracy, especially when conversion or migration occur, because the transparency of its code allows for a practical demonstration that nothing could be altered, lost, planted, or destroyed in the process (Mocas, 2004; Carrier, 2003b).

Authenticity, to digital forensics experts, means that the data or contents of the record are what they purport to be and were produced by or came from the source they are claimed to have been produced by or come from. Again, the term "source" is used to refer to a person (physical or juridical), a system, software, or a piece of hardware. As in the archival concept, authenticity implies integrity, but the opposite is not automatically true, that is, integrity does not imply authenticity (because identity must also be confirmed). In fact, the digital forensics view of integrity is much more nuanced than the archival view, for which integrity is simply the quality of being complete and unaltered in all essential respects, a definition that equally applies to data, documents, records, copies, or records systems.

In digital forensics, integrity is distinguished in several types. Data integrity implies that data are not modified either intentionally or accidentally without proper authorization, and is based on bitwise integrity, that is, on the fidelity not only of the bits but of their order. To clarify, in the analogue environment, a document may fade to the point of being unreadable, although it maintains the same content/data in the same order in which they were first affixed to the medium. In contrast, in the digital world, if the original bits are ordered, for example, 101, the value conveyed is 5, but if we change the

[7] The difference between substantive and demonstrative evidence is that the former is admitted for its content, while the latter only for the mere fact of its existence, and in support of other substantive evidence.

order to 110, the value is 6, and, if we change again to 011, the value is 3. The same bits have different value if their order changes. Thus, loss of fidelity implies different content. To prove authenticity one has to demonstrate that there has not been loss of data integrity. How can one do that? Intentional alteration is preventable through permission and access controls, but accidental alteration avoidance requires that additional hardware and/or software be in place. Both types of alteration require, in addition to methods for preventing them, methods of determining whether the record has been altered, maliciously or otherwise. For this, one cannot rely on file size, dates or other file properties, but needs audit logs and strong methods like Checksum and HASH Algorithms.

A second type of integrity digital forensics experts are concerned with is duplication integrity, that is, the fact that, given a data set, the process of creating a duplicate of the data does not modify the data either intentionally or accidentally, and the duplicate is an exact bit copy of the original data set (Mocas, 2004). This type of integrity is extremely important because one can only preserve digital records by reproducing them. However, when lawyers talk about duplication, they usually refer to making "copies", while forensic experts refer to "images" (or disk images). The difference is fundamental.

A copy is a selective duplicate of files. When we copy our digital files from one location to another, we are copying indexed files that are visible to us through our computer's file system. We cannot copy deleted files (once they are removed from the trash folder) because deletion makes them non-retrievable, invisible to the user. However, those files do still exist. Therefore copying provides an incomplete picture of the digital device. Furthermore, the action of copying changes elements of file and system metadata that may be instrumental in assessing evidence. In contrast, a disk image is a bit by bit reproduction of the storage medium, a full disk copy of all sectors of a storage device, including indexed files, deleted files and slack space, regardless of operating system or storage technology, made prior to performing any analysis of the disk. Creating a disk image is important in forensics to ensure that disk information is not inadvertently changed, to reproduce forensic test results on the original evidence, and to capture information normally invisible to the operating system when in use (including memory, page files, boot sector, BIOS). In addition, digital forensics experts link duplication integrity to time and use of time stamps for that purpose. The reason is that every time one accesses a computer something changes, thus, not two images taken at different times—even in a close sequence—are identical.[8]

Another type of integrity is computer integrity, which means that the computer produces accurate results when used and operated properly, and that it was so employed when the evidence was generated. This is similar to the concept of system integrity, which means that the system in question would perform its intended function in an unimpaired manner, free from unauthorized manipulation, whether intentional or accidental. Both imply hardware and software integrity. To be able to establish computer and system integrity one needs to verify that 1) sufficient security measures are in place to prevent unauthorized or untracked access to the computers, networks, devices, or storage, and 2) stable physical devices will maintain the value they were given until authorized to change: users/permissions, passwords, firewalls, and system logs. The latter are sets of files *automatically* created to track the actions taken, services run, or files accessed or modified, at what time, by whom and from where. They are categorized in Web logs (Client IP Address, Request Date/Time, Page Requested, HTTP Code, Bytes Sent, Browser Type, etc.), Access logs (User account ID, User IP address, File Descriptor, Actions taken upon record, Unbind record, Closed connection), Transaction logs (History of actions taken on a system to ensure Atomicity, Consistency, Isolation, Durability; Sequence number; Link to previous log; Transaction ID; Type; Updates, commits, aborts, completes), and Auditing logs. The latter are increasingly required by law to demonstrate the integrity of the system and, when properly configured and restricted, provide checks and balances, are able to determine effective security policies, to trap errors that occur, to provide instantaneous notification of events, to monitor many systems and devices through 'dashboards,' to support the determination of the accountability of people, to provide the necessary snapshot for post-event reconstruction ('black-box'), and to answer Who–What–Where–When questions, but only if retained for sufficient time.

Regardless of the elements of the computer/system that are examined to verify it, computer/system integrity can be inferred on the basis of repeatability, verifiability, objectivity and transparency. More generically, an inference of system integrity can be made if the theory, procedure or process on which the system design is based 1) has been tested or cannot be tampered with; 2) has been subjected to peer review or publication (or follows a standard); 3) it's known or potential error rate is acceptable; and 4) is generally accepted within the relevant scientific community (Mocas, 2004).

The final type of integrity is process integrity, that is, the respect of formalized legal requirements for the collection, recovery, interpretation and presentation of digital evidence. The assessment of process integrity is based on two fundamental principles, the principle of non-interference and the principle of identifiable interference. The former means that the method used to gather and analyze digital data or records does not change the digital entities; the latter means that, if the method used does alter the entities, the changes are identifiable (Mocas, 2004; Carrier, 2003a). These principles, which embody the ethical and professional stance of digital forensics experts, are consistent with the traditional impartial stance of archivists, as well as with their responsibility of neutral third party, or trusted custodians (Duranti, 2009).

For digital forensics, in the process of presenting evidence at trial, authentication is proof of authenticity by means of an authoritative declaration, but such declaration is provided by a witness who can testify about the existence and/or substance of the record on the basis of his/her familiarity with it, or, in the absence of such person, by a digital forensics expert showing that the computer process or system produces accurate results when used and operated properly and that it was so employed when the evidence was generated. In digital forensics, the strength of circumstantial digital evidence could be increased by metadata which record 1) the exact

[8] For a comprehensive comparison of the concepts and implications of digital reproduction between digital records management and digital forensics, see (Xie, 2011).

dates and times of any document sent or received; 2) which computer(s) actually created them; and 3) which computer(s) received them. Also a chain of legitimate custody (or chain of evidence, in legal terms) is ground for inferring authenticity and authenticate a record, and so is a digital chain of custody, that is, the information preserved about the record and its changes, showing that specific data was in a particular state at a given date and time. Additionally, a declaration made by an expert who bases it on the trustworthiness of the record-keeping system and of the procedures controlling it (quality assurance) is recognized as valid authentication, and so is circumstantial evidence that a system would perform its intended function in an unimpaired manner, free from unauthorized manipulation of the system, whether intentional or accidental.

## 3. Testing trustworthiness of digital documentary evidence[9]

Trust is at the root of acceptability of documentary materials as evidence at court. The common law determines the admissibility of evidence by the application of three rules: (1) the *authentication rule*; (2) the *best evidence rule*; and (3) the *hearsay rule*. The application of these rules is severely challenged by an inconsistent understanding of the nature of digital materials. In this section we will discuss the rules of evidence in the common law tradition generally, and the Canadian common law tradition specifically, as they apply to and affect admissibility of digital documentary material.

### 3.1. *The authentication rule*

According to the traditional rules for admissibility of documentary evidence, data/documents/records proffered as evidence must be capable of being shown to be authentic, that is, they must be proven to be what they purport to be. The party adducing the evidence is responsible for establishing a foundation of authenticity.[10] While this may be done in a number of different ways, typically the onus falls on the opponent to challenge the trustworthiness of the evidence and raise a reasonable doubt that the evidence is not what it purports to be. Some legal professionals have questioned whether it is possible for the opponent to raise a reasonable doubt about the authenticity of digital evidence, given the complexity of digital materials and the systems which produce and store them. A challenge to the claim of authenticity of digital material may require access to the system that generated the information to determine whether, in fact, it was operating properly at the time the evidence was generated. These professionals advocate the need for a shift in the focus of the admissibility rules from a records focus to a system focus (Peritz, 1986; Gahtan, 1999; Arkfeld, 2006; Buskirk and Liu, 2006; Paul, 2004, 2008; Chasse, 2007, 2011). In fact, the current statutes and rules of evidence have led one legal scholar to argue that there is an "authenticity crisis" (Paul, 2008), while another author contends that the judicial system may not be experiencing so much an authenticity crisis as a reliability crisis (Parry, 2009).

### 3.2. *The best evidence rule*

In Canada that shift has begun.[11] Traditional documentary evidence must adhere to the "best evidence" rule, interpreted as a requirement for the original, unless the original document/record is unavailable for accepted reasons. Digital entities pose a challenge for this traditional rule. Research has shown that the concept of original is meaningless in the digital environment (Duranti and Thibodeau, 2006; Duranti, 2005; Duranti and Preston, 2008), although one can speak of records having "the force of originals" (Paul, 2008). When the best evidence rule gained the force of law, it was to minimize the risk of admitting unreliable and inaccurate records resulting from hand copying. However, all digital duplicates are, or appear to be identical (although some of their metadata will be different). As seen earlier, reliability in the digital environment comes not from the record itself but from the integrity of the system which generates and stores it and from the controls exercised on the creation, maintenance and use of the record in such a system. In Canada, the electronic evidence provisions were drafted by the Uniform Law Conference of Canada in 1998 to address this issue. The resulting *Uniform Electronic Evidence Act (UEEA)* now incorporated into the *Canada Evidence Act* (s. 31), and many of the provincial and territorial acts, established that: (1) authentication is of the computer system, not the record; (2) the best evidence rule is satisfied by evidence showing the integrity of the system; (3) no discussion is needed of the hearsay rule or

[9] As stated earlier, the following discussion concerns the use of digital documentary evidence in common law traditions. Although civil law systems of trial rely heavily on documents and documentary evidence, and require that those documents be authentic, it is in the common law systems, based on the foundational belief that the trustworthiness of evidence can best be determined by testing the evidence, that digital documentary evidence is now generating much discussion. Trustworthiness of evidence at common law is accomplished traditionally through oral testimony and cross-examination of live witnesses, and by means of an intricate set of statutory and common law rules, developed over centuries, which governs the use of real and documentary evidence. The separation between the trier of law and trier of fact (although they may indeed be the same person) guides discussion of admissibility of evidence (the responsibility of the trier of law) and the weight of that evidence (the degree of probative value, or credibility of proof – the responsibility of the trier of fact). For a comparison of common law and civil law traditions, see (Paciocco, 2010), "Understanding the Accusatorial System," Canadian Criminal Law Review 14 pp. 307–325.

[10] Although according to George Paul, "If we are to be intellectually honest, there is almost no preliminary burden of proving digital information is authentic." (Paul, 2008, p. 49).

[11] In 2010 these authors conducted a research project: The Canadian legal framework for evidence and the Digital Economy: A disjunction? Principal Investigator: Anthony F. Sheppard, Professor of Law, UBC; Co-Investigator: Dr. Luciana Duranti, Professor of Library, Archival and Information Studies (SLAIS), UBC; researchers, Corinne Rogers and Donald Force, PhD students, SLAIS, UBC. Funded by the Social Sciences and Humanities Research Council (SSHRC), Knowledge Synthesis Grants on the Digital Economy.

its exceptions[12] for computer records; (4) no discussion of weight is needed. These electronic records provisions stipulate that systems integrity be the standard by which the best evidence rule is superseded for digital evidence.[13] However, while the traditional best evidence rule seems to be inapplicable in the digital environment, its intent needs to be captured and expressed in rules aiming to achieve functional equivalence (Chasse, 2007; Duranti and Endicott-Popovsky, 2010; Sheppard and Duranti, 2010).

### 3.3. *The hearsay rule*

Traditionally, at common law, documents are considered to be hearsay because they can only 'say' what somebody else 'told them', and are not admissible as evidence as they cannot be cross-examined. The Supreme Court of Canada has defined hearsay as "a statement offered in evidence to prove the truth of the matter asserted within it, but made otherwise than in testimony at the proceeding in which it is offered" (*R. v. O'Brien*, [1978] 1 S.C.R. 591at 593-94, (1977), 35 C.C.C. (2d) 209 at 211 S.C.C.). Whether digital documents are different from paper documents in this regard is a matter of dispute. Some prosecutors in the U.S. have argued that the *hearsay rule* applies only to human declarants, while most federal courts have considered computer reports as hearsay. Others distinguish between computer-stored materials, which may be considered hearsay on the grounds that they contain human statements, and computer-generated digital objects, which are not considered hearsay, because their content does not result from human intervention. In *State v. Kandutsch*, 756 N.W.2d 811 2011 WL 2820791, the court rejected the prosecution's argument that the digital data of an electronic monitoring device were inadmissible because they constituted hearsay evidence (Brenner, 2011). The relevance of or purpose for which a digital entity is being offered into evidence may also affect its classification as hearsay or non-hearsay.

Hearsay becomes admissible in a court of law if it qualifies for an exception to the hearsay rule. One such exception is the statutory business records exception, which considers documents to be admissible because inherently reliable if they qualify as business records. Business records are defined as documents generated in the usual and ordinary course of business by an individual who had a duty to make them and did so at or near the time of the documented event or transaction. Many attorneys assume that digital documents will meet these criteria, but it is not necessarily so (Fosmire, 2006). The business records exception to the hearsay rule considers the reliability of business records from the perspective of traditional paper recordkeeping practice, and examines the records themselves and the circumstances of their creation. When the rule was first applied to digital records, their management meant "batch processing" in a mainframe computer. Thus, the business records provisions reflect the technology of the time of their enactment – computer technology still serving traditional concepts of business records, essentially the equivalent of paper records accelerated in their application by mainframe computers. Digital records management today is based on concepts relating to the information systems in which documents reside – the reliability of business records depends also on the reliability of the systems that produce and maintain them. Records managers (and archivists) understand and work daily with these distinctions but few lawyers have the same opportunity. Therefore judges are not presented with the evidence or the arguments that would enable them to use the law of evidence more compatibly with digital technology (Chasse, 2007, 2010).

The discussion of the rules of admissibility at common law has touched upon some of the ways in which the nature of electronic records and digital technologies is challenging traditional rules of evidence and procedure. The traditional best evidence rule is no longer relevant because of the absence of originals in the digital environment. The authentication rule also is inadequate, because it cannot be established that an electronic record is the same as its first instantiation simply by looking at the record itself, but it is necessary to refer either to an unbroken line of traces left by all those who interacted with the record or to the legitimate custody of a professional who can account for them (MacNeil, 2000; Duranti and Thibodeau, 2006; Duranti, 2009). Furthermore, the complexity and variety of digital information systems and the often uncontrolled ways in which they are used makes it difficult to identify records within them and the business activities to which they are linked, thereby challenging the application of the business records exception to the hearsay rule. Finally, ever-changing technology speeds up the obsolescence not only of earlier record-making processes, but also of the laws regulating admissibility.

Ken Chasse poses a radical question – are the traditional best evidence, authentication and hearsay rules necessary for admitting digital evidence? He concludes that they are not. Furthermore, system integrity bridges the gap between legal and records management rules, and so the call for "system integrity" should require compliance of electronic record systems with recognized standards of records management (Chasse, 2007, 2010, 2011). If such questions are still being debated with respect to digital records that can be considered to be traditional computerized records contained in in-house recordkeeping systems, the problems are compounded by the increasing adoption of virtualization and cloud technologies.

## 4. Trust in the cloud

Having discussed the characteristics of digital records, the related trust framework established by archival science and digital forensics, and the challenges encountered by the existing legal system in common law countries in establishing

[12] Documents presented by litigants for the truth of their contents are considered to be hearsay, that is, a statement "or communicative conduct made by persons otherwise than in testimony at the proceedings in which it is offered" (Bryant et al., 2009). Hearsay is not admissible unless it falls under an exception to the hearsay rule. The most common exception is the business records exception, codified in section 30 of the Canada Evidence Act (Government of Canada, 2008).

[13] Their effectiveness has been questioned however. Despite their passage into statutory law twelve years ago, there have yet to be judicial decisions providing analysis of their key phrases, such as "the integrity of the electronic documents system." Chasse (2011) para. 11. See also Duranti et al. (2010).

the trustworthiness of digital evidence, we can now return to the aspect of digital technology that was discussed in the beginning as creating both excitement and anxiety for businesses and individuals, and new headaches for the legal profession, namely, cloud computing. Individuals and organizations, large and small, are drawn increasingly by the lure of cloud computing for the many benefits it offers. Scalable, agile, efficient, on-demand computing resources mean that email, photos, documents and records can be easily stored and shared through a seemingly endless number of hosted web applications, and that sophisticated software, platforms, and infrastructure are available to the budget-conscious and technology-resource limited. Cloud architectures offer on-demand access to services across a network of standard internet-accessible devices – mobile phones, tablets, laptops – and a vast array of other devices, such as game consoles, MP3 players, and e-business technologies. Resources are shared among users, and resource use is monitored and invoiced based on usage for service. We use – and increasingly rely on – cloud services for communication (email is the number one use), backup and storage, collaboration, distribution, recordkeeping and preservation. But for every benefit there is a corresponding risk that may or may not be recognized.

Cloud computing is defined by the National Institute of Standards and Technology as "a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" (Jansen and Grance, 2011). According to a study in the U.S. from 2008, 69% of online Americans have used at least one web-based, or cloud, service. Four years is an eternity in respect of technology adoption – these results must now be considered conservative. A global study released in March 2011, reported that Canadians average the most time on the Internet of any national group – a staggering 44 h per week (Denham, 2011). This trend "is fuelling a mass migration of information, once stored on the hard drives of personal computers, to remote servers in a domain controlled by online service providers" (Nied, 2011).

The model of cloud computing is reminiscent of the mainframe environment of the 1960s, except that in this case we are not putting our trust in the proprietary and highly controlled environment of the company mainframe, but in a global service provider like Amazon or Google, whose agendas and priorities as they build out their infrastructures are very different from our own. The trust relationship demands careful analysis and consideration.

There are four standard deployment models for cloud architecture that broadly characterize the management and disposition of computational resources for service delivery. Each has corresponding benefits and risks to be analyzed in the context of trust requirements. A private cloud infrastructure is operated for a single organization, that is, data in a private cloud does not share resources with data belonging to other individuals or organizations. A private cloud may be managed by the organization or by a third party, and may be hosted within the organization's IT infrastructure, or externally. Public cloud infrastructure is made available to the general public over the Internet. By definition external to the customers' organization, public clouds are owned and operated by third-party providers and usage is subject to detailed service level agreements. Between these two extremes are community clouds and hybrid clouds. A community cloud infrastructure is shared by two or more organizations with common privacy, security, and regulatory considerations. It may be managed by the organizations or a third party, and may be hosted internally or externally. The most complex is the hybrid cloud, composed of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (Jansen and Grance, 2011).

Clouds conform to one of three service models, which dictate an organization's scope and control over the computational environment. These service models can be actualized in each deployment model.

- Software as a Service (SaaS) offers the consumer on-demand access to one or more applications and the computational resources to run them. The cloud provider carries out management, control, and security of network, servers, operating systems, applications, and storage.
- Platform as a Service (PaaS) offers the consumer on-demand access to the computing platform upon which applications can be developed and deployed. The consumer controls applications and environment settings, and security is split between the cloud provider and cloud consumer.
- Infrastructure as a Service (IaaS) offers the consumer on-demand access to the basic computing infrastructure of servers, software, and network equipment. The consumer does not manage or control the underlying cloud hardware and software infrastructure components, but has broad freedom and control over operating systems, storage, deployed applications, and some networking components (e.g. host firewalls). Security of consumer-chosen elements is carried out mainly by the consumer (Jansen and Grance, 2011).

The theoretical trust framework developed in the previous section can be applied to highlight specific challenges to trusting data, information, and records to the cloud. Key issues of ownership, jurisdiction, and privacy have yet to be resolved. Longer term concerns around responsibility for maintenance, access, and preservation, all of which correspond to issues of trust, are looming on the horizon. The following list identifies some concerns but is by no means exhaustive:

- The servers in which data and records are stored may be, but likely are not, in the same country or jurisdiction in which they were created. In the event of litigation or other dispute, in what jurisdiction will they be governed?
- Do you even know where your data is stored? As the cloud storage market continues to grow, this becomes increasingly unclear. New storage providers are appearing who aggregate unused storage from third parties. The entrance of a peer-to-peer model for storage adds further complexity to teasing out the tangled web of provenance, custody, control, and legal responsibility (Darrow, 2012).

- Will trade secrets, if entrusted to cloud storage, remain secrets? Having already been shared with a third party, can they still be considered secret?
- How will cloud service providers protect content from data breaches? There is a school of thought that says you should be concerned not about *if* a data breach occurs, but *when* it occurs. How will your cloud service provider handle a breach? Will your provider even admit to a breach?
- What happens to content if a cloud service provider goes offline (this could be due to bankruptcy or criminal investigation), or if the server containing your records is sequestered for an investigation? Even if you can recover your content, can you then be assured of its trusted chain of custody? How do you even prove an unbroken chain of custody?

Returning then to the premise of this article (section 1), if trust in records rests on four types of knowledge about the records' custodian – namely reputation, performance, competence, and confidence, we must ask hard questions of the providers to whom we entrust our records and data. International research projects into the nature of digital records have developed guidelines and solutions to managing authenticity, accuracy and reliability in digital records systems,[14] but solutions are often out of reach financially for many organizations driven by the bottom line. National and international standards of records and information management provide guidance but adherence is not legally required in most sectors. Cloud computing offers to ease the financial burden of many aspects of records management, but in the process raises a host of new and troubling questions that must be answered if we are to be able to trust our documentary output. Technology will not stand still to wait for our legal and regulatory system to catch up.

As Leslie Johnston has stated in the Library of Congress blog, Signal:

> *We can't be afraid of cloud computing. Given the volumes of data coming our way and mounting researcher demands for access to vast quantities of data, the cloud is the only feasible mechanism for storing and providing access to the materials that will come our way. We need to focus on developing authentication, preservation and other tools that enable us to keep records in the cloud* (Johnston, 2011).

In other words, when it comes to digital records, trust is not all!

**Luciana Duranti** (luciana.duranti@ubc.ca), *Professor, University of British Columbia, School of Library, Archives, and Information Studies, Irving K. Barber Learning Centre, Suite 470-1961 East Mall, Vancouver, BC V6T 1Z1, +1 604 822 2587.*

**Corinne Rogers** (cmrogers@interchange.ubc.ca), *PhD candidate, University of British Columbia, School of Library, Archives, and Information Studies, Irving K. Barber Learning Centre, Suite 470-1961 East Mall, Vancouver, BC V6T 1Z1, +1 604 929 0243.*

## REFERENCES

Arkfeld MR. Electronic discovery and evidence. Phoenix, AR: Law Partner Publishing, LLC; 2006.

Bok S. Lying: moral choice in public and private life. New York: Vintage Books; 1999.

Borland J. Trusting archivists. Archivi and Computer 2009;XIX(1): 94–106.

Brenner S. CYB3RCRIM3: expert testimony, hearsay and the electronic monitoring device. CYB3RCRIM3. Available at: http://cyb3rcrim3.blogspot.com/2011/07/expert-testimony-hearsay-and-electronic.html; 2011 [accessed 30.07.11].

Bryant AW, Lederman SN, Fuerst MK, Sopinka J. The law of evidence in Canada, Markham, Ont. Dayton, Ohio: LexisNexis; 2009.

Buskirk E van, Liu VT. Digital evidence: challenging the presumption of reliability. Journal of Digital Forensic Practice 2006;1:19–26.

Carrier B. Defining digital forensic examination and analysis tools using abstraction layers. International Journal of Digital Evidence 2003a;1(4):1–12.

Carrier B. Open source digital forensics tools: the legal argument. Available at: www.digital-evidence.org/papers/opensrc_legal.pdf; 2003b.

Chasse KL. Electronic evidence; 2010.

Chasse KL. Electronic records as documentary evidence. Canadian Journal of Law and Technology 2007:141–62.

Chasse KL. The inadequacy of analysis of electronic records management for evidence and discovery. Asia Law Info. Available at: http://article.chinalawinfo.com/Article_Detail.asp?ArticleID=59257; 2011 [accessed 17.04.12].

Darrow B. "Skype of cloud storage" symform nets $11M in funding. GigaOM. Available at: http://gigaom.com/cloud/skype-of-cloud-storage-symform-nets-11m-in-funding/; 2012 [accessed 26.04.12].

Denham E. Finding our way through the clouds. Available at: http://www.oipc.bc.ca/pdfs/Speeches/FindingOurWayThroughTheClouds.pdf; 2011 [accessed 22.04.12].

Duranti L. From digital diplomatics to digital records forensics. Archivaria 2009;68(Fall):39–66.

Duranti L. The long-term preservation of authentic electronic records: findings of the InterPARES project. San Miniato: Archilab; 2005.

Duranti L, Endicott-Popovsky B. Digital records forensics: a new science and academic program for forensic readiness. Journal of Digital Forensics, Security and Law 2010;5(2):1–12. Available at: http://www.jdfsl.org/subscriptions/JDFSL-V5N2-Duranti.pdf [accessed 01.03.11].

[14] Two such projects are the International Research on Permanent Authentic Records in Electronic Systems (InterPARES), a three-phase, thirteen year project just completed at the University of British Columbia, directed by Luciana Duranti, and engaging research teams from more than 25 countries (www.interpares.org), and the Digital Records Forensics project (www.digitalrecordsforensics.org), also led by Dr. Duranti. The InterPARES project developed the theoretical and methodological knowledge essential to the long-term preservation of authentic records created and/or maintained in digital form. The Digital Records Forensics project (2008–2011) was a collaboration between the UBC Faculty of Law, the School of Library, Archives, and Information Studies (SLAIS), and the Vancouver Police Department researching the identification of records among all the digital objects produced by complex systems, and the determination of their authenticity. This project has resulted in a new stream of graduate study at SLAIS, in collaboration with the University of Washington, that combines archival theory, information assurance, and digital forensics (Duranti and Rogers, 2011; Duranti and Endicott-Popovsky, 2010).

Duranti L, Preston R. Research on permanent authentic records in electronic systems (InterPARES) 2: experiential. Interactive and dynamic records. Padova: Associazione Nazionale Archivistica Italiana; 2008.

Duranti L, Rogers C. Educating for trust. Archival Science 2011; 11(3–4):373–90. Available at: http://www.springerlink.com/index/10.1007/s10502-011-9152-3 [accessed 16.04.12].

Duranti L, Rogers C, Sheppard AF. Electronic records and the law of evidence in Canada: the uniform electronic evidence act twelve years later. Archivaria 2010;70(Fall):95–124.

Duranti L, Thibodeau K. The concept of record in interactive, experiential and dynamic environments: the view of InterPARES. Archival Science 2006;6(1):13–68.

Fosmire MS. Refining the standard: authenticating computer-based evidence. Law and Technology Resources for Legal Professionals. Available at: http://www.llrx.com/features/computerbasedevidence.htm; 2006 [accessed 13.11.10].

Gahtan AM. Electronic evidence. Scarborough, ON: Carswell; 1999.

Google. Privacy policy. Google|policies & principles. Available at: http://www.google.com/intl/en/policies/privacy/; 2012 [accessed 25.04.12].

Government of Canada, L.S.B.. Consolidated federal laws of Canada, Canada evidence act. Available at: http://laws-lois.justice.gc.ca/eng/acts/C-5/page-10.html#docCont; 2008 [accessed 26.04.12].

InterPARES. InterPARES 3 project: glossary. Available at: http://www.interpares.org/ip3/ip3_terminology_db.cfm?letter=p&term=38; 2011 [accessed 09.05.11].

Jansen W, Grance T. Guidelines on security and privacy in public cloud computing. Available at: http://www.nist.gov/customcf/get_pdf.cfm?pub_id=909494; 2011 [accessed 22.04.12].

Johnston L. From records to data: it's not just about collections any more. The Signal: Digital Preservation. Available at: http://blogs.loc.gov/digitalpreservation/2011/11/from-records-to-data-it%E2%80%99s-not-just-about-collections-any-more/; 2011 [accessed 25.04.12].

Kenneally E. Gatekeeping out of the box: open source software as a mechanism to assess reliability for digital evidence. Virginia Journal of Law and Technology 2001;13(Fall). Available at: www.vjolt.net/vol6/issue3/v6i3-a13-Kenneally.html.

MacNeil H. Providing grounds for trust: developing conceptual requirements for the long-term preservation of authentic electronic records. Archivaria 2000;50:52–78.

Maes J. SOPA, PIPA, Megaupload.com, and the United States government. The Washington Times. Available at: http://communities.washingtontimes.com/neighborhood/political-potpourri/2012/feb/3/sopa-pipa-megauploadcom-and-united-states-governme/; 2012 [accessed 25.04.12].

Mocas S. Building theoretical underpinnings for digital forensics research. Digital Investigation 2004;1(1):61–8. Available at: http://www.sciencedirect.com/science/article/B7CW4-4BMXXJS-C/2/9154d2932943f309d86f8a748ac40ab3.

Nied M. The internet, cloud computing, and the charter right to privacy: the effect of terms of service agreements on reasonable expectations of privacy in criminal cases. Internet and E-Commerce Law in Canada 2011;12(5). Available at: http://defamationlawblog.files.wordpress.com/2009/07/internet-cloud-computing-and-charter-right-to-privacy-matthew-nied.pdf [accessed 25.04.12].

Office of the Information Commissioner of Canada. A dire diagnosis for access to information in Canada. Available at: http://www.oic-ci.gc.ca/eng/med-roo-sal-med_spe-dis_2009_4.aspx; 2009 [accessed 23.04.12].

Paciocco DM. Understanding the accusatorial system. Canadian Criminal Law Review 2010;14(3):307–25.

Parry ZB. Digital manipulation and photographic evidence: defrauding the courts one thousand words at a time. Journal of Law, Technology & Policy 2009;(1):175–202. Available at: http://www.jltp.uiuc.edu/archives/Parry.pdf; 2009.

Paul GL. Foundations of digital evidence. Chicago, IL: American Bar Association; 2008.

Paul GL. The "authenticity crisis" in real evidence. The Practical Litigator 2004;15(6):45–52.

Peritz RJ. Computer data and reliability: a call for authentication of business records under the federal rules of evidence. Northwestern University Law Review 1986;80: 956–1002.

Sheppard AF, Duranti L. The Canadian legal framework for evidence and the digital economy: a disjunction? University of British Columbia; 2010.

Sztompka P. Trust a sociological theory. Cambridge, UK; New York, NY: Cambridge University Press; 1999.

Twitter Blog. Tweets still must flow. Twitter Blog. Available at: http://blog.twitter.com/2012/01/tweets-still-must-flow.html; 2012 [accessed 25.04.12].

Xie S. Building foundations for digital records forensics: a comparative study of the concept of reproduction in digital records management and digital forensics. American Archivist 2011;74(2):576–99. Available at: http://archivists.metapress.com/content/E088666710692T3K [accessed 25.04.12].