Authenticity of Digital Records: An Archival Diplomatics Framework for Digital Forensics

Luciana Duranti, Adam Jansen,
University of British Columbia, Vancouver, Canada
luciana.duranti@ubc.ca
adam@dkives.com

Abstract:
The field of forensic analysis of digital records is undergoing a rapid maturation process in order to maintain pace with the change of technology. The evolving best practices in digital forensic investigation that concern collection, handling, examination and presentation of evidence are predominantly practitioner driven, with little research directed towards the issues surrounding the maintenance and long term preservation of the authenticity of the evidence collected. Yet, the authenticity of such evidence is essential to the judicial process for, without authentic evidence, determination of guilt or innocence cannot be certain. As evidence is increasingly in digital form, higher standards of conduct are demanded of those examining and preserving it. An interdisciplinary approach to the digital forensic processes involving the use of archival diplomatics can assist forensic experts in the determination of the authenticity of digital evidence that has the nature of records, as well as the identification of key attributes that must be preserved over the long term to demonstrate and maintain the records continuing authenticity. The scientific rigor of archival diplomatics would allow forensic experts and the custodians of collected evidence to demonstrate, over the long term and through system and technical migrations, the continuing authenticity of both the records they have collected from a crime scene as evidence and the records they have produced during the examination, submission and maintenance of such evidence. This paper discusses how the concepts of archival diplomatics, digital forensics and the law of evidence can be integrated to form the new discipline of Digital Records Forensics.

Keywords: Digital Forensics, Digital Preservation, Archival Diplomatics, Records, Authenticity

Introduction

> "Without the rigorous process that leads to proven scientific discovery, decision-makers in the courts and elsewhere are left to rely on supposition or worse yet intuition in the pursuit of justice." (DFRWS, 2001)

With increased use of computer based systems in the execution of business, both public and private, the need to rely upon computer-generated and stored materials as evidence has increased, as demonstrated by studies conducted by International Data Corporation (IDC) which have predicted a 138 percent compounded annual growth rate (CAGR) in the use of email from 2001-2010 (CNN, 2001), as well as a study out of the University of California-Berkeley, which found an 87% increase in digital information from 1999-2003 (UC-Berkeley, 2003). This exponential growth has resulted in a rapid maturation of the field of digital forensics to address the immediate needs of forensics investigators. However, precisely because of the reasons driving such growth, a majority of the published research has been on the practitioner driven processes of collection, handling, analysis and presentation of evidence, with little attention paid to the longer-term issues of the maintenance and preservation of evidence and to the theory, principles, and methodological concepts on which these processes should be based. The need for a comprehensive understanding of what record authenticity means in the digital environment and in relation to long term preservation is becoming increasingly important as cases involving digital evidence are becoming larger and more complex and the time between initial incident and final resolution of court cases is taking longer, often lasting a decade or more (Jansen, 2010).

The lack of a rigorous, scientific approach to forensic analysis based on theory, along with the absence of consistent terminology has been highlighted by several of the leading judicial and forensic experts. In testimony before the US Senate, Chief Judge Emeritus Harry Edwards provided testimony on the "paucity of interdisciplinary scientific research to support forensic disciplines" and "the absence of solid scientific and applied research focused on new technology and innovation" (Edwards, 2009: 2). The first Digital Forensics Research Workshop (DRFWS) emphasized the need to "establish a

research community that would apply the scientific method in finding focused near-term solutions driven by practitioner requirements and addressing longer term needs" (DRFWS, 2001: iii). Sujeet Shenoi, in his opening address to the 2011 International Federation for Information Processing (IFIP) Working Group 11.9 on Digital Forensics conference, challenged the participants to conduct and publish more peer-reviewed research in order to advance the field. Already Eoghan Casey had raised concerns about digital forensics being "a nebulous discipline that lacks standards or even clear definitions of fundamental terminology" (Casey, 2004).

In the United States, adding to this need for a consistent theoretical approach to the forensic processes is the increased application of the 'Daubert rules' to determine the admissibility of expert testimony (Kenneally, 2001). These rules are four general guidelines established in *Daubert v. Merrell Dow Pharmaceuticals, 509 US 579* (1993) to assist judges in determining whether the methodologies and techniques employed to identify scientific evidence were sound and widely accepted. They essentially stated that an **inference** of a valid process could be made if:

- the theory, procedure or process has been tested and the test is valid and repeatable
- it has been subjected to peer review or publication (standard)
- its known or potential error rate is acceptable
- it is generally accepted within the relevant scientific community

:

These guidelines are of particular concern to the forensics field in the US as *Kumho Tire Company v. Carmichael 526 U.S. 137* (1999) extended them to include expert testimony in other technical and specialized fields (Meyers, 2004). Application of the concepts introduced by the Daubert rules is however spreading well beyond the jurisdiction in which they were first articulated and requires the development of an overarching intellectual framework within which all forensic processes can be situated and articulated as an internally consistent system of ideas and practices. This requires interdisciplinary research in areas of established knowledge capable of supporting the purposes of forensic activity.

The Digital Records Forensics Project

The Digital Records Forensics (DRF) project at the University of British Columbia was developed in 2008 to begin addressing the need for an interdisciplinary body of theory and methods supporting the digital forensics processes. The DRF project is a collaboration among the University of British Columbia (UBC) School of Library, Archival and Information Studies, the UBC Faculty of Law and the Vancouver Police Department. Leveraging on previous research into the long term preservation of the authenticity of digital records conducted by the International research on Permanent Authentic Records in Electronic Systems (InterPARES) project, the DRF project is integrating the bodies of knowledge of archival diplomatics, digital forensics and the law of evidence to develop the new scientific discipline of Digital Records Forensics.

The very first obstacle the DRF project had to deal with was a reconciliation of the terminology that is at the core of the project itself. For example, while all disciplines involved in the research deal with "preservation", they define it quite differently. Most digital forensic models identify preservation as one of the steps in the forensic process. Carrier's model refers to preservation as isolation and collection (Carrier, 2003). Reith, Carr, and Gunsch define the preservation activity as to "isolate, secure and preserve the state of physical and digital evidence," including "preventing people from using the digital device or allowing other electromagnetic devices to be used within an affected radius" (Reith et al., 2002: 4). DFRWS recognizes the need for preservation, but does not provide any significant detail on what preservation means with regards to digital evidence; rather, its model describes the relevant issues as case management, imaging technologies, chain of custody and time synchronization. The archival view of digital preservation is quite different, as it involves the whole of the principles, policies, strategies and activities designed to ensure the physical and technological stabilization of records for the purpose of extending indefinitely their life and protecting the accuracy and authenticity of and maintaining the accessibility to their intellectual content. Thus, preservation includes, among many other activities, *description*, which provides an account of the context, attributes, and relationships of the documents, and the development and maintenance of *retrieval systems*. The continuing authenticity of records has relied for centuries on the concept of an unbroken chain of custody. However, as the long term storage and retrieval of digital objects require both "physical and representational transformations, the traditional concept of an unbroken chain of custody needs to be expanded to encompass the processes that are necessary to ensure that an electronic record is

transmitted over time without inappropriate alteration" (Preservation Task Force, 2001: 8). Integrating the archival understanding of preservation into the forensic models will bring a more sophisticated understanding of what is required to effectively preserve evidence over the long term, an understanding that heavily relies on the archival diplomatics concept of authenticity.

Traditional Diplomatics

The first scientific method for establishing the authenticity of records of questionable origin was developed by the French Benedictine monk Dom Jean Mabillon, who in 1681 wrote *De re Diplomatica libri VI,* which analyzed medieval charters in relation to others of similar type looking for deviations from the routines used by the entities they claimed to be authored by. Mabillon penned his masterwork in response to a challenge by the Jesuit Bollandist Daniel van Papenbrock questioning the authenticity of the charters on which the French Abbey of Saint Denis based its patrimonial rights (Skemer, 1989). However, the scientific validity of the methodology he developed and the theoretical framework that he built for it were recognized by van Papenbrock and were used for centuries in judicial disputes about the authenticity of documentary evidence, which were called "Bella Diplomatica" or "wars about records", from the title of Mabillon's book. From the second half of the 18th century, the concepts and methods of diplomatics were taught in the faculties of law of all major European universities and became the foundation of the law of evidence as we know it today.

Over the centuries, as the introduction of new technologies (e.g. the printing press, photocopiers) expanded the creation and use of records, the body of knowledge of diplomatics continued to evolve to accommodate the development of new record types, records creation processes, etc. while continuing to maintain the ability of its concepts and methods to determine the capacity of a record to stand as evidence of the action in the course which it was created. The postulate on which diplomatics is based is that "regardless of differences in nature, provenance or date, from a formal point of view all records are similar enough to make it possible to conceive of one typical, ideal documentary form containing all possible elements of a record" (Duranti and Thibodeau, 2006: 16). In the course of the InterPARES project, researchers proved that the process of decomposing records into their constituent components and analyzing each component independently of the record's content in order to determine the authenticity of the whole is as applicable to today's digital records as it was to medieval documents. However, as contemporary records are much more dependent on their context than their medieval counterparts, the theory of archival science was integrated by InterPARES researchers into diplomatic theory, and this resulted in a new discipline that was called archival diplomatics.

When a Digital Entity is a Record

Archival diplomatics is the integration of traditional diplomatic concepts and methods with modern archival theory "based on jurisprudence, the history and theory of administration, and an extensive and centuries old body of written reflection and experience about the nature of records and record-keeping practices in bureaucratic organizations" (MacNeil, 2004: 205). It is defined as "the discipline which studies the genesis, forms, and transmission of archival documents [or records], and their relationship with the facts represented in them and with their creator, in order to identify, evaluate, and communicate their true nature "(Duranti, 1989a:17). Like diplomatics, archival diplomatics emphasizes the importance of identifying records among other kinds of information, because they are considered, in civil law countries, the "perfect" form of proof of actions and transactions, and, in common law countries, an exception to the hearsay rule, which precludes admissibility of any other kind of documents as evidence. Archival diplomatics has been refined over the course of the InterPARES research project (1998-2012) and continues to evolve with the DRF project. One of the key outcomes of this process of refinement was the breaking down of the 'ideal' digital record into its constituent parts, all considered requirements for the existence of a record:

1) a stable content and a fixed form, meaning that the entity's content must be stored so that it remains complete and unaltered, and its message can be rendered with the same documentary form or presentation it had when first set aside;

2) explicit linkages to other records within or outside the digital system, through a classification code or other unique identifier;

3) an identifiable legal-administrative, provenancial, and procedural context;

4) an identifiable author (i.e. the person or organization issuing the record), addressee (i.e., the person or organization for whom the record is intended), and writer (i.e. the person responsible for the articulation of content);

5) an action, in which the record participates or which the record supports either procedurally or as part of the decision making process ; and

6) a medium, that is a support or carrier to which the record is affixed (Duranti and Thibodeau 2006, MacNeil, 2000).

On the basis of the diplomatics fundamental assumption that the elements of a record form or documentary presentation (e.g. the elements of a letter, a memo, a report, a GIS) reveal the presence of its necessary components listed above and their meaning, the InterPARES project developed a *Diplomatic Template of Analysis* (IP2, 2008) that identifies all the possible elements of a record's external and internal form and allows for a quick and accurate identification of written information as a record

Documentary Form and Authenticity

Archival diplomatic theory states that when a record is proven authentic, one can presume that its integrity has been maintained (MacNeil, 2004). The opposite however is not true. While authenticity implies integrity, integrity does not imply authenticity – one can preserve the integrity of a forgery, yet doing so does not make of it an authentic record. Therefore, there exists another characteristic of a record that, when combined with integrity, allows for the validation of its authenticity: identity. The identity of a digital record comprises the whole of the distinguishing attributes that together uniquely characterize it and differentiate it from other records. In order to prove the authenticity of a digital record, it is necessary to establish its continuing *identity* and demonstrate its *integrity*. Identity allows the examiner to differentiate one record from another; to determine the who, what, where, when, why, and to establish the degree of perfection of the record, that is, whether it is a draft, an original or a copy and, if the latter, what type of copy and what version of it. From a diplomatic perspective, the identity of a record is revealed by its documentary form or presentation.

The documentary form of a record is the expression of a set of rules of representation that are meant to convey a message and is composed of both extrinsic elements and intrinsic elements. The extrinsic elements define the appearance of the record; while the intrinsic elements articulate its content by providing its structure. Together, they reveal its legal, administrative, provenancial, procedural and documentary context. From a conceptual viewpoint, "intrinsic elements of form are those which make a document complete, and extrinsic elements are those which make it perfect, that is, capable of accomplishing its purpose," (Duranti, 1991: 6).

The InterPARES *Diplomatic Template for Analysis* identified among the extrinsic elements of digital records presentation features (i.e., general characteristics, like text, graphics, images or sound, and specific characteristics, such as fonts, colors, hyperlinks, layouts, and resolution/scale/sample rate), electronic signatures, digital time stamps, and special signs (identifiers denoting the persons involved in the origination or execution of the record, like digital watermarks and logos) (IP1, 2000). Intrinsic elements establishing the identity of digital records include date, superscription (i.e., the mention of the name of the author in the upper part of the record, as in an e-mail header), inscription (i.e., the mention of the name of the addressee), subject, salutation, preamble, subscription, etc.

It is not the purpose of this paper to sum up the concepts of archival diplomatics but to give a sense of the level of detail diplomatic analysis relies upon, which is not limited to the form of the record, but concerns the persons (or actors, being participant in the action that the record puts into being or participate in), the action to which the record relates, by putting it into being (i.e. dispositive record, like a contract), proving it after it occurred (i.e. probative record, like a certificate), supporting it with information (i.e. supporting record, like a GIS), discussing it (i.e. narrative record, like a report), guiding it (i.e. instructive record, like a score or a digital form), or enabling it (i.e. enabling record, like a digital workflow); the procedure or process in which the record takes part, the behaviour of different types of records (e.g. dynamic, interactive, experiential) and the different types of trustworthiness (i.e. reliability, accuracy and authenticity) a record needs to be tested for (Duranti and Thibodeau, 2006; Duranti, 2007; Duranti, 2009, Duranti, 2010).

To assist in the establishment of the identity of a record, InterPARES further developed the concept of record profiles, which had been previously introduced in the first iteration of the DoD5015.2 standard by Duranti and Eastwood (Duranti, Eastwood and MacNeil, 2002). Record profiles document the identity of digital records by linking context and form in an inextricable way,

Authenticity and Preservation

Custodians can only preserve records as trustworthy (i.e. reliable, accurate and authentic) as they are when first created. It is therefore the custodian's responsibility to establish the identity of the records prior to acquiring them and to maintain that identity, together with their integrity, afterwards (MacNeil, 2004). In the digital environment, this is a tall order, because it is not possible to preserve digital records; it is only possible to preserve the ability to reproduce them (Duranti and Thibodeau, 2006). As it will always be necessary to retrieve the binary bits and process those bits through the use of intermediaries (i.e. hardware and software) in order to render the evidence into a human readable format, it falls upon the custodian to ensure that the necessary intermediaries will exist when needed. To render representations with an accuracy that is able to withstand a diplomatic analysis requires the custodian to store the binary content of the record, including indicators of all the elements of documentary form necessary to convey the essence of the record, in a manner that ensures the record will be rendered with the same presentation and in the same context that gave it meaning.

Because of technological obsolescence, the conversion and migration processes required to preserve the records will undeniably alter the underlying bits that 'make' the record, but here the force of the archival diplomatics theory comes into play. In fact, such theory allows for the authenticity of the record to become an inference that one draws from the documentation that is maintained by the custodian about its components, characteristics, creation, maintenance and preservation. This inference would be supported by the identity and integrity metadata included in the record profile that would be an integral part of any record that is presumed authentic.

Conclusion

In the more than three hundred years since its origin, diplomatics has been widely used by European archivists as a dependable, tested methodology for determining the authenticity of a record. Over the past ten years, InterPARES has extended diplomatic theory, principles and methods to digital records, integrating them with archival science to form archival diplomatics in order to establish the means of preserving the authenticity of digital records over the long term. More recently, the DRF project has tested the concepts of archival diplomatics on the digital forensic process, bringing an interdisciplinary, academic research-based rigour to the practice of digital forensics, thereby addressing some of the concerns raised by Edwards, Casey, and Shenoi. The research conducted by InterPARES and DRF should support digital forensic work with regard to at least two of the four Daubert guidelines, by providing a theoretical framework that has been tested and has a wide-spread acceptance within the scientific community, and meets international standards. Further research and implementation of archival diplomatic concepts, principles and methods within the computer forensics domain will be necessary not only to meet all the requirements of the Daubert test but also, and more importantly, to ensure that authentic digital evidence is preserved for as long as needed with its identity and integrity intact.

The body of knowledge built by the DRF project through the integration of archival diplomatics, digital forensics and the legal principles underling the law of evidence is being articulated in a number of courses at the University of British Columbia School of Library, Archival and Information Studies (SLAIS) as part of a newly proposed stream within the Master of Archival Studies program. The proposed courses will cover the theoretical and methodological content of a new discipline called "digital records forensics" and will contribute to the formation of a new kind of archival and forensic professional, able to rely on an established body of theory to continuously develop new practices capable of dealing with new and emerging technologies and permanently preserve the authenticity of the records that they generate.

References

Casey, E. (2004) “The Need for Knowledge Sharing and Standardization”, *Digital* Investigation, No. 1, pp.1-2.

Carrier, B. (2003) “Getting Physical with the Digital Investigative Process”, *International Journal of Digital Evidence*, Vol. 2, No. 2, pp 1-20.

CNN Tech (2001) “Email Mailboxes to Increase 1.2 Billion Worldwide by 2005”, 19 September 2001 [online] http://articles.cnn.com/2001-09-19/tech/email.usage.idg_1_email-market-email-usage-accessing-email?_s=PM:TECH.

Daubert v. Merrell Dow Pharmaceuticals, Inc. Daubert (1993) 509 U.S. 579.

Digital Forensics Research Workshop (2001) DTR-T001-01 Final: *A Roadmap for Digital Forensic Research – Report from the First Digital Forensic Workshop (DFRWS),* [online] http://www.dfrws.org/2001/dfrws-rm-final.pdf.

Duranti, L. (1989) “Diplomatics: New Uses for an Old Science, Part I”, *Archivaria*, Summer, No. 28, pp.7-27.

Duranti, L. (1989) “Diplomatics: New Uses for an Old Science, Part II”, *Archivaria*, Fall, No. 29, pp.4-17.

Duranti, L. (1999) “Concepts and Principals for the Management of Electronic Records, or Records Management Theory is Archival Diplomatics”, *Records Management Journal,* Vol. 9, No. 3, pp. 148-171.

Duranti, L. (2007) “The InterPARES 2 Project (2002-2007): An Overview”, *Archivaria*, No. 64, pp. 113-121.

Duranti, L. (2009) “From Digital Diplomatics to Digital Records Forensics”, *Archivaria*, No. 68, pp. 39-66.

Duranti, L. (2010) “Concepts and Principles for the Management of Electronic Records, or Records Management Theory is Archival Diplomatics”, *Records Management Journal*, Vol. 20, No. 1, pp. 78-95.

Duranti, L., Eastwood, T. and MacNeil, H. (2002) *Preservation of the Integrity of Electronic Records*, Kluwer Academic Publishers, Norwell, Massachusetts.

Duranti, L. and Thibodeau, K. (2006) "The Concept of Record in Interactive, Experiential and Dynamic Environments: The View of InterPARES", *Archival Science,* Vol. 6, no. 1, pp 13-68.

Edwards, H. (18 Mar 2009) *Strengthening Forensic Science in the United States: A Path Forward. Statement before the United State Senate Committee on the Judiciary.* [online] http://judiciary.senate.gov/pdf/09-03-18EdwardsTestimony.pdf.

Gerber, M. and Leeson, J. (2004) “Formalization of Computer Input and Output: The Hadley Model”, *Digital Investigation,* No.1, pp. 214-224.

Jansen, A. (2010) “Digital Records Forensics: Ensuring Authenticity and Trustworthiness of Evidence over Time”, *Proceedings of the Fifth International Workshop on Systematic Approaches to Digital Forensic Engineering,* pp. 84-88.

InterPARES 1 (IP1) Authenticity Task Force (2000) “Appendix 1: Template of Analysis”, *The Long Term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, University of British Columbia, [online] http://www.interpares.org/book/interpares_book_j_app01.pdf.

InterPARES 1 (IP1) Preservation Task Force (2001) “Appendix 6: How to Preserve Authentic Electronic Records”, *The Long Term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, University of British Columbia, [online] http://www.interpares.org/book/interpares_book_o_app06.pdf.

InterPARES 2 (IP2) (2008) Luciana Duranti and Randy Preston, eds., *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records Appendix 7: Diplomatic Analysis Template*, [online], University of British Columbia, [online] http://www.interpares.org/ip2/display_file.cfm?doc=ip2_book_appendix_07.pdf.

Kenneally, E. (2001) “Gatekeeping Out of the Box: Open Source Software as a Mechanism to Assess Reliability for Digital Evidence”, *Journal of Law and Technology*, Vol. 6, No. 13, [online] http://www.vjolt.net/vol6/issue3/v6i3-a13-Kenneally.html.

Kumho Tire Co. v. Carmichael, (97-1709) 526 U.S. 137 (1999) 131 F.3d 1433, reversed.

Lemieux, V. (2001) “Let the Ghosts Speak: An Empirical Exploration of the Nature of the Record”, *Archivaria,* No. 51, pp. 81-111.

MacNeil, H. (2000) “Providing Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Electronic Records”, *Archivaria*, No. 50, pp. 52-78.

MacNeil, H. (2004) “Contemporary Archival Diplomatics as a Method of Inquiry: Lessons learned from Two Research Projects”, *Archival Science,* No. 4, pp. 199-232.

Meyers, M. and Rogers, M. (2004) “Computer Forensics: The Need for Standardization and Certification”, *International Journal for Digital Evidence*, Vol. 3, No. 2.

Reith, M, Carr, C and Gunsch, G. (2002) “An Examination of Digital Forensic Models”, *International Journal of Digital Evidence*, Vol. 1, No. 3 [online] http://www.utica.edu/academic/institutes/ecii/ijde/articles.cfm?action=issue&id=3.

Skemer, D. (1989) “Diplomatics and Archives”, *American Archivist,* Vol. 52, Summer, pp. 376-382.

UC-Berkeley, School of Information Management and Systems (2003) “How Much Information?”, [online] http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/execsum.htm.