Digital Records Forensics:

Ensuring Authenticity and Trustworthiness of Evidence Over Time

Adam Jansen
Ph.D. Student, School of Library, Archival and Information Studies
University of British Columbia
Vancouver, British Columbia
Adam@Dkives.com

***Abstract*—As digital evidence becomes increasingly common in the court systems, educated, trained professionals are required to manage the lifecycle of evidence from initial collection through final disposition. This paper describes a *work-in-progress* between the University of British Columbia and the Vancouver Police Department toward the creation of a new discipline, called Digital Records Forensics, focusing on the authenticity, trustworthiness and admissibility of evidence over time. The Digital Records Forensics Project is engaging in a wide range of research activities including: multi-national, cross-domain interviews, extensive interdisciplinary literature reviews, case studies and digital forensics workflow modeling.**

Keywords- digital forensics, computer forensics, digital preservation, evidence, Diplomatics

I. INTRODUCTION

Over the past decade, the field of digital forensic investigation has begun to establish a formalized framework of operation amongst practitioners. As digital evidence is increasingly being utilized in the court system, and understood to a greater extent by those judges and law makers on the forefront of the technology curve, higher standards of conduct are being demanded of those collecting, examining, handling and presenting evidence. The result of these higher standards can be seen in the increase in amount of case law defining admissibility of evidence [1,2,3,4]. As an outgrowth of this case law, proposed digital forensics models [5,6,7,8] have gone to great lengths to define and establish the expectations amongst digital forensics examiners and court systems as to the required process for evidence to be considered authentic (defined in this paper as being what it purports to be, free from tampering) and trustworthy (defined as being dependable and reliable).

One of the major complicating factors in the digital forensics process, and therefore the standardization of procedures used by practitioners and the expectations of the courts, is the ever increasing rate of change of technology. The rate of change is significant enough that "keeping up" with technology changes has become a major dilemma to digital forensics examiners [8]. With a majority of the focus spent on staying current with the technology curve, little thought (and less research) has been given to those pieces of evidence that have already been heard in court, yet by law must be maintained in an authentic and trustworthy state beyond the court hearing. That is, the evidence must be kept authentic and trustworthy even after it has been submitted to the court. Within the typical legal system, evidence must be kept for years beyond the end of the sentence for re-examination by the court in the case of an appeal, or indefinitely in the case of unsolved crimes [9].

Based on the extreme nature of some crimes, that could translate into electronic evidence being kept for seventy-five years or more. Given the aforementioned increasing rate of technological changes, it is becoming increasingly difficult to access digital records (record defined as any document made or received in the course of a practical activity and retained as *evidence* of that activity) from ten years ago, let alone fifty. How many labs still maintain 5 ¼" floppies and copies of the WordStar word processing application popular in the 1980s? Little effort has been put into the problem of how to maintain evidence in an authentic and trustworthy state through the inevitable generations of technological changes. Based on the experience of the past twenty years (one state archives estimates that their state government has already lost 50% of its electronic records of permanent value due to loss, deterioration technological obsolesce) [10], it is highly unlikely that the evidence submitted today will be readable by technology twenty years from now.

The collection of evidence is further complicated by the transitory and often ephemeral nature of records on the internet. Given that Facebook pages can be updated from cellular phones, by the time a subpoena is obtained from a judge, the Facebook page in question could have updated dozens of times. Technologies such as Tiny URL (a web based service that 'shrinks' long URLs into very short URLs), and the distributed nature of DNS registries have further complicated the ability to capture authentic, trustworthy evidence off of the internet. One illustration of this difficulty is a recent case on spamming that dragged on for two years. By the time the case finally came to trial, a large percentage of the domain names for the web

pages in question had ceased to exist, with the records of the past owners unobtainable from the DNS registrar; resulting in a reduction in the potential fine of $40 million dollars to one of $300,000 [11].

Within the past few years, several pertinent projects have emerged from the archives field. These studies focused on the preservation of digital records that are of interest to the digital forensics field. Two of the larger scale projects are InterPARES and the Washington State Digital Archives. The InterPARES research project, a multi-national research project (lead by the University of British Columbia) has spent the past ten years researching the policies, strategies and plans of action necessary to ensure the longevity and authenticity of electronic records [12]. The Washington State Digital Archives, one of the first statewide digital archives (as of December 2009, the archives contained over 84 million digital objects), delineated three major components of information systems that need to be addressed in order to ensure the accessibility of electronic records over time: the metadata collected, the file formats used and the media upon which the records are stored [13].

II. DIGITAL RECORDS FORENSICS PROJECT

The Digital Records Forensics Project (DRF for short) is a collaborative, interdisciplinary project between the University of British Columbia's School of Library, Archival and Information Studies (SLAIS), the Faculty of Law, and the Vancouver Police Department Computer Forensics Division [14]. The research being conducted throughout the course of this project seeks to address several distinct challenges to the long term admissibility of digital evidence currently being faced by forensics practitioners, records managers, archivists and legal professionals. Among the challenges to be addressed are:

- Identification of pertinent records amongst the myriad of digital objects produced in today's society;
- Handling and authentication of transitory and ephemeral nature of web–based records
- Determining authenticity of records maintained outside of their original environments
- Authentication of records of unknown origin, and
- Authentication of records that are in proprietary or obsolete formats or media.

A. Project Objectives

The goal of the DRF is to develop a new graduate level degree program called Digital Records Forensics, that combines the study of digital forensics, archival studies, Diplomatics (Diplomatics being the critical study of records in order to determine their authenticity [15]), the Law of Evidence (the Canadian rules and procedures governing the admissibility of evidence into a court of law). In order to achieve this goal, the DRF has developed the following objectives:

- Develop the concepts and methods necessary for records managers, archivists, digital forensics practitioners, law enforcement and legal professionals to recognize records from amongst the plethora of digital objects produced once they are removed from their native environment
- Develop the concepts and methods necessary to determine the authenticity of digital records that are no longer in their native environment, original format or original media
- Develop methods for the extraction of records from their native environment in such a way that their authenticity for the long term can be assumed
- Develop the methods necessary to maintain the legal admissibility, over the long term, of digital evidence collected or created by digital forensics examiners so that their authenticity and trustworthiness is beyond reproach, and
- Identify and develop the theoretical and methodological content for a new discipline that combines knowledge from the fields of diplomatics, archival science, digital forensics and the Law of Evidence that is focused on the maintenance of the authenticity, trustworthiness and admissibility of digital evidence over of the long term.

B. Project Methodology

In order to achieve these objectives, the DRF has employed a multi-phased approach to iteratively build upon the knowledge gained in the InterPARES project. During the first phase, the DRF team conducted an in-depth study of the existing literature concerning digital evidence in order to establish a clear and thorough understanding of the problems presented by current legal requirements for the admissibility of digital evidence, as well as the validity of the proposed solutions based on the DRF's theoretical framework.

The second phase of the project comprised of a review of the existing scholarly literature and professional writings in the archival science, computer forensics, digital forensics and legal fields. This literature review allowed the team to identify the various branches of each respective profession that commonly has exposure to digital evidence, in order to create an up-to-date, comprehensive, bibliographic database of writings from those branches. This bibliography serves as the fundamental resource for the emerging trends in digital forensics for the research project [14]. The works contained in the bibliography

formed the basis of a series of six semi-structured interview questionnaires, each targeted toward a specific sector or area of expertise: court clerks, digital forensic practitioners, judges, lawyers (both prosecution and defense), law enforcement officials, and records managers. Subject matter experts were identified and interviewed for each of these areas, with interviews tape recorded and transcribed. The purpose of the interviews was to determine:

- Which criteria these various experts consider the basis for determining the authenticity and trustworthiness of the digital evidence
- The methods they view as necessary to maintain the authenticity and trustworthiness of digital evidence from moment of collection through its final use
- The major issues challenging the maintenance of authentic, trustworthy evidence, and
- The knowledge, skills and abilities required for a digital records forensics professional to succeed.

To date, several dozen interviews have been conducted with a roughly equal number of representatives from each profession. The results of the interviews are being analyzed by the team to identify key trends. Further information in regards to the interviews can be found on the DRF website [14].

Concurrent with the analysis of the questionnaires, an ethnographic study is examining the context and procedures used in the evidence room of the Vancouver Police Department's Computer Forensics Unit, as well as the responses given by the interviewees to the questionnaires.

Under the direction of the DRF's experts in Law of Evidence and digital forensics, the graduate research assistants (GRAs) analyzed and described the hierarchy for policy changes and decision making, current court procedures regarding admissibility of digital evidence, and problems noted by those who handle the digital evidence at each step of the process. Additionally, the team Diplomatics expert has interviewed classic Diplomatics scholars to determine what aspects of classic Diplomatics can be applied to digital evidence, and what component should be taught to digital records forensics professionals.

The findings from this phase were used to create a comprehensive model of the digital forensic process, from collection through submission to court and final destruction of the evidence. To model this process workflow, the DRF utilized the IDEF0 method to map the various inputs, outputs, mechanisms and controls that have bearing on individual, discrete functions within the digital forensics process, from initial collection through the filing with the court clerk [16]. IDEF0 was designed for the United States Air Force as a way to model the decisions, actions and activities of an organization. Using IDEF0 modeling has provided the DRF with a highly detailed look into the end to end digital forensics process, including the interaction between the various functions, as well as the controls that drive each individual step along the way.

The next phase of the project will entail an examination of records from the Computer Forensics Unit of the Vancouver Police Department. These records have been identified by the VPD as particularly problematic due to issues of technological obsolescence, unstable media, legacy systems, internet based records, or records of unknown origin. While virtual machines can help address some issues of technological obsolesce of software, they are less promising with issues of hardware, particularly obsolete storage devices. Furthermore, emulation has not proven to be an acceptable preservation option [12]. These pieces of evidence will be examined, based on the knowledge gained in the previous phases of the research, utilizing analytical methods from all the disciplines involved in the project, including philological analysis of text, as needed. Solutions to the issues for each of the 'problem areas' identified by the VPD will be implemented with copies of the evidence and tested against the Rules of Evidence, elaborating on existing policies and procedures. Where necessary, new concepts and procedures will be developed utilizing knowledge gained from the literature review and from the other professions.

The final phase of the project will be the distillation of the knowledge accumulated during the research and experience gained on the case study into the core content of a new Digital Records Forensics discipline. The content developed by the DRF will be used to form the basis of a proposal to form a new interdisciplinary graduate program at the School of Library, Archival and Information Studies at the University of British Columbia. It is the intent that graduates of this program will have the knowledge, skills and ability to successfully develop the policies and procedures necessary for the maintenance of the authenticity and trustworthiness of evidence, and ultimately their admissibility in court, over time. Additionally, graduates of this program will be able to work and communicate directly with digital forensics examiners, legal professionals and records managers in order to manage the entire lifecycle of the evidence process, from initial collection through the ultimate disposition of the evidence when no longer needed.

III. INITIAL FINDINGS

As this project is a work in process and in the early stages of the research, four interesting challenges, mentioned by multiple interviewees, have been noticed by the author. With more interviews still to be conducted, it would be premature to state that these issues will be a primary focus moving forward; yet these challenges bear mentioning at this point as potential issues, and in the event that the pattern of 'shared pain' continues to be noted in further interviews, further examination and research will be warranted.

First, storage of digital evidence is becoming a huge issue for forensic examiners. While just five years ago, Gigabyte databases were uncommon, now it is not unusual to encounter Terabyte databases. Not only does this require vast amounts of storage within the examiners unit, the time required to image, transfer, load and move that amount of data is rapidly increasing as well, necessitating a review of the workflow process to account for the longer 'down' times while waiting to proceed.

Second, going hand in hand with the vast increases in storage, is the corresponding increase in the volume of the files that are being encounter during digital forensics analysis. One case studied involved examining server logs that generated three terabytes a week (containing over 28 million entries), which while large, are not uncommon in larger data centers [17]. Databases are also rapidly increasing in the number of records they contain. Similar increases are also being encountered in domestic investigations as well due to high speed internet and larger, cheaper hard drives allowing the accumulation and movement of huge numbers of files over the internet.

Third, criminals are becoming increasingly savvy about their targets and their methods of operation. This is particularly evident in large, internet based, service oriented businesses. This class of businesses derives their revenue from the uptime of their servers; and their IT managers can typically quote the amount of money the company loses for every minute the servers are down. Hackers understand this and are increasingly creating havoc on the way out, knowing that the first instinct of the IT manager will be to restore the system to operational order (often from images or backup tapes) as quickly as possible; thereby erasing all tracks of the hacker in the process [18].

Lastly, the rapidly changing and often ephemeral nature of internet based records appears to present a significant challenge to digital forensic examiners. Unlike computer forensic evidence collection, internet based records can be based in other states, countries or split among multiple locations. In addition to the examples previously given, perpetrators of internet based crimes are becoming increasingly savvy with moving webpages and domain registrations in order to obscure the ownership of the records [Simon]. Often large backlogs of digital forensics examinations that exist in many law enforcement agencies makes evidence collection in internet based types of crimes more difficult; by the time the case reaches the top of the queue, the website might have changed significantly, or simply ceased to exist.

IV. Future Research Directions

This paper presents a high level description of the Digital Records Forensics Project, comprised of the School of Library, Archival and Information Studies, as well as the Faculty of Law from the University of British Columbia and the Computer Forensics Unit of the Vancouver Police Department, in its endeavor to define the requirements for a new discipline -- Digital Records Forensics. Future work from this research project will continue to refine the IDEF model developed for this project, the expansion of the case study at the Vancouver Police Department and the development of a graduate level degree in Digital Records Forensics. Continuation and expansion of the interviews being conducted will provide further insight into the depth of the challenges mentioned above. Should these challenges be validated, further research into addressing these challenges will be in order. Additionally, several of the Ph.D. students working as GRAs on the project are using specific issues discovered during investigation phases of the project as the basis of their doctoral dissertations.

In conclusion, the research conducted to date by the DRF team has shown a clear indication of a lack of trained, knowledgeable professionals to ensure the authenticity and trustworthiness of evidence over the long term. As there currently exists no opportunity for practitioners to obtain the education and training necessary to provide for the maintenance of authenticity and trustworthiness over time and space, findings to date have shown there exists the need for the development of the theoretical and methodological content for a new discipline -- Digital Records Forensics [14].

References

[1] *Frye v. United States.* 293 F. 1013 (D.C. Cir. 1923).

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

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

[4] R. v. Penny, 2002 NFCA 15.

[5] N. L. Beebe and J. G. Clark, "A heirarchical, objectives-based framework for the digital investigations process", Digital Investigation 2, No. 4, 2005, p147-167.

[6] S. Ciardhuain, "An extended model of cybercrime", International Journal of Digital Evidence 3:1, Summer 2004 Retrieved 17 November 2009 from the World Wide Web: www.ijde.org.

[7] M. Reith, C. Carr, and G. Gunsch, "An examination of digital forensic models", International Journal of Digital Evidence 1:3, Fall 2002, Retrieved 17 November 2009 from the World Wide Web: http://www.utica.edu/academic/institutes/ecii/ijde/articles.cfm?action=article&id=A04A4

[8] Digital Forensic Research Workshop (DRFWS). DTR-T001-01, DFRWS Technical Report: A Road Map for Digital Forensic Research, 2001, Retrieved Dec 22, 2009 from the World Wide Web: www. DFRWS.org.

[9] Sgt. Mark Johnstone, Vancouver Police Department, Digital Records Forensics Project Interview, November 25, 2009.

[10] A. Jansen, Experiences in Digital Preservation (Fall 2009), Guest Lecture, ARST555 Digital Preservation, University of British Columbia.

[11] M. Simon, Creation Logic, Digital Records Forensics Project Interview, November 24, 2009.

[12] L. Duranti and R. Preston, International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2 Experimental, Interactive and Dynamic Records, Italy:Podva, 2008. See also www.interpares.org.

[13] Office of the Secretary of State, Washington State Digital Archives Feasability Study, 2003. Retrieved on January 10, 2010 from the World Wide Web: www.digitalarchives.wa.gov/.

[14] Digital Records Forensics Project. Retrieved on January 15, 2010 from the World Wide Web: www.digitalrecordsforensics.org/.

[15] O. Guytojeannin, "The expansion of diplomatics as a discipline", American Archivist, Vol. 59, Fall 1996, p. 414-421.

[16] Knowledge Based Systems, Inc. Retrieved 14 January 2009 from the World Wide Web: www.idef.com.

[17] S. Peisert, A Model of Forensic Analysis Using Goal-Oriented Logging, Retrieved 22 November 2009 from ProQuest Dissertations and Theses Database.

[18] D. A. Allen, Developing a Proactive Digital Forensics System. Retrieved 22 November 2009 from ProQuest Dissertations and Theses Database.