Digital Records Forensics Project

Forming the eXtreme Professional

Combining Archival and Diplomatics Theory with the Law of Evidence and Computer Science to Create the **Digital Records Forensics Expert**

UBC Faculty of Law Colloquium
9 April 2010
The University of British Columbia
Vancouver, BC, Canada

Luciana Duranti, Anthony Sheppard & Alexandra Allen

Digital Records Forensics Project

The most challenging issues presented by digital technology:

1. The identification of "records" among all the digital objects it produces
2. the determination of their “authenticity”

Issue 1. is addressed by Digital Diplomatics

Issue 2. is addressed by Digital Forensics

Digital Records Forensics Project

Digital Diplomatics is a contemporary development of a centuries-old discipline that studies the nature, genesis, formal characteristics, structure, transmission and legal consequences of records

Digital Forensics is the use of scientifically derived and proven methods toward the collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence derived from digital sources for the purpose of facilitation or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations

Digital Records Forensics Project

The legal systems, both common and civil law, consider records to be a very special kind of documentary evidence.

In civil law environments, a record is admissible as evidence in court simply on the basis of the recognition of its record nature.

In common law environments, in addition to *relevance*, disputed records may require further steps to gain admissibility, such as proof of *authenticity*, and compliance with the *best evidence* and the *hearsay* rules.

Digital Records Forensics Project

Archival Definition of Record

a document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference

Business Records

documents considered admissible under an exception to the hearsay rule because they were made or received in the usual and ordinary course of business

Digital Records Forensics Project

The issue of **what is a record in the digital environment** keeps coming up at trials and in political discussions.

- British Columbia Rail case: as questions were raised of preserving “records,” the Liberal MLA Ralph Sultan asked “What is the definition of a record?” referring “to the controversy over to what extent e-mails qualify” (Vancouver Sun, January 29, 2010).
- The Supreme Court of Canada will be hearing a defamation case on whether hyperlinks in a text constitute a repetition of a defamatory statement (Vancouver Sun, April 2, 2010).

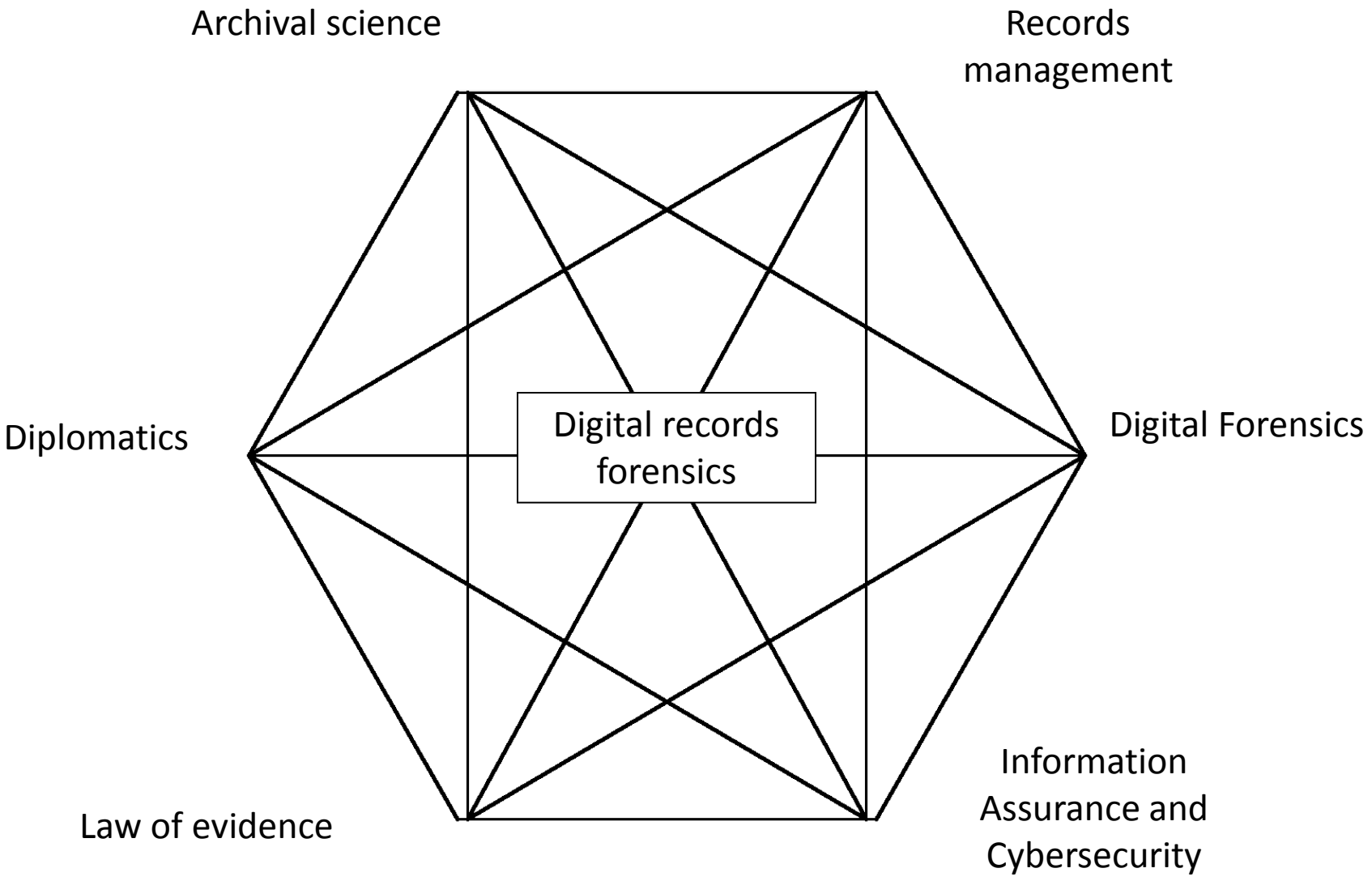
The Record as "Best Evidence"

- Required by Common Law – the original is preferred; if unavailable a copy may be admissible
- No originals in digital environment
- Experts must attest to both authenticity of record and integrity of system

Digital Records Forensics Project

The described issues can only be addressed by developing a body of interdisciplinary knowledge from already established disciplines and educating a new breed of professionals in what we have come to call "**Digital Records Forensics**"

Digital Records Forensics



Digital Records Forensics Project

The objectives of the Digital Records Forensics Project:

- to develop concepts and methods that will allow the records management, archival, legal, judicial, law enforcement and digital forensics professions to recognize records among all digital data objects produced by complex digital technologies once they have been removed from the original system;
- to develop concepts and methods to determine the reliability, accuracy and authenticity of records no longer in the original digital environment;

Digital Records Forensics Project

The objectives of the Digital Records Forensics Project:

- to identify, develop and organize the content of a new science and discipline called “Digital Records Forensics;” and

- to develop the intellectual components of a new program of education for Digital Records Forensics experts.

Digital Records Forensics Project

The DRF Project is a **collaboration** between

- The UBC School of Library, Archival & Information Studies
- The Faculty of Law, and
- the Computer Forensics Division of the Vancouver Police Department

Methodology

- Literature review
- Digital Records Forensics Activity Model
- Case Law Database
- Terminology Database
- Questionnaires and Interviews
- Ethnographic study with the Vancouver Police Department

http://www.digitalrecordsforensics.org/

Digital Records Forensics Project

Conceptual Findings

- What is a record
- What is authenticity

Methodological Findings

- Recordkeeping/Preservation

What is a Record: the Archival/Diplomatics View

- **Act**: an action in which the records participates or which the record supports (naturalness and impartiality)
- **Persons Concurring to Its Creation**: author, writer, originator, addressee, and creator
- **Archival Bond**: explicit linkages to other records inside or outside the system (uniqueness)
- **Identifiable Contexts**: juridical-administrative, provenancial, procedural, documentary, technological (interrelatedness)
- **Medium**: necessary part of the technological context, not of the record
- **Fixed Form and Stable Content**

Fixed Form

- An entity has 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 (different digital presentation: Word to .pdf)
- An entity has fixed form also if the same content can be presented on the screen in several different ways in a limited series of possibilities: we have a different documentary presentation of the same stored record having stable content and fixed form (e.g. statistical data viewed as a pie chart, a bar chart, or a table)

Stable Content

- An entity has stable content if the data and the message it conveys are unchanged and unchangeable, meaning that data cannot be overwritten, altered, deleted or added to

- **Bounded Variability**: when changes to the documentary presentation of a determined stable content 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 intention of the author or to different operating systems or applications

The Parts of a Digital Record

- **Formal Elements**: constituent parts of the record documentary form as shown on its face, e.g. address, salutation, preamble, complimentary close
- **Metadata**: the attributes of the records that demonstrate its identity and integrity
- **Digital Components**: stored digital entities that either contain one or more records or are contained in the record and require a specific preservation measure

Digital Records Forensics Project

- **Stored record**: it is constituted of the digital component(s) used in re-producing it, which comprise the data to be processed in order to manifest the record (content data and form data) and the rules for processing the data, including those enabling variations (composition data)

- **Manifested record**: the visualization or instantiation of the record in a form suitable for presentation to a person or a system. Sometimes, it does not have a corresponding stored record, but it is re-created from fixed content data when a user's action associates them with specific form data and composition data (e.g. a record produced from a relational database)

Digital Records Forensics Project

Static Records: They do not provide possibilities for changing their manifest content or form beyond opening, closing and navigating: e-mail, reports, sound recordings, motion video, snapshots of web pages

Interactive Records: They present variable content, form, or both, and the rules governing the content and form of presentation may be either fixed or variable. They can be divided into dynamic and non-dynamic

Digital Records Forensics Project

- **Non-dynamic**: the rules governing the presentation of content and form do not vary, and the content presented each time is selected from a fixed store of data. Ex. Interactive web pages, online catalogs, records enabling performances—they are records
- **Dynamic**: the rules governing the presentation of content and form may vary—they are either information systems or potential records

Digital Records Forensics Project

Records Functions

- ***Ad substantiam*** (e.g., contracts): the record is the action
- ***Ad probationem*** (e.g., registries): the record proves the action
- **Supporting**: generated to be used in the course of activity (ies) as a source of information, often by multiple users (e.g., GIS)
- **Narrative**: generated on a purely discretionary basis only as a means of communication (e.g., most e-mails, memos, web sites)

Digital Records Forensics Project

- **Instructive**: provide guidance on the way in which external data or documents are to be presented (e.g., scores, scripts, regulations, manuals of procedure, instructions for filling out forms)
- **Enabling**: enable the performance of artworks, the execution of business transactions (interacting business applications), the conduct of experiments (a workflow generated and used to carry out an experiment of which it is instrument, byproduct and residue), the analysis of observational data (interpreting software), etc. Most of them are stored only records.

Digital Records Forensics Project

What is a Record—Digital Forensics View

- **Computer Stored:** They contain human statements and are considered hearsay (tested for truthfulness and accuracy under the business records exception to the hearsay rule): e.g. e-mail messages, word processing documents, and Internet chat room messages.
- **Computer Generated**: They do not contain human statements, but they are the output of a computer program designed to process input following a defined algorithm (tested for authenticity on the basis of the functioning of the computer program): e.g. server log-in records from Internet service providers, ATM records.
- **Computer Stored & Generated:** e.g. a spreadsheet record that has received human input followed by computer processing (the mathematical operations of the spreadsheet program).
- **Dynamic Records:** Records in live systems

Digital Records Forensics Project

- It is essential to reconcile the forensic view with the archival/diplomatic view in light of the application of the hearsay rule to computer generated records, and of the determination of whether they constitute substantive evidence (revealing intentionality) or demonstrative evidence (showing capability)
- This reconciliation must be made in relation to the concept of trustworthiness as it is understood by the two fields in order for the legal system to establish at any given time whether it is concerned with issues of reliability or authenticity

Trustworthiness (the archival/diplomatics view)

Reliability

The trustworthiness of a record as a statement of fact,

based on:

- the competence of its author
- the controls on its creation

Accuracy

The correctness and precision of a record's content

based on:

- the competence of its author
- the controls on content recording and transmission

Authenticity

The trustworthiness of a record that is what it purports to be, untampered with and uncorrupted

based on:

- identity
- Integrity
- reliability of the system

Authenticity

Identity: The whole of the attributes of a record that characterize it as unique, and that distinguish it from other records (e.g. date, author, addressee, subject, identifier).

Integrity: A record has integrity if the message it is meant to communicate in order to achieve its purpose is unaltered (e.g. chain of custody, security, technical changes).

Authentication

A means of declaring the authenticity of a record at one particular moment in time -- possibly without regard to other evidence of identity and integrity.

Example: the digital signature. Functionally equivalent to medieval seals (not signatures): verifies origin (identity); certifies intactness (integrity); makes record indisputable and incontestable (non-repudiation)

But, medieval seals were associated with a person; digital signatures are associated with a person and a record

Digital Records Forensics Project

The Digital Forensics View

Reliability: the term is used in reference to the source of the records, and defined in digital forensics in a way that points to a reliable software. This would be an open source software, because the processes of records creation and maintenance can be authenticated with evidence either by describing a process or system used to produce a result or by showing that the process or system produces an accurate result

Digital Records Forensics Project

The Digital Forensics View

Accuracy: a component of integrity, it is one of the key qualities of the evidence. Data are guaranteed accurate if they are repeatable. "Repeatability," which is one of the fundamental precepts of digital forensics practice, is supported by the accurate documentation of each and every action carried out on the evidence.

Open source software is the best choice for assessing accuracy, especially when conversion or migration occurs, because it would allow a practical demonstration that nothing could be altered, lost, planted, or destroyed in the process

Digital Records Forensics Project

The Digital Forensics View: *data integrity* vs *duplication integrity*

Data integrity: the fact that data are not modified either intentionally or accidentally "without proper authorization."

Duplication integrity: 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." Digital forensics experts also link duplication integrity to time and have considered the use of time stamps for that purpose.

Digital Records Forensics Project

The Digital Forensics View: principles of *non-interference* and *identifiable interference*

Non-interference: the method used to gather and analyse [or acquire and preserve] digital data or records does not change the original digital entities [always kept]

Identifiable interference: if the method used does alter the original entities, the changes are identifiable

These principles, which embody the ethical and professional stance of digital forensics experts, are consistent with the traditional impartial stance of the archivist , as well as with his/her new responsibility of neutral third party, or trusted custodian

The Digital Forensics View

Authenticity: "the data or content 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. In digital forensics, the term "source" is used in a general way to refer to either a person (physical or juridical), a system, software, or a piece of hardware.

Like in diplomatics, authenticity implies integrity, but the opposite is not true, that is, integrity does not imply authenticity.

Digital Records Forensics Project

The Digital Forensics View

Authentication: Proof of authenticity 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 computer programmer 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.

The strength of circumstantial digital evidence could be increased by metadata which records (1) the exact dates and times of any messages sent or received, (2) which computer(s) actually created them, and (3) which computer(s) received them.

Digital Records Forensics Project

The Digital Forensics View

Authentication: a chain of legitimate custody is ground for inferring authenticity and authenticate a record.

Digital chain of custody: the information preserved about the data and its changes that shows specific data was in a particular state at a given date and time

Biometric identification systems and cryptography are not considered the prevalent means of authentication.

The Digital Forensics View: Authentication (cont.)

A declaration made by an expert who bases it on the trustworthiness of the system containing the record and of the procedures controlling it is a prevalent method.

System Integrity: The quality that a system has when it performs its intended function in an unimpaired manner, free from unauthorized manipulation of the system, whether intentional or accidental.

The assessment of system integrity is based on the rules used for scientific evidence

Digital Records Forensics Project

The Digital Forensics View

- the theory, procedure or process for making or keeping the record has been tested or cannot be tampered with
- it has been subjected to peer review or publication
- the known or potential error rate is acceptable
- it is generally accepted within the relevant scientific community

Digital forensic experts look for **repeatability, verifiability, objectivity** and **transparency**

Implications of conceptual Findings

An additional objective for the DRF:

To ensure that the Law of Evidence maintains an awareness of the changing nature of documentary evidence determined by digital technologies and by its rapid obsolescence and considers adjusting its requirements and procedures to the changing characteristics of such evidence

Digital Records Forensics Project

Methodological Findings:

http://www.digitalrecordsforensics.org/

The professions involved with record evidence have a disciplinary perspective.

Chain of custody or chain of continuity are the basis for presumption of reliability and authenticity.

Terminology is conflicting: e.g. copy vs. image

Issue of application of hearsay rule to computer generated records

Digital Records Forensics Project

Preservation requirements are longer, becoming indefinite

Every person interviewed has supported the idea of developing digital records forensics knowledge and a program of education delivering it.

The courts still have paper minds (e.g. hard-drive as a filing cabinet)