Digital Records Forensics Project

DIGITAL RECORDS FORENSICS: A NEW SCIENCE AND ACADEMIC PROGRAM FOR FORENSIC READINESS

Conference on Digital Forensics, Security and Law (ADFSL)
21 May 2010
St. Paul, MN, USA

Luciana Duranti, The University of British Columbia
Barbara Endicott-Popovsky, University of Washington

Digital Records Forensics Project

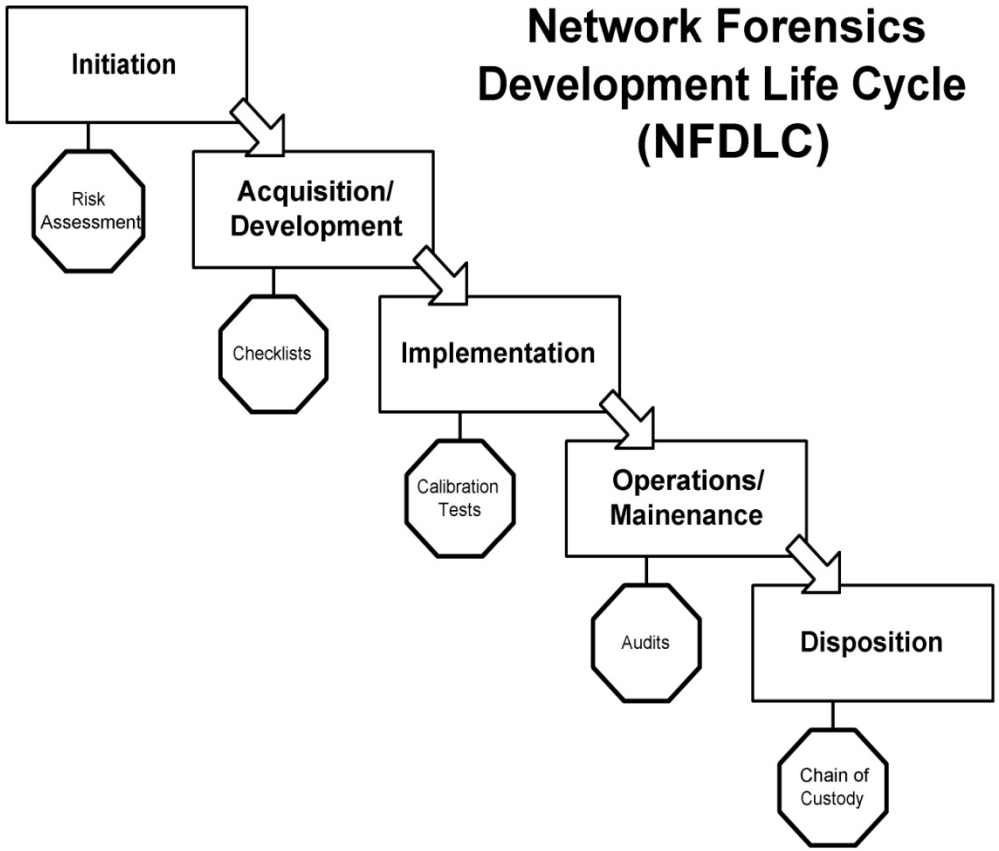
In a typical incident response….

- Technicians must choose:
  - Expend effort collecting forensically sound data, or
  - Simply restore network as quickly as possible
    - » Evidentiary files altered in the process
    - » Forensic value limited
- Expediency wins…and so do attackers!

New Zealand vs. Russian Cases

Characteristics	NZ Hacker Case	Russian Hacker Case
Type of attack	Typical script kiddie intrusion scenario	Online criminal automated auction scam
Damages	$400,000	$25 million
Investigator time	417 hours	9 months
Consequences	Community service	3 & 4 years in Federal prison

Digital Records Forensics Project



Digital Records Forensics Project

ISDLC (Life Cycle) Phases	NFDLC Additional Procedures
Initiation Phase: preliminary risk assessment	Determine what aspects of a network warrant digital forensic protection
Acquisition/ Development Phase	Adhere to Rules of Evidence in system requirements
Implementation Phase	Perform baseline testing Perform network/mechanism verification/calibration tests
Operation/ Maintenance Phase	Conduct verification/calibration audits
Disposition Phase	Incorporate chain of custody/ evidence preservation procedures

Digital Records Forensics Project

Collecting evidence across the lifecycle of forensic data

- Gather
- Store
- Retrieve
- Make available
- Secure
- Create repeatable process

Underlying Principles

IA Principles	RM Principles
Integrity	Integrity and reliability
Availability	Useability
Authentication	Authenticity
Non-repudiation	
Confidentiality	

Linkages and context

- content - facts about the activity are accurate and complete,
- context - circumstances of its creation and use and
- structure - relationships between data/information/records must be transparent

Digital Records Forensics Project

Contemporary archivists must act as neutral third parties, trusted recordkeepers and trusted custodians.

The most challenging issues they are presented with:

1. The identification of "records" among all the digital objects produced by digital technology
2. the determination of their “authenticity”

Both issues are addressed by Digital Diplomatics and Digital Forensics, two bodies of knowledge that share much more

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 of digital evidence, its validation, identification, analysis, interpretation, documentation, and presentation to the relevant entities 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. It is also the foundation of "forensic readiness".

Digital Records Forensics Project

Archivists today are called to act as forensics experts, e.g. assessing the identity and integrity of records stored in a variety of obsolete or nearly obsolete hardware/software and/or formats, or on portable media, and attesting to it

Digital forensic experts are called to act as archivists, e.g. identifying what digital materials fall under the definition of business records, and keeping them intact for as long as needed. They are also called to attest to and sometimes provide quality assurance for digital system that produce and/or contain records.

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: the judge pointed out that legislation speaks of preserving "records," and 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 is deciding whether hyperlinks in a text are akin to footnotes or make of the material to which they connect the reader a component of the document being read (Vancouver Sun, April 2, 2010).

Digital Records Forensics Project

The trustworthiness of digital records is equally problematic:

- The law prefers that the original of any document, regardless of medium or form, be used as evidence at trial, but in the digital environment, we no longer have originals.
- We cannot keep digital records. We can only maintain our ability to reproduce or even to re-create them as needed, or image the container of the records at a given point in time.
- As a consequence, the **authenticity** of digital records is difficult to establish on the records themselves and becomes ***an inference that one draws from the circumstances surrounding the creation, maintenance and preservation of the records.***
- This is a problem not limited to law enforcement, but affecting any human activity based on the trustworthiness of records, including research, therefore, it is an archivists' concern.

Digital Records Forensics Project

These issues and concerns can begin to be addressed by integrating digital diplomatics and digital forensics, starting with concepts and methods that have been the concern of both disciplines:

- the **nature of records** and
- their **trustworthiness**

What is a Record: the Diplomatics View

A document made or received as an instrument and residue of business activity, and kept for further action or reference

- **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 in the Digital Environment

- 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 in the Digital Environment

- 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 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, but the rules governing the content and form of presentation are fixed. Ex. Interactive web pages, online catalogs, records enabling performances

What is a Record—Digital Forensics View

A document produced in the usual and ordinary course of activity for the purposes of such activity by a person responsible for doing so.

- **Computer Stored:** They contain human statements and, in common law, 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).

Records Trustworthiness: The Diplomatics View

Reliability: The trustworthiness of a record as a statement of fact, *based on* the competence of its author and the controls on its creation

Accuracy: The correctness and precision of a record's content, *based on* the competence of its author, and 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 its* identity, integrity and the reliability of the system in which it resides

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.

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

Records Trustworthiness: The Digital Forensics View

Reliability: the trustworthiness of a record as to its *source*, defined in digital forensics in a way that points to either a reliable person or a reliable software.

This would be an open source software, because the processes of records creation and maintenance can be 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

Accuracy

A component of authenticity and, specifically, integrity. Digital entities are guaranteed accurate if they are repeatable.

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 evidence.

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

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.

Forensic Principles for Protecting Integrity

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

Identifiable interference: if the method used does alter the 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, of trusted custodian

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. Again, the term "source" is used 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.

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.

Other Means of Authentication

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

Digital chain of custody: the metadata preserved about the record and its changes that shows specific content and elements of form were in a particular state at a given date and time.

A declaration made by an expert who bases it on the **trustworthiness of the recordkeeping system** and of the procedures controlling it (quality assurance).

Other Means of Authentication (cont.)

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

Inference of system integrity: 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.

The assessment of system integrity is based on the Daubert Rules

The Daubert Rules

Extended from scientific to technical evidence (to assess demonstrative evidence, not substantive—capability rather than intentionality):

- 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 (standard)
- the known or potential error rate is acceptable
- it is generally accepted within the relevant scientific community

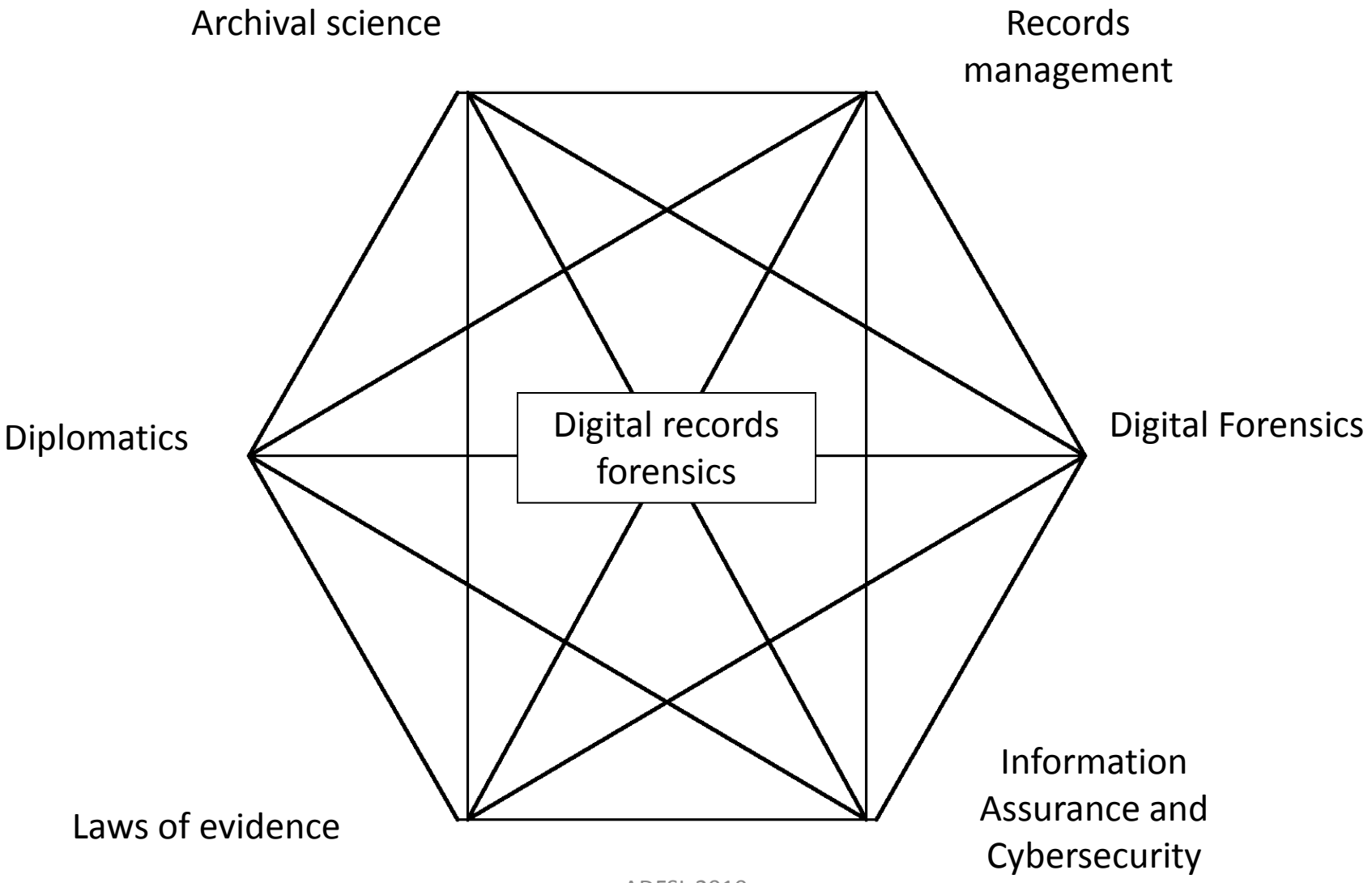
In a system, digital forensic experts look for **repeatability, verifiability, objectivity** and **transparency**

Digital Records Forensics Project

Objectives of the Digital Records Forensics Project:

- to carry out an analytical comparison and integration of the concepts and methods of Digital Diplomatics and Digital Records Forensics
- to further enrich this integrated body of knowledge with Archival Science (including records management), Laws of Evidence, and Information Assurance and Cybersecurity concepts and methods
- to identify, develop and organize the content of a new science and discipline called “Digital Records Forensics”
- to develop the intellectual components of a program of education for Digital Records Forensics experts, as a specialization of archival programs

Digital Records Forensics



Transition Required

- From an investigative and response mechanism to one of prevention, compliance and assurance (Barbin and Patzakis)

- Proactively influencing system design and the active life of the record

Digital Records Forensics Project

Digital Records Forensics Science BOK

- The Law of Evidence
- Diplomatics (specifically Digital Diplomatics)
- Digital Forensics
- Archival Science
- Information Technology
- Organizational Information Assurance

Digital Records Forensics Project

New Professional Master's Degree:

- Integration of MSIM at UW and MS Archival Science UBC
- Professionals educated in core theoretical and methodological knowledge that identifies their profession
- Learn international standards as well as specific, local and unique aspects of the local juridical-administrative environment in which they will work
- Educated in the scholarly and practical nature of their work.

Digital Records Forensics Project

"people need to understand what exactly a record is. And then understand the manner in which it's maintained. So you'd have to have the knowledge of what it is you're trying to maintain and then the knowledge of the systems that are maintained. So, yes, there's some very specific knowledge needed"

Mark Johnstone, Sergeant, Forensics Services Division, Financial Crime Unit, Vancouver Police Department

Digital Records Forensics Project

The DRF Project (2008-2011) is a collaboration between

- The UBC Archival Studies programs in the School of Library, Archival & Information Studies
- The UBC Law of Evidence Department in the Faculty of Law
- The University of Washington Information Assurance and Cybersecurity program in the School of Information, and
- the Computer Forensics Division of the Vancouver Police Department

http://www.digitalrecordsforensics.org/