New Approaches to Digital Forensics

Darren Cerasi
Director, I-Analysis Pte Ltd

CS 5 - 4
Digital Forensics and Risk Management

- Digital Forensics is an integral part of the risk management process
- Digital Forensics is a defined part of the CISM certification
- Domain 5—Incident Management & Response
Digital Forensics and Risk Management

Domain 5 – Task Statements:

“Establish the capability to investigate information security incidents (e.g., forensics, evidence collection and preservation, log analysis, interviewing).”

Domain 5—Knowledge Statements:

“Knowledge of forensic requirements for collecting, preserving and presenting evidence (e.g. admissibility, quality and completeness of evidence, chain of custody)”
Digital Forensics – What Is It?

Forensics is the use of science and technology to investigate and establish facts in a court of law.

Digital forensics is the application of computer investigation and analysis techniques in the interests of determining potential legal evidence.
Digital Forensics – Why?

- Defensible Evidence Collection
- Preservation of Evidence
- Regulatory Compliance
- Meet Legal Obligations
Digital Forensics – Collections

**Digital forensics** is the most reliable and defensible collection strategy that can be adopted.

Digital forensics allows you to make a full and complete image of the entire contents of the hard disk, including deleted files and partial files.

Digital forensics software used in conjunction the appropriate hardware does not change any of the data and retains it in a read-only and tamper evident file. All the data that is collected is verified to ensure that **authenticity and integrity is not in question.**
New features allow you to **only collect what is needed** through the use of filters e.g. file extension, date range, hash value.

Searches can be run across the entire content of a hard drive or just specific files to identify electronic evidence.

In conjunction with a documented process, **it is the best way to collect digital evidence.**
Collection Options – Drive to Drive

One of the oldest ways to make a forensic image of a hard disk. The computer would be booted up using a boot floppy diskette or boot CD.

We would then pray that we did not overwrite the evidence!
The computer is booted up using a boot floppy diskette or boot CD and a network connection would be made to the forensic computer.
Collection Options – Write-blockers

The computer hard disk is removed from the chassis or the laptop and connected to either a hardware write-blocker, a disk duplicator or directly to a forensic computer with a software write-blocker enabled.
The computer hard disk is generally not removed from the chassis or the laptop. The subject device is booted with a forensic boot CD or USB drive.

These are generally based on a Linux distribution and do not mount drives by default.
Collection Options – Network Collections

The computer is left on and connected to the network. Typically an agent is deployed to the subject device and the evidence is acquired over the network.

For the organisations there are enterprise class tools such as EnCase Enterprise and AccessData Enterprise.

For poor consultants like me there are other options!
Locations of Evidence
Types of Evidence
Digital Forensics – Challenges

A number of challenges arose such as:

- Data volume
- Encryption
- Mobile devices
- Diverse and complex IT environment
- Archives, back-up tapes
- On-line storage (cloud)
- Inadequate solutions
- Keyword driven
- No reduction of data
- Clients have no control of data
- Review process very slow, difficult and costly
Data Volume

1 TB = 1,024 GB
1 PB = 1,024 TB
1 EB = 1,024 PB

Source: IDC 2008
Data Volume

It is estimated that **210 billion e-mail messages** and 32.2 billion instant messages are sent each day!

The typical corporate email account will send and receive **18.5MB of data per day** according to The Radicati Group.

Wal-Mart, a retail giant, handles more than **1 million customer transactions every hour**.
For too long, digital forensics analysis has been driven by **keywords**. Though certainly helpful it is not a complete analysis.

If the evidence is not properly process, keywords cannot access the following data stores:

- Outlook PST files
- Attachments to emails
- Zip and other archives
- Registry artifacts
- Instant messaging chat logs
- Databases
- Backups of mobile devices
A smarter and more holistic approach would be to process the data first and then proceed with analysis. The advantage of this approach is that it reduces the volume of data and converts it, if necessary, to formats more suitable for review and analysis.

Typical processes include:

- Date range culling
- File filtering
- De-Duplicating
- Indexing
- Topic Identification
- Email Threading
- Email clustering
- Find similar documents
EnCase Forensic is an excellent digital forensics tool however it is not the be all and end all of digital forensics. There are numerous tools out there, both commercial, open source and free that should be part of any forensic practitioner’s arsenal.
eDiscovery – What Is It?

**Electronic discovery** (or eDiscovery) refers to discovery in civil litigation which deals with information in electronic format also referred to as Electronically Stored Information (ESI).

In Singapore electronic evidence is commonly known as **Electronically Stored Documents** (ESD).

Electronic information is **different from paper** information because of its intangible form, volume, transience and persistence.

Electronic information is usually accompanied by **metadata**, which is not present in paper documents. The **preservation of metadata** from electronic documents creates special challenges to prevent spoliation. (Partial Source: Wikipedia)
eDiscovery – What Is It?

Practice Direction No.3 of 2009 - Discovery and Inspection of Electronically Stored Documents came into force on 1st October 2009.

Currently an opt-in framework that provides explanation and guidance to lawyers on how to deal with the complex issue of electronic evidence.

The Practice Direction covers the discovery and inspection of electronic documents and also provides an eDiscovery Protocol for parties to agree to when dealing with electronically stored documents.
Solicitors have a duty to advise their clients that they **must not** deliberately destroy documents (including electronic documents) relevant to the issues in matters that are in your possession, custody or power.

Documents that are “**relevant**” to the issues in this matter include documents that do not support or adversely affect your case, or are confidential or informal.

A party “deliberately” destroys relevant documents if he intends to put these documents **out of reach of the other party in pending or anticipated litigation.**

Discovery Obligations

CANNOT!
Electronic Discovery Reference Model
Identification Meeting

Each and every IT environment is different so an early meeting with external counsel, in-house counsel and IT staff is best in order to ensure that all the data needed to be collected has been duly identified.

This early meeting will help counsel better understand the size of the matter and whether they need to engage external consultants to assist with the process.
At this stage it is also helpful for legal counsel to remind their clients of their discovery obligations.

Being able to produce meetings of minutes and where possible a data map helps clients demonstrate to the courts that they have taken their discovery obligations seriously.

Usually clients know where the data is located. There have been some cases where that was not the case. Some solutions can scan and index an entire network and identify potentially relevant documents.
Electronic Discovery Reference Model

- Information Management
- Identification
- Preservation
- Collection
- Processing
- Review
- Analysis
- Production
- Presentation

VOLUME

RELEVANCE

Electronic Discovery Reference Model / © 2009 / v2.0 / edrm.net

Analysis @Discovered
Data Collections

“For the avoidance of doubt, *electronically stored documents* residing in folders or directories in storage locations, media or devices, including folders or directories where temporarily deleted files are located (for example the Recycle Bin folder or Trash folder) are within the scope of general discovery;

*deleted files or file fragments* containing information which are recoverable through the use of computer forensic tools or techniques during a forensic inspection of the unallocated file space or file slack are not within the scope of general discovery.”
Data Collections

“Nothing in this paragraph shall prevent the party giving discovery from reviewing the discoverable electronically stored documents or the results of any reasonable search for the purpose of identifying privileged documents. However, such review for the purpose of identifying privileged documents shall not extend to the deletion, removal or alteration of metadata information.”

There is thus a requirement that none of the data be deleted, removed or altered in any way.

Sounds like digital forensics right?
Electronic Discovery Reference Model
Processing, Review & Analysis

**Processing**
Reduces the volume of data and converts it, if necessary, to formats more suitable for review and analysis.

Typical processes include:
- Date and file extension filtering e.g. Include .doc, exclude.mp3
- Removal of known files e.g. system files through hashing
- De-Duplicating
- Indexing
- Topic Identification
- Email Threading
- Language Identification
When thinking about which eDiscovery solution to use, a good analogy would be cars. These come in different sizes, come with different features, and can be acquired in different ways.

The same applies to eDiscovery solutions as some have better features than others whereas others scale better. Some solutions can be purchased and brought in-house whereas others are hosted solutions.

Each have their pros and cons so selecting the right one is important.
## Processing, Review & Analysis

**Mail Source**
Choose the mail source that you want to convert. Set the "Include mail in trash" checkbox if you want to process mail in the Trash folder. If available, set the "Search for mail location" checkbox to make the program search for the disk folder containing your mail) refine the search by indicating the account profile/identity name.
Processing, Review & Analysis
Review & Analysis

At this stage, lawyers evaluate the data for relevance, privilege and issues related to the matter.

This can be done in many different ways such as:
- Keyword Searching
- Content and Context Analysis
- Concept Searching
- Email Discussion
- Timeline Analysis
- Near de-duplication
- Advanced Analytics
- On the fly translation
Review & Analysis

Reviewers are also able to better organise their data and to assign tasks to speed up the review of the documents. Documents eventually need to be reviewed one by one and tagged for later production.

This can be done in many different ways such as:

- Grouping documents by language or custodian
- Assigning certain batches of documents to domain experts
- Redacting parts of or entire documents for privilege and relevance
- Tagging documents per issue
摘要

1964年，美国人乔治·德沃尔设计了第一台电子可编程的工业机器人，并于1961年发表了该项机器人专利。从此以后，机器人技术突飞猛进，如今，基于网络的机器人遥控操作和协同技术又成为机器人研究领域的热点之一。开展基于网络的异构机器人集成技术的研究，对于生产技术的提升，对以数字化、柔性化、敏捷化为基本特征的基于网络的制造技术都有非常重要的意义。本论文针对异构机器人集成作了深入的研究，论文的主要内容归纳如下。

第一章对基于网络的遥控操作机器人研究的发展和现状作了全面的分析，在此基础上，提出异构机器人集成技术的主要障碍在于系统的封闭性，解决障碍的方案在于充分利用开放式系统控制器，而利用开放式控制器进行集成的关键技术是要具备一个开放式结构的软件环境的观点。
<table>
<thead>
<tr>
<th>docid</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>3725347</td>
<td></td>
</tr>
<tr>
<td>3723333</td>
<td></td>
</tr>
<tr>
<td>3725355</td>
<td></td>
</tr>
<tr>
<td>3725418</td>
<td></td>
</tr>
<tr>
<td>3725421</td>
<td></td>
</tr>
<tr>
<td>3725423</td>
<td></td>
</tr>
</tbody>
</table>

Cc: Varadarajan Sivalapperi; Rakesh Bhatt; Larry Barela

Subject: CR 5.3 release

Sent: 06/01/2006 11:57:12 AM -0600 (MDT)

Attachments: 5.3 Release.xls; CaseShareLogo.bmp; note

Attached is a simple spreadsheet showing the features, enhancements and bug fixes that we plan to have in the 5.3 release, currently scheduled for deployment on June 21. We do not yet have a complete environment set up yet for testing all of these features, but you can get an overview of many of the features by using the test site located at https://test.caseshare.com/cr/index.asp?uniqueid=408641427 ... please note the some items are still in development and testing and that you might get
Electronic Discover Reference Model
Production

This is the stage after which all the documents have been reviewed and lawyers are ready to disclose them to opposing counsel or prepare their bundle.

This involves choosing an appropriate format for the documents to be exchanged or filed.

Typical processes would include:
- Converting documents to a usable format
- Bates Stamping (document numbering)
- Creation of List of Documents or Index
- Applying watermarks
- Choosing what metadata to display
  e.g. Custodian name, Doc ID
Microsoft Outlook

From: Mike Simonsen [msimonsen@brownmartin.com]
To: Mike Tamas
Cc: Tom Bums
Subject: FW: Wall Street Earnings Numbers
Attachments: Docs.zip

-----Original Message-----
From: Mike Simonsen [msimonsen@brownmartin.com]
Sent: Thursday, January 10, 2008 1:55 PM
To: Mike Tamas
Subject: FW: Wall Street Earnings Numbers

I need to meet with you ASAP to start discussing a legal strategy on the case regarding the failed earnings. Looks like it is getting real fast forward. Two included below an email from 2005 along with some additional attachments that may have a negative impact on our case.

Can we meet ASAP?

-----Original Message-----
From: Mike Tamas [mailto:mktamas@tamascorp.com]
Sent: Thursday, September 26, 2007 1:14 PM
To: Robert Berg
Subject: RE: Wall Street Earnings Numbers

Robert,

What is your suggestion? Should we meet tomorrow to discuss project feedback?

Thanks, Mike

-----Original Message-----
From: Robert Berg [mailto:robertberg@tamascorp.com]
Sent: Thursday, September 26, 2007 2:04 PM
To: Mike Tamas

I
Ghost In the Shell
The reality is that the lawyers are taking over due to the proliferation of data that has lead to an escalation in litigation.

These are some of the trends that we foresee developing in the near future:

- RAM forensics
- Database forensics
- Mobile device forensics (tablets)
- Network forensics
- Network appliance forensics
- Cloud forensics
- ISO standards
- Academic qualifications
- Greater court acceptance of ESD