nuix

Simple. Powerful. Precise.

nuix

Simple. Powerful. Precise.

Nuix Workshop – Introduction to Digital Forensics

ITU

CCIT

Universidad de los Andes

vive digital Colombia

TODOS POR UN NUEVO PAÍS
PAZ EQUIDAD EDUCACIÓN

MINTIC

INTRODUCTION

James Billingsley
Principal Solutions Consultant, Nuix

James has ten years experience in the field of Computer Forensics. A Certified EnCase Examiner James worked for a number of years as a senior Computer Forensics Investigator providing expert witness in UK Courts.

As part of a Security Investigation & Assessment team James worked as a senior Breach Investigation Consultant leading PCI Forensic Investigations.

As a Relativity Certified Administrator, James worked as an senior eDiscovery Consultant supporting legal reviews hosted on the Relativity platform.

James has co-authored software tools focusing on Internet Browser Forensics which are used globally by a number of law enforcement agencies and international corporations.

What is Computer/Digital Forensics?

- Computer forensics, also called cyber forensics, is the application of scientific method to computer investigation and analysis in order to gather evidence suitable for presentation in a court of law or legal body. The goal of computer forensics is to perform a structured investigation while maintaining a documented chain of evidence to find out exactly what happened on a computer and who was responsible for it.

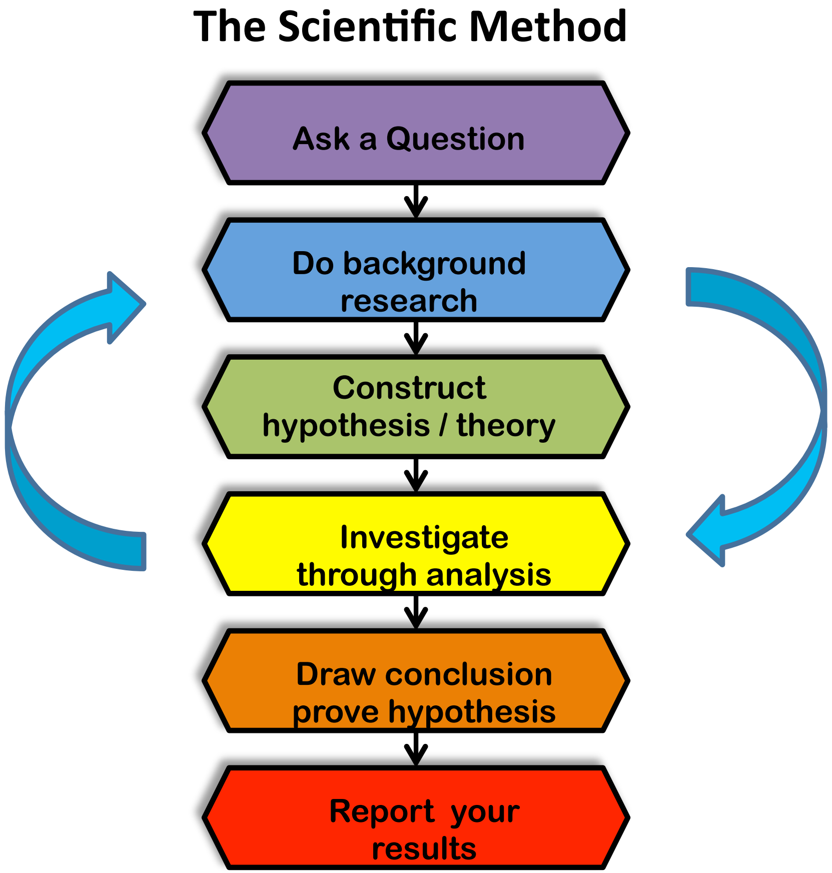
- Computer forensics has become its own area of scientific expertise, with accompanying coursework and certification.

- Computer forensics is a science and therefore requires **Scientific Method**....

SCIENTIFIC METHOD?

- The **scientific method** is a recognised body of techniques for investigating incident or occurrence, acquiring new knowledge, or correcting and integrating previous knowledge. To be termed scientific, a method of enquiry must be based on empirical and measurable evidence subject to specific principles of reasoning.

- **Scientific method** is a model applied to all areas of scientific examination. These elements are valuable to computer forensic science


The Scientific Method
Ask a Question
Do background research
Construct hypothesis / theory
Investigate through analysis
Draw conclusion prove hypothesis
Report your results

Industry Guidelines

Principle 1:

No action taken by law enforcement agencies or their agents should change data held on a computer or storage media which may subsequently be relied upon in court.

Principle 2:

In circumstances where a person finds it necessary to access original data held on a computer
or on storage media, that person must be competent to do so and be able to give evidence explaining the relevance and the implications of their actions.

Principle 3:

An audit trail or other record of all processes applied to computer-based electronic evidence should be created and preserved. An independent third party should be able to examine those processes and achieve the same result.

Principle 4:

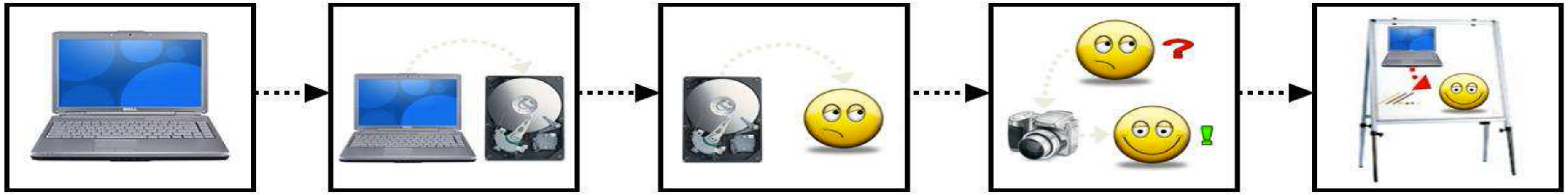
The person in charge of the investigation (the case officer) has overall responsibility for ensuring that the law and these principles are adhered to.

Source http://www.acpo.police.uk/documents/crime/2011/201103CRIECI14.pdf

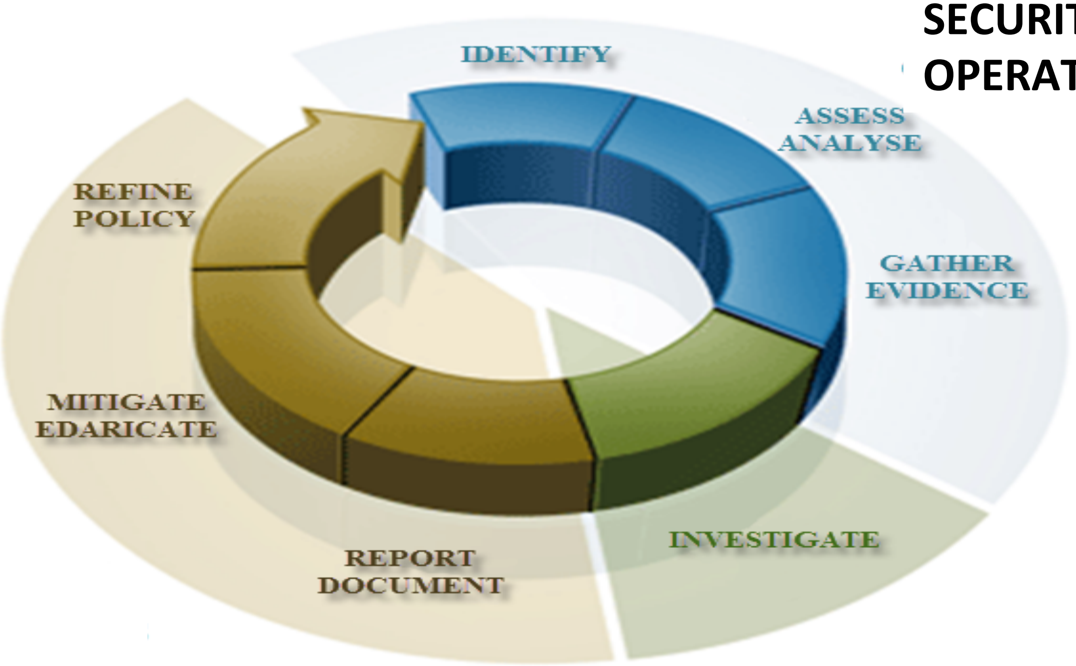
The Forensic Approach

The benefits of the application of digital forensics to computer based investigations underpin the following:

- Security of evidence / incident
- Integrity of investigative steps
- Deeper analysis unallocated space / file slack *(The whole story)*
- Auditable response
- Repeatability of action
- Best evidence practice



MANAGING SECURITY INCIDENT?



SECURITY OPERATIONS

REACTIVE

PROACTIVE

NETWORK OPERATIONS

EVIDENCE HANDLING & CHAIN OF CUSTODY

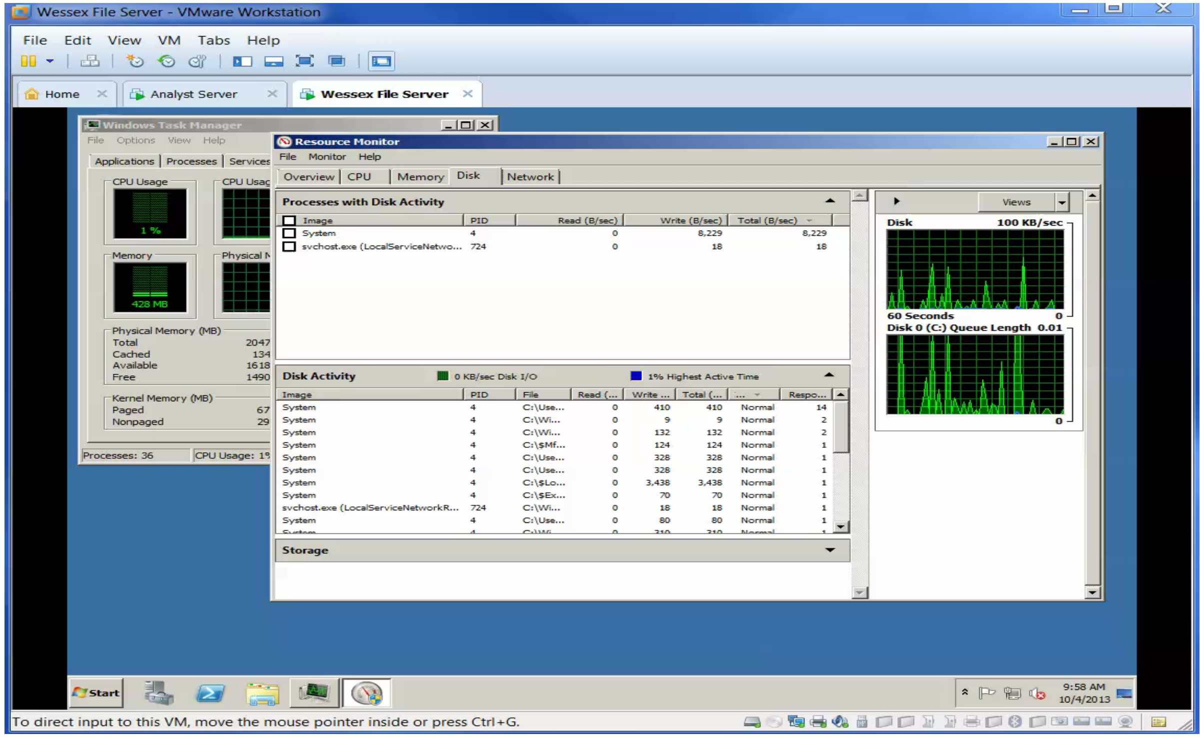
- The first step in any investigation is the search & seizure of exhibits which may contain crucial evidence!
- Decisions that you, make may result in loss of crucial evidence.
- Points to consider
  - DNA and/or fingerprints
  - Prevent tampering & preserve original condition
  - Record details & actions – paperwork!
  - Store in a secure location
  - What is capable of storing data?
  - Losing data – shutdown or not?

Doing nothing even causes changes

- Even at rest a computer is using memory and performing disk writes. This is essential to the operating system.

- The capture shows disk activity on a computer with no user activity and no applications running.

- **<u>Now consider</u>**
- Malware
- Anti forensic applications
- Cluster overwrites



Physical Memory (MB)

Total	2047
Cached	134
Available	1618
Free	1490

Kernel Memory (MB)

Paged	67
Nonpaged	29

Processes: 36 | CPU Usage: 1%

Processes with Disk Activity

☐ Image	PID	Read (B/sec)	Write (B/sec)	Total (B/sec)
☐ System	4	0	8,229	8,229
☐ svchost.exe (LocalServiceNetwo...	724	0	18	18

Disk Activity 🟩 0 KB/sec Disk I/O 🟦 1% Highest Active Time

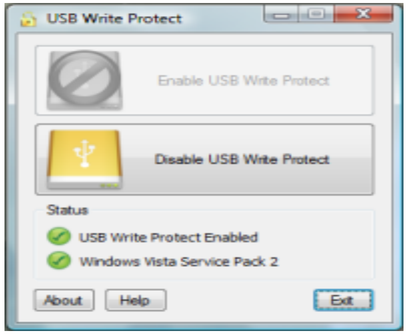
Image	PID	File	Read (...	Write ...	Total (...	...	Respo...
System	4	C:\Use...	0	410	410	Normal	14
System	4	C:\Wi...	0	9	9	Normal	2
System	4	C:\Wi...	0	132	132	Normal	2
System	4	C:\$Mf...	0	124	124	Normal	1
System	4	C:\Use...	0	328	328	Normal	1
System	4	C:\Use...	0	328	328	Normal	1
System	4	C:\$Lo...	0	3,438	3,438	Normal	1
System	4	C:\$Ex...	0	70	70	Normal	1
svchost.exe (LocalServiceNetworkR...	724	C:\Wi...	0	18	18	Normal	1
System	4	C:\Use...	0	80	80	Normal	1
System	4	C:\Wi...	0	310	310	Normal	1

Storage

COLLECTING DATA – FORENSIC METHOD

- Whether we are considering logical or physical collection we must ensure that we collect data in accordance to industry guidelines and take every step to protect the data from any change due to our action. In accordance to guidelines if this is impractical we must ensure we understand the implications of our actions.

- A write blocker is a hardware or software device that prevents ANY write activity to a connected device or resource. We can then use a forensic application or DD command to collect the data into a forensic container.

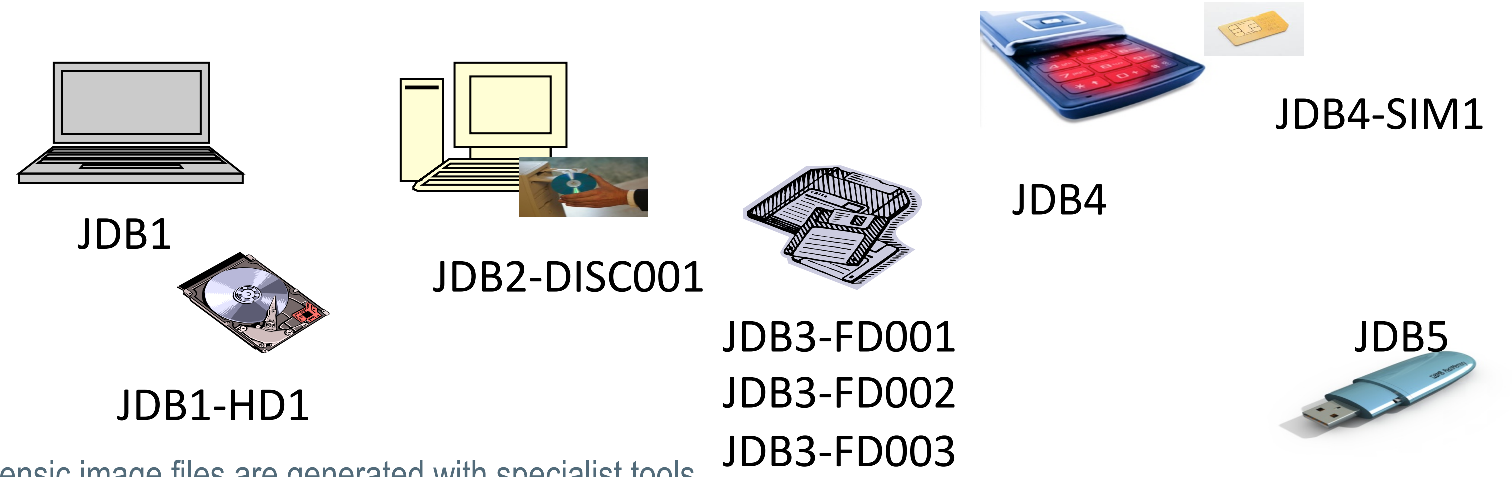

USB Write Protect

Status

USB Write Protect Enabled

Windows Vista Service Pack 2

Forensic Image Files



JDB1

JDB1-HD1

JDB2-DISC001

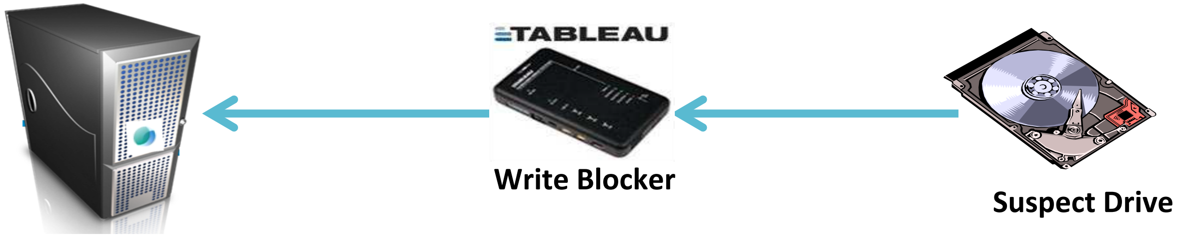
JDB3-FD001
JDB3-FD002
JDB3-FD003

JDB4

JDB4-SIM1

JDB5

- Forensic image files are generated with specialist tools
- Are an exact 'bit for bit' acquisition of the data
- All devices should be unique referenced



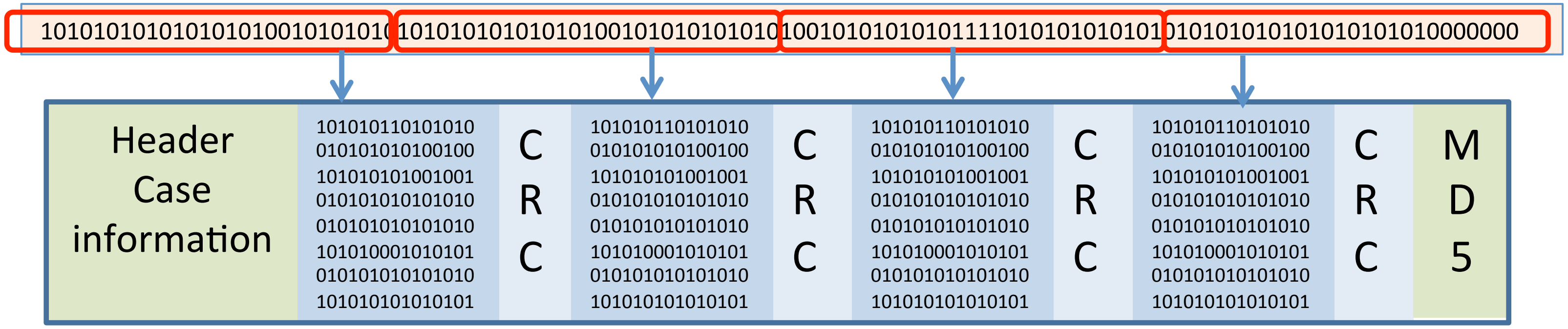
Write Blocker

Suspect Drive

Investigator

Forensic Image - container

- Data is collected from the source device at binary/disk level by pre defined size and each section is checked with a CRC checksum. The whole image is then verified with an MD5 checksum



1010101010101010101001010101(1010101010101010010101010101010|1001010101010101111010101010101|0101010101010101010101010000000

Header Case information

101010110101010
010101010100100
101010101001001
010101010101010
010101010101010
101010001010101
010101010101010
101010101010101

CRC

101010110101010
010101010100100
101010101001001
010101010101010
010101010101010
101010001010101
010101010101010
101010101010101

CRC

101010110101010
010101010100100
101010101001001
010101010101010
010101010101010
101010001010101
010101010101010
101010101010101

CRC

101010110101010
010101010100100
101010101001001
010101010101010
010101010101010
101010001010101
010101010101010
101010101010101

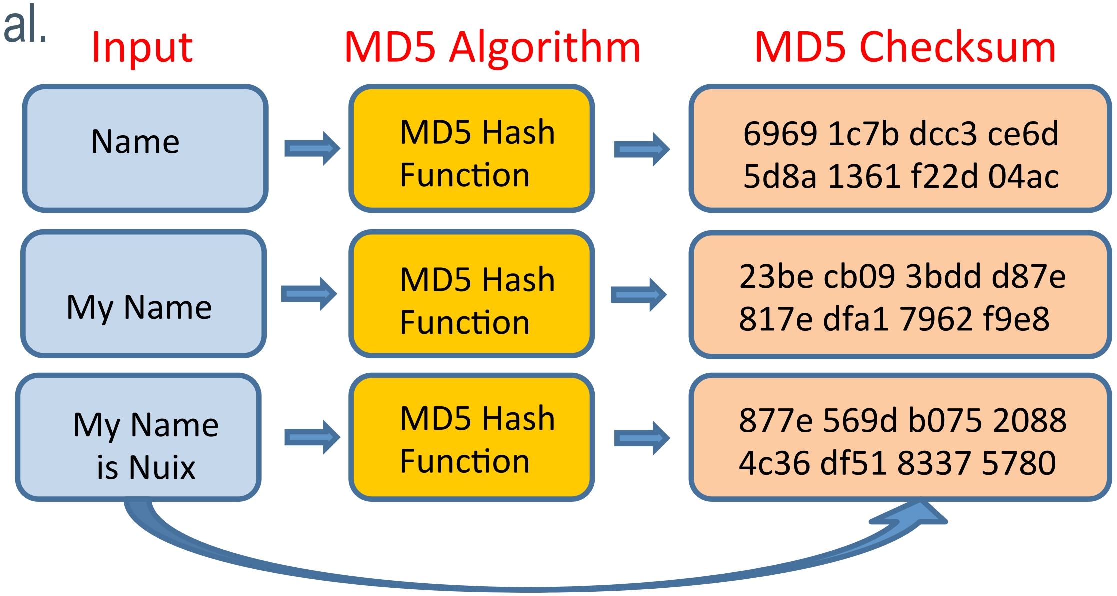
CRC

MD5

- Should any single value be change then the CRC would fail and the MD5 checksum would present a different value. Therefore verification would fail and the collection process would be undermined.

MD5 – CREATING INTEGRITY

- MD5 is an algorithm that is used to verify data integrity through the creation of a 128-bit message digest from data input (which may be a message of any length). The result is as unique to that specific data as a fingerprint is to the specific individual.



Input	MD5 Algorithm	MD5 Checksum
Name	MD5 Hash Function	6969 1c7b dcc3 ce6d 5d8a 1361 f22d 04ac
My Name	MD5 Hash Function	23be cb09 3bdd d87e 817e dfa1 7962 f9e8
My Name is Nuix	MD5 Hash Function	877e 569d b075 2088 4c36 df51 8337 5780

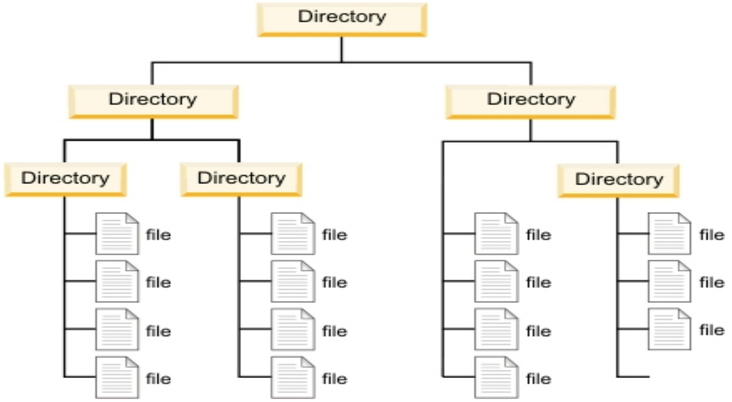
Identical data will provide identical MD5

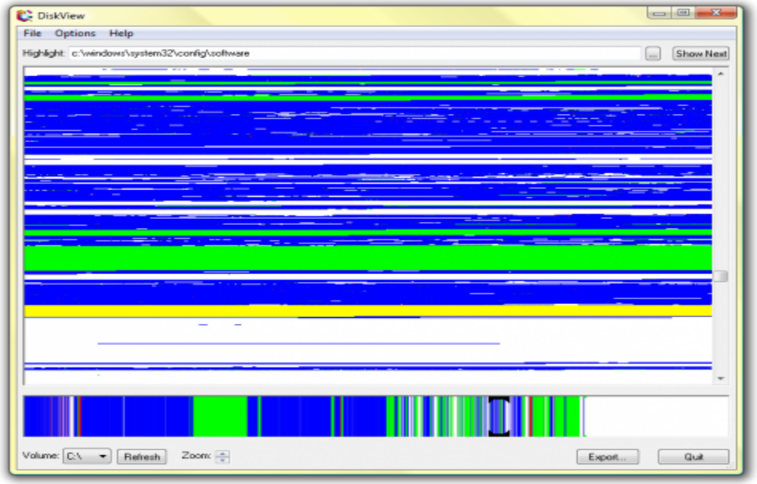
Investigations – An Intelligent approach

- Investigations frequently involve large numbers of devices including multiple computers, mobile devices and a variety of digital storage media.
- Traditional methods of analysing each data repository individually are immensely time consuming and often ineffective.
- Typical collection of devices for investigation analysis
  - ➢ Suspect's personal possessions
  - ➢ Apple Mac book Laptop (HFS+)
  - ➢ Apple iPhone (iOS)
  - ➢ External Hard Drive
  - ➢ Company/Employer data relating to suspect
  - ➢ Microsoft Windows Desktop PC
  - ➢ Microsoft Exchange Mailbox
  - ➢ Folder and files stored on a Windows Network share
  - ➢ RIM Blackberry mobile phone

- Nuix is engineered to triage, process, analyze and bring to the surface critical evidence from entire data sets.
- This saves time and effort, freeing investigators to test hypotheses, follow evidence trails and find links between suspects.

LOGICAL VS PHYSICAL





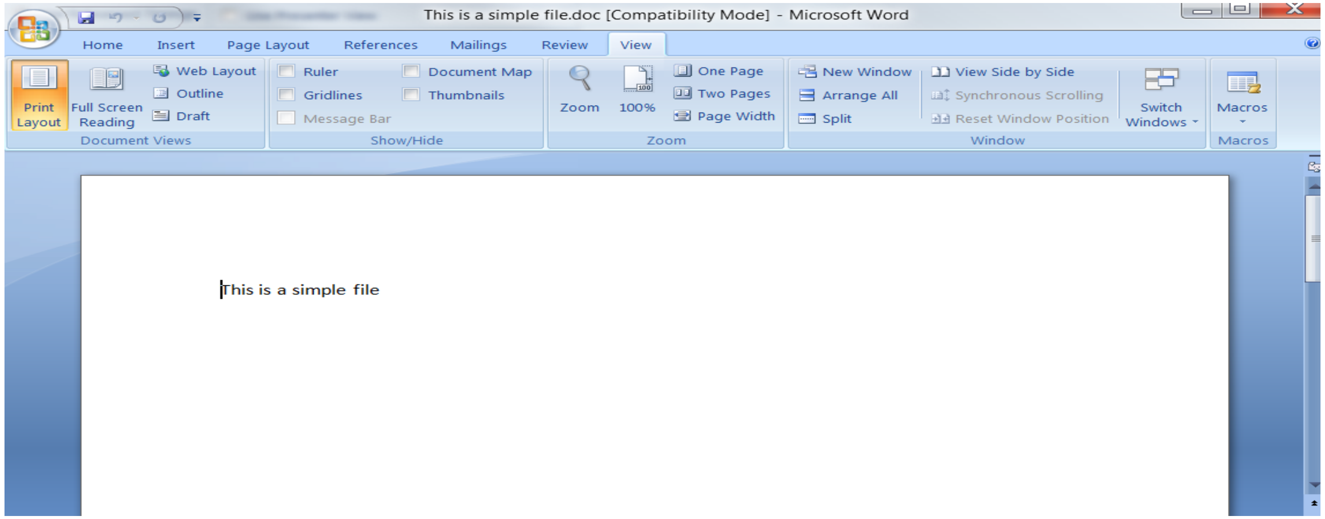
Highlight: c:\windows\system32\config\software

Logical application only allows an analyst to investigate live files and folders and whilst investigation is undertaken important attributes are changing

This Sysinternals utility will make a graphical map of a hard drive - we can see all clusters and view information about every single one of them. Capturing the data at disk level in a forensic container ensures no changes can be made to the data

WHAT'S IN A FILE?

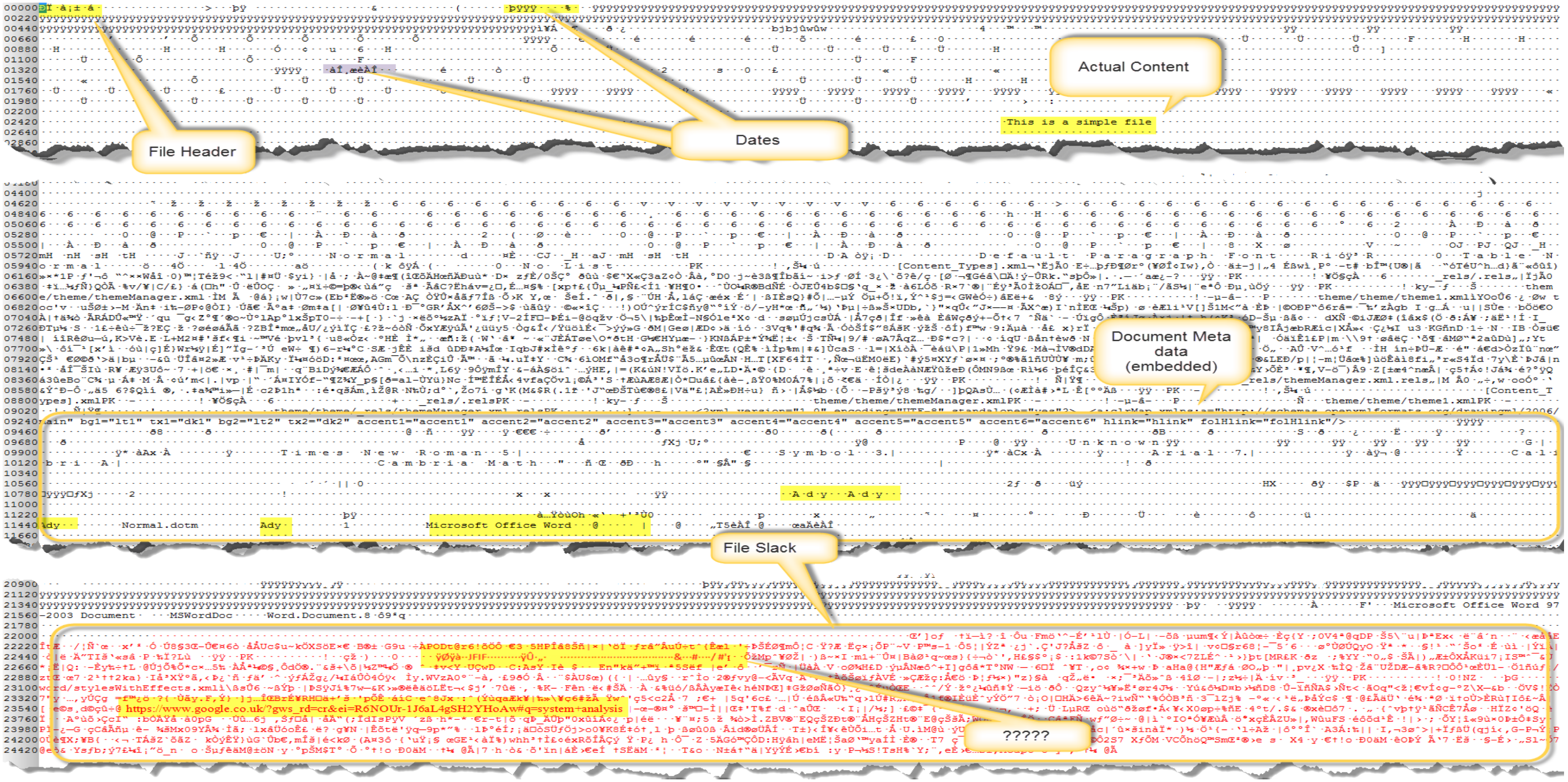
- Lets take a look at a simple word document. From a logical view we can see the content and some simple meta data



This is a simple file

- Now lets take a look at the document from a forensic image

What's in a file?



File Header

Dates

Actual Content

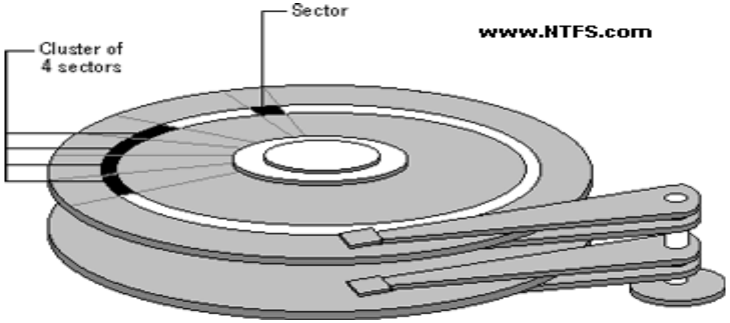
Document Meta data (embedded)

File Slack

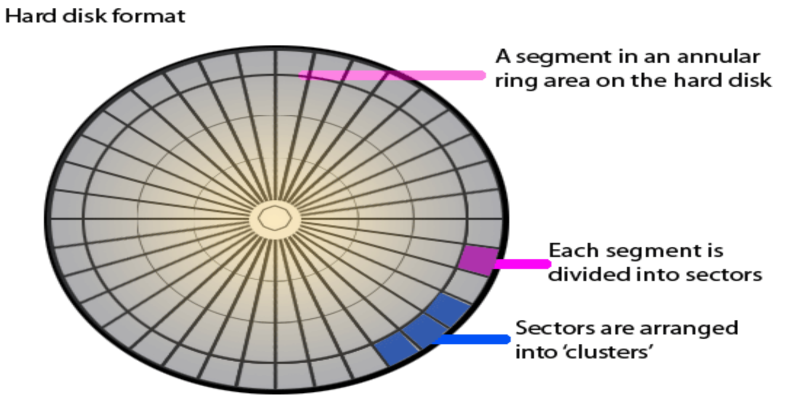
?????

File systems – how is data stored?

- To gain a better understanding of how data is recovered we must appreciate how data is stored and managed by an operating system.



www.NTFS.com

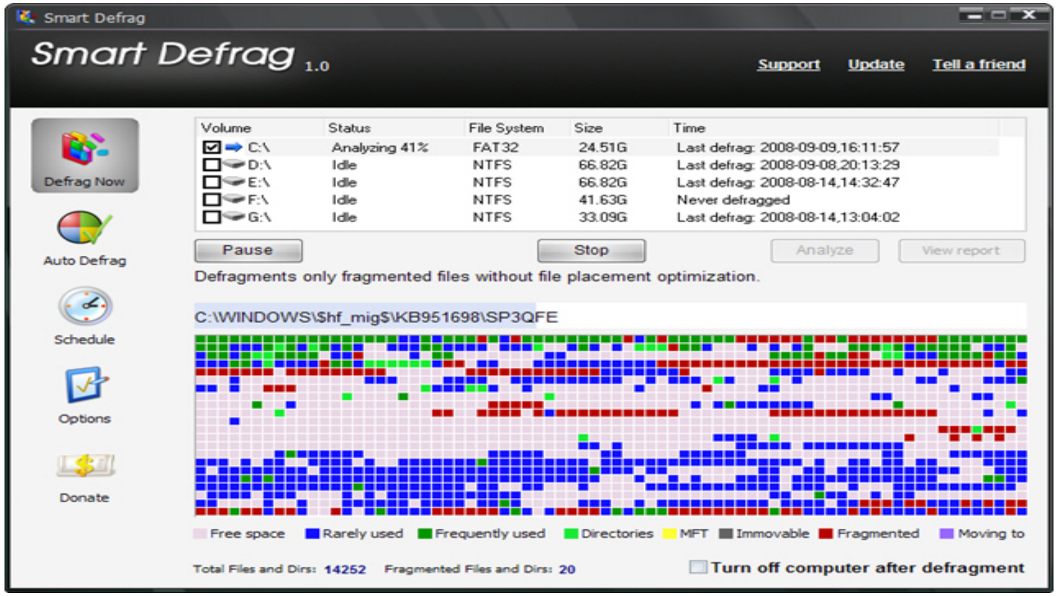

Hard disk format
A segment in an annular ring area on the hard disk
Each segment is divided into sectors
Sectors are arranged into 'clusters'

(c) www.teach-ict.com

- The above example uses a traditional single HDD however the same principle applies to other data medium e.g. USB, Solid State and RAID configuration.

FILE SYSTEMS – HOW IS DATA STORED?

- This is sometimes easier to represent and more familiar when we use an application to show the fragmentation of files across the hard drives



Volume	Status	File System	Size	Time
☑ C:\	Analyzing 41%	FAT32	24.51G	Last defrag: 2008-09-09,16:11:57
□ D:\	Idle	NTFS	66.82G	Last defrag: 2008-09-08,20:13:29
□ E:\	Idle	NTFS	66.82G	Last defrag: 2008-08-14,14:32:47
□ F:\	Idle	NTFS	41.63G	Never defragged
□ G:\	Idle	NTFS	33.09G	Last defrag: 2008-08-14,13:04:02

Defragments only fragmented files without file placement optimization.

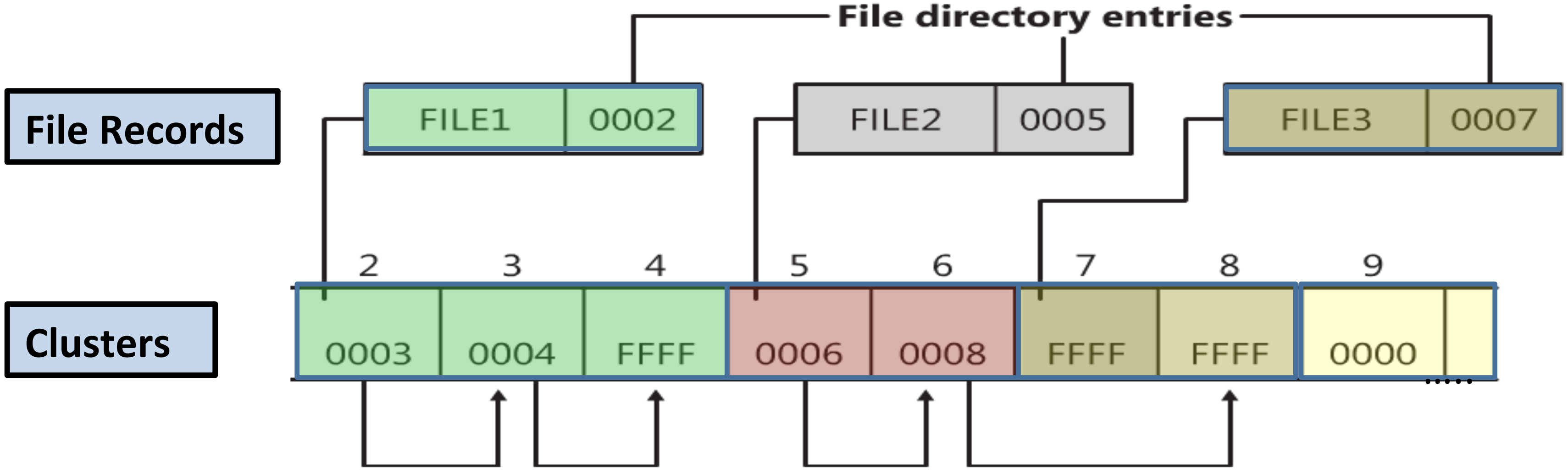
C:\WINDOWS\hf_mig\KB951698\SP3QFE

Total Files and Dirs: 14252 Fragmented Files and Dirs: 20

□ Turn off computer after defragment

File systems – how is data stored?

- The allocation of the sectors and cluster is managed by the operating system. On FAT it is the File Allocation Table and NTFS is the Master File Table. The system records much information about the files it is storing in these tables and this is referred to as Meta Data. The table records the whereabouts of all files on a system and also which clusters are available for future use.



FILE SYSTEMS – HOW IS DATA STORED?

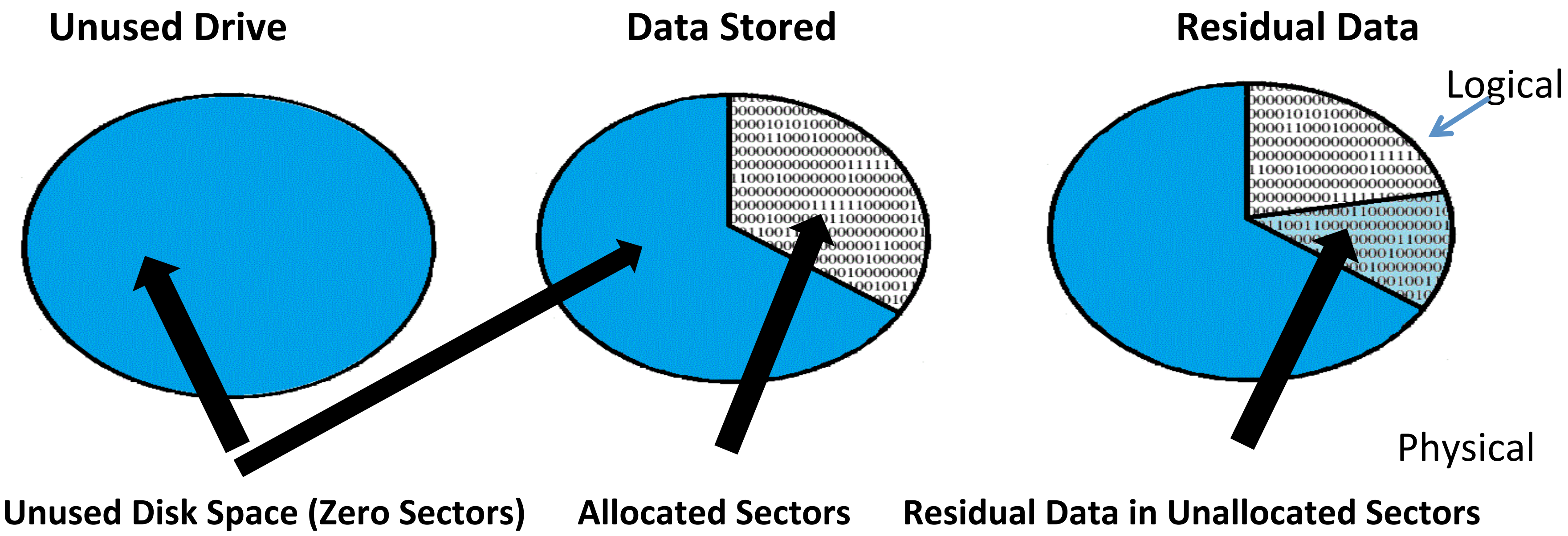
When files are erased or deleted the content of the file is not actually erased. Unless security grade file deletion software is used data from the 'erased file' remains behind in an area called unallocated storage space. The same is true concerning file slack that may have been attached to the file before it was deleted. As a result, the data remains behind for discovery through the use of data recovery and/or computer forensics software utilities.

Unallocated file space and file slack are both important sources of leads for the computer forensics investigator.

Until the first file is written to the data storage area of a computer storage device, the clusters are unallocated by the operating. As files are created by the computer user, clusters are allocated in the file table to store the data. When the file is 'deleted' by the computer user, the clusters allocated to the file are released by the operating system so new files and data can be stored in the clusters when needed. However, the data associated with the 'deleted' file remains behind. This data storage area is referred to as unallocated storage space and it is fragile from an evidence preservation standpoint. However, until the unallocated storage space is reassigned by the operating system, the data remains behind for discovery and extraction by the computer forensics specialist.

UNALLOCATED SPACE

- As data is deleted through system or user activity then more and more data becomes recoverable from unallocated sectors



Unused Drive

Data Stored

Residual Data

Logical

Physical

Unused Disk Space (Zero Sectors)

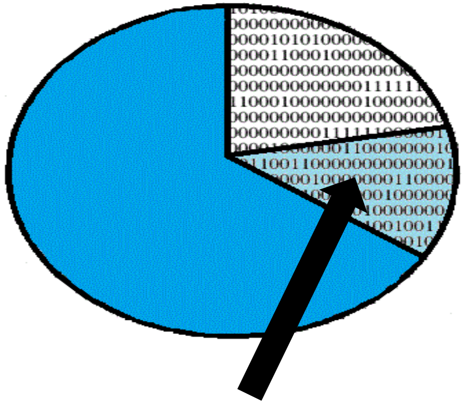
Allocated Sectors

Residual Data in Unallocated Sectors

FILE RECOVERY - CARVING

Carve file system unallocated space

- Data carving, or file carving is a process of reading files without reference to a file system. The technique can be applied to any type if disk that stores data on sector boundaries which includes camera memory, USB devices as well as hard drives. It is based on the fact that most files start with a recognisable data signature



Unallocated Sectors with residual data

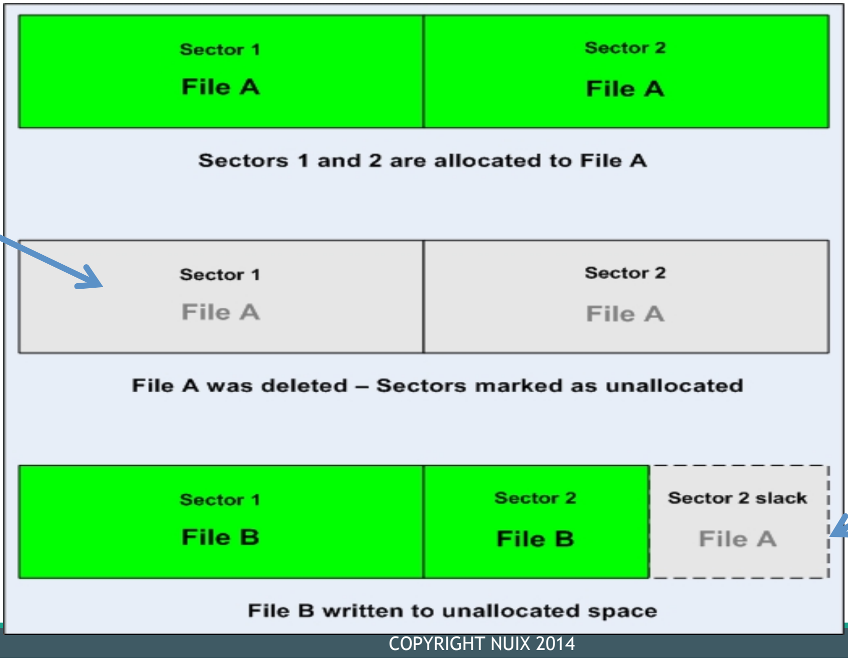
File Slack Space

Extract end of-file slack space from disk images

- The unused space in a disk cluster. The DOS and Windows file systems use fixed-size clusters. Even if the actual data being stored requires less storage than the cluster size, an entire cluster is reserved for the file. The unused space is called the *slack space.*

Actual data remains on the drive and the sectors are ready to be reused by the file system
[File entry is deleted from the file index]

New file written to sectors is smaller than file A



Sectors 1 and 2 are allocated to File A

File A was deleted – Sectors marked as unallocated

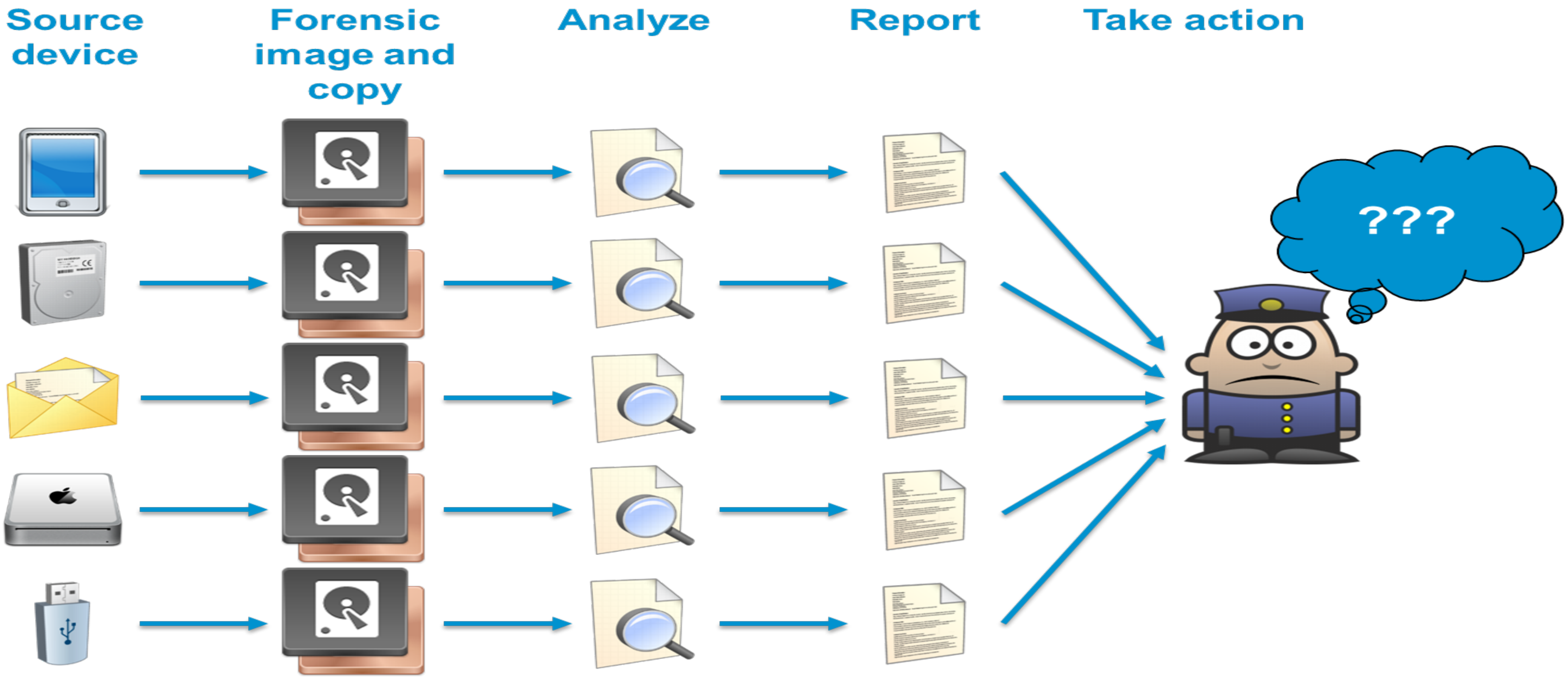
File B written to unallocated space

Sector size = 512 bytes

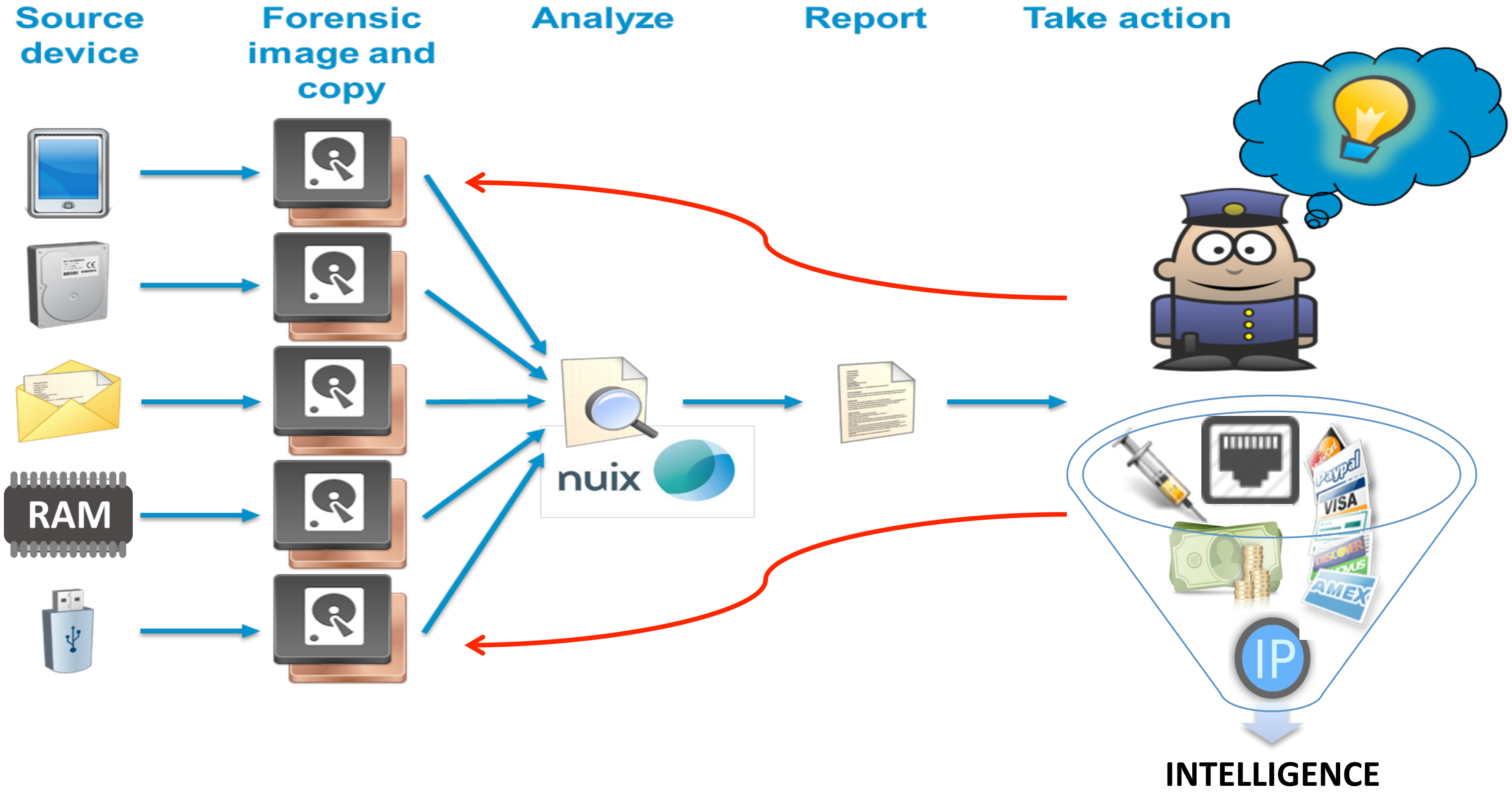
File A = 1024 bytes
File B = 750 bytes
Slack space = 274 bytes

Residual data recovered by forensic application

Current Approach To Forensic Investigation



Single window approach



The good news

- Nuix addresses all of the topics we have discussed along with many more and automates them into its process.
- Allows the user to feel assured that forensic integrity is maintained and data is presented in a format that is ready to be immediately searched and investigated.
- Lets take a look!

Why Nuix?

- Speed
  - Nuix's speed of scanning and content indexing is 10-50 times faster than competitors on comparable hardware through our fault-tolerant, parallel-processing binary indexing engine
  - It is unlikely that any organization will ever be able to bridge the gap.
- Scale
  - Nuix scales to tens of TBs per case and is able to federate searches and report across any number of databases (potentially up to petabytes).
- Scope of data
  - Extract text and metadata from 100s of file types – ranging from email files to forensic artifacts, log files, windows registry, etc…
- Simplicity
  - Nuix is simple to install and use and can be up and running in minutes, allowing incident responders to focus on chase more quickly.
- Internationalization
  - Fully I18N compliant from day 1 – localized UI for Chinese, Japanese, Arabic, Spanish, Portuguese, etc..

Extract text and metadata from 100s of different file types

Email & Loose Files

Microsoft:

- EDB, STM, EWS (Microsoft Exchange)
- PST,OST (Microsoft Outlook storage files)
- MSG (Microsoft Outlook single mail files)

Lotus:

- NSF (Lotus Notes / Domino)

Misc. Other:

- MBOX, DBX, MBX (Microsoft Outlook Express)
- EML, EMLX, BOX, SML
- Webmail – HTML scraped from browser cache

Document Types:

- HTML , Plain text, RTF, PDF
- DOCX, DOC, DOT (Microsoft Word)
- XLSX, XLS,XLT (Microsoft Excel)
- PPTX, PPT,POT,PPS (Microsoft PowerPoint)
- WKS,XLR (Microsoft Works spreadsheets)

Image Types:

- PNG, JPEG, JP2, TIFF, GIF, BMP, PBM, PPM, PGM, RAW, WBMP, WMF, WMZ, EMF, EMZ

Incident Response

Forensic Image Files:

- Encase Images (E01, L01)
- Access Data (AD1)
- Linux DD Files
- Mobile Images (Cellebrite / XRY / Oxygen)

Log Files:

- Windows Event Logs (EVT/EVTX)
- Web Logs (IIS, Apache, FTP)

Network Captures:

- PCAP Files

System Files:

- EXE/DLLs
- LNK Files
- Windows Registry Hives

File System Artifacts:

- $LogFile, $UserJrml, Object ID
- File slack space
- Carved, unallocated blocks

Structured Data:

- MS SQL (Live & MDF/LDF are text stripped)
- SQLite

Misc.

Browser Artifacts:

- IE, Safari, Chrome, Firefox

Container Files

- ZIP, RAR, LZH, LHA, ARC, TAR, GZ, BZ2, ISO

Virtual Machine Images

- VDK, VMDK (Virtual Disk Images)
- Parallels

Archive Systems

- EMC EmailXtender (*.emx)/Source One
- Symantec 2007, 8, 9, 10
- HP EAS

DMS Systems:

- MS SharePoint

Unknown File Types:

- Unknown file types are text stripped.

NUIX SEARCHABLE INDEXES

- As Nuix processes data it extracts valuable information and places the data into separate searchable indexes which can be searched against in whole or individually.

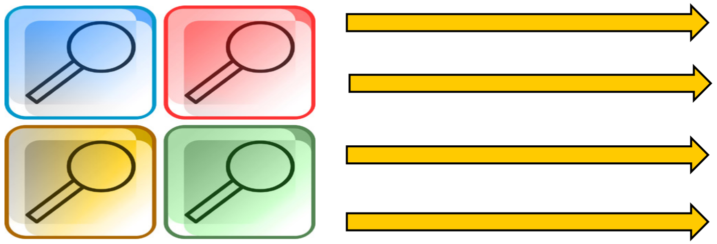
- Powerful dynamic component that allows investigators to be flexible and intuitive in the approach to data management from ECA through analysis.

- Allows for the application of specific function to relevant data throughout workflow.

- Enable investigators to quickly target and hydrate function to relevant material through ECA and NVA.

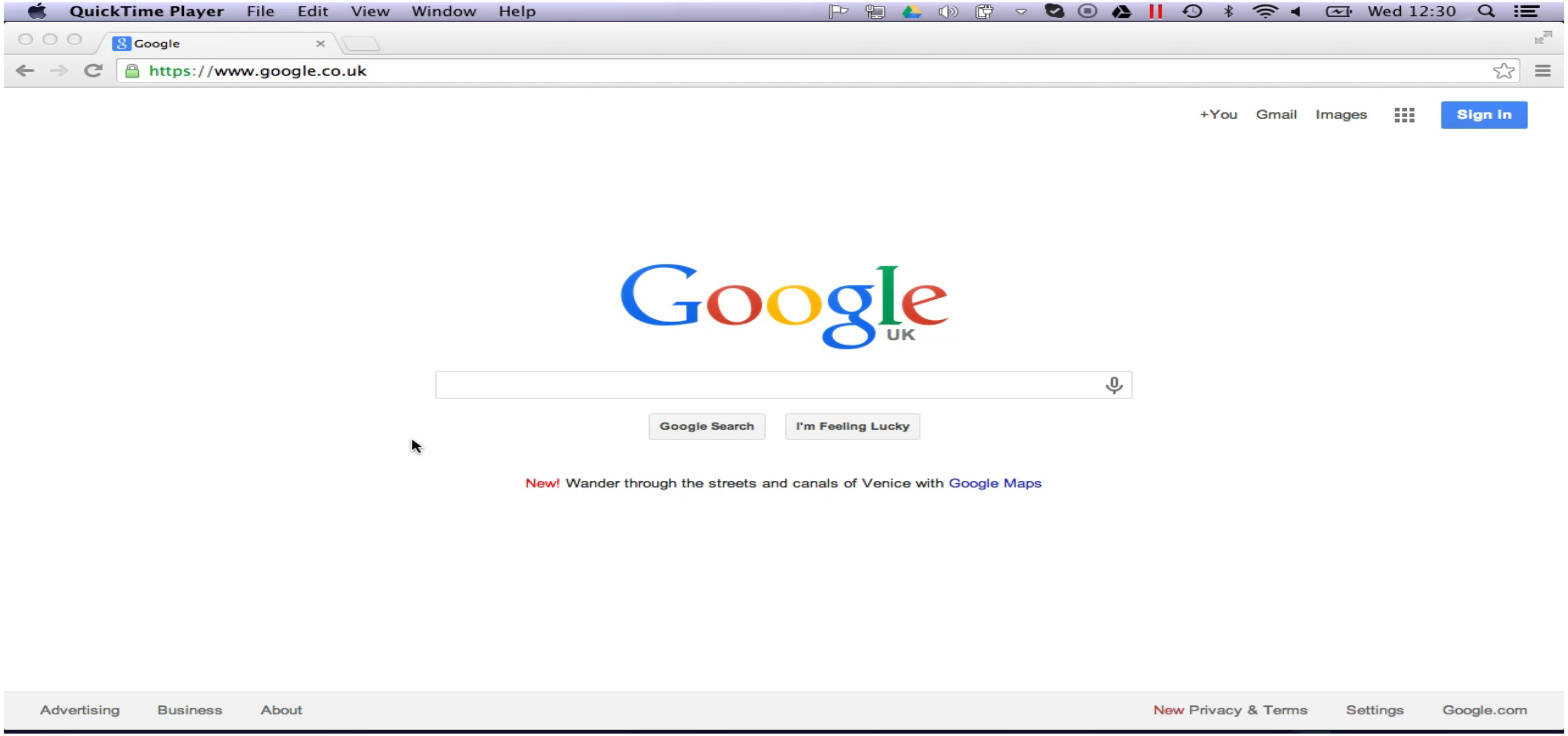
Nuix Indexes

- The database architecture of Nuix offers the investigator powerful options in order to get to relevant data very quickly and decreasing false hits from search criteria.

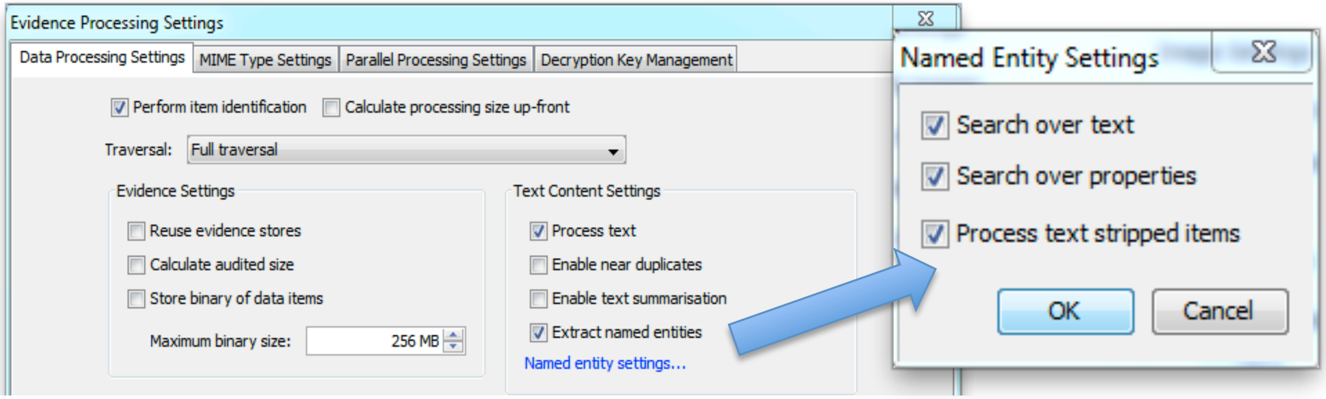


- We can draw comparison to a well known entity that we use every day - Google

Smart Searching



Extracted Entities



Evidence Processing Settings

Data Processing Settings | MIME Type Settings | Parallel Processing Settings | Decryption Key Management

☑ Perform item identification ☐ Calculate processing size up-front

Traversal: Full traversal

Evidence Settings

☐ Reuse evidence stores

☐ Calculate audited size

☐ Store binary of data items

Maximum binary size: 256 MB

Text Content Settings

☑ Process text

☐ Enable near duplicates

☐ Enable text summarisation

☑ Extract named entities

Named entity settings...

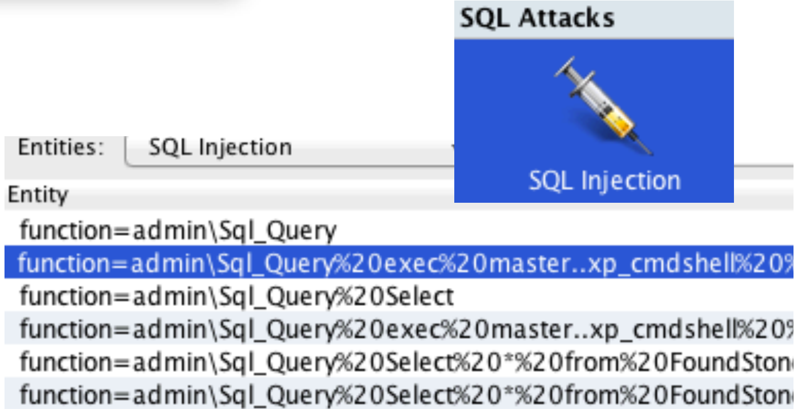
Named Entity Settings

☑ Search over text

☑ Search over properties

☑ Process text stripped items

- Further analysis opportunity
  - Most IR analysis is GREP
  - RAM dumps
- Controlled by mime type



SQL Attacks

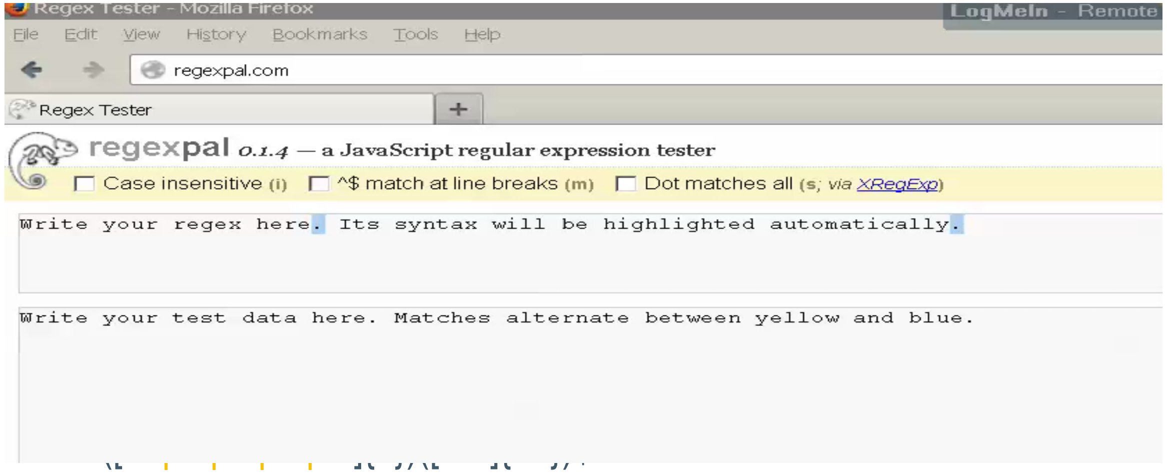
Entities: SQL Injection

Entity

```
function=admin\Sql_Query
function=admin\Sql_Query%20exec%20master..xp_cmdshell%20%
function=admin\Sql_Query%20Select
function=admin\Sql_Query%20exec%20master..xp_cmdshell%20%
function=admin\Sql_Query%20Select%20*%20from%20FoundSton
function=admin\Sql_Query%20Select%20*%20from%20FoundSton
```

Using regular expression to find data

- This is a simple example of how the use of Regular expression can locate data. All Mastercard numbers start with a 51,52,53,54, or 55 followed by 14 numbers. This rule can be applied to a regular expression as follows:


Regex Tester - Mozilla Firefox
LogMeIn - Remote
File
Edit
View
History
Bookmarks
Tools
Help
regexpal.com
Regex Tester
+
regexpal 0.1.4 — a JavaScript regular expression tester

□ Case insensitive (i) □ ^$ match at line breaks (m) □ Dot matches all (s; via XRegExp)

```
Write your regex here. Its syntax will be highlighted automatically.
```

```
Write your test data here. Matches alternate between yellow and blue.
```

- The number must start with 2 digits that are defined in the expression followed by 14 digits that are between 0 and 9.

Online tester http://regexpal.com/

nuix

FIND OUT MORE:

- nuix.com
- twitter.com/nuix
- facebook.com/nuixsoftware
- linkedin.com/company/nuix
- youtube.com/nuixsoftware
- nuix.com/blog