Investigazioni ed Analisi Forense

Clusit
Education

Relatore

- Ing. Pasquale Stirparo
  - Digital Forensics Engineer @PSS Srl
  - Consulente per Polizia, Carabinieri, Guardia di Finanza
  - GCFA, ECCE, OPST, OWSE
  - Socio IISFA
  - Socio CLUSIT

Digital Forensics

L'uso di metodi scientifici (identificazione, raccolta, validazione, preservazione, analisi, interpretazione, documentazione e presentazione delle evidenze digitali derivate da "fonti digitali") che hanno lo scopo di facilitare la ricostruzione di azioni non autorizzate, dannose o di eventi criminali.

Digital Forensics

- Inizialmente la Digital Forensics è stata usata solo per i crimini tecnologici.
  - Intrusioni informatiche;
  - Web defacement;
  - Danneggiamento/Furto di dati;
  - Pedofilia online;
- Negli altri casi i computer sono stati *semplicemente ignorati*.

Digital Forensics

- Alcuni casi “non informatici” che abbiamo risolto negli ultimi anni:
  - **Frode telefonica:** analisi di dispositivi “GSM-box”-like al fine di individuare il Modus Operandi tecnologico.
  - **Spionaggio Industriale:** supporto ad azienda nella risoluzione e conseguenti azioni in Tribunale (furto di disegni e progetti industriali).
  - **Antipedofilia digitale:** analisi di evidence elettroniche a supporto dell’AA.GG., verso PC e smartphone.

Digital Forensics

- È quindi lampante come l'analisi delle evidenze digitali si rende **necessaria** anche per **crimini che nulla hanno a che fare con la tecnologia**.
- Dal **palmare della Lioce** al **delitto di Garlasco**…sino agli **atti di bullismo su Facebook** ed altre cronache recenti.
- Non sono stati portati all'attenzione del grande pubblico **molti altri casi**, risolti per merito delle evidenze digitali.

Digital Forensics

- Adesso la Digital Forensics “è di moda”!!!

- Questo è **un bene** in quanto vi è:
  - Maggiore scambio di informazione;
  - Nuovi tool e nuove tecnologie;
  - Un più rapido sviluppo;
  - Una Maggiore sensibilità al problema;

Digital Forensics

- Questo è **un male** perché:
  - Tutti vogliono lanciarsi in questo mercato;
  - Ci sono molti “presunti esperti”, improvvisati e molto spesso privi dei necessari skill, strumenti, laboratori ed esperienza sul campo;
  - Tutti promettono tool “facili da usare”;
  - Il fatto di scrivere “forensics” su un programma di 10 anni fa non lo rende necessariamente più adatto allo scopo.

La Metodologia

Physical vs Digital Evidence

- **Physical Evidence:** *Physical objects that can establish that a crime has been committed, can provide a link between a crime and its victim, or can provide a link between a crime and its perpetrator.*
- **Digital Evidence:** *Digital data that can establish that a crime has been committed, can provide a link between a crime and its victim, or can provide a link between a crime and its perpetrator.*

Physical vs Digital Evidence

- **Physical Crime Scene:** *The physical environment where physical evidence of a crime or incident exists.*

- **Digital Crime Scene:** *The virtual environment created by software and hardware where digital evidence of a crime or incident exists.*

B. Carrier, E. H. Spafford, "Getting Physical with the Digital Investigation Process", International Journal of Digital Evidence, Fall 2003.

Approccio alla scena del crimine

La Metodologia

- Assenza di uno standard comune “compensato” dalla presenza di *Best Practice* riconosciute a livello internazionale;
- Dal 2009 sono iniziati i lavori per sviluppare delle linee guida standard dall’ISO/IEC
- Famiglia ISO/IEC 27000 *"Information technology - Security techniques - Information security management systems - Overview and vocabulary*

La Metodologia

- Assenza di uno standard comune “compensato” dalla presenza di *Best Practice* riconosciute a livello internazionale;
  - Purtroppo, molti ne ignorano ancora l’esistenza;
- Dal 2009 sono iniziati i lavori per sviluppare delle linee guida standard dall’ISO/IEC
- Famiglia ISO/IEC 27000 *"Information technology - Security techniques - Information security management systems - Overview and vocabulary*

ISO27037

- *Guidelines for identification, collection and/or acquisition and preservation of digital evidence;*

- Copre diversi scenari, dall'acquisizione live a quella classica, dai computer ai dispositivi mobili, dall'identificazione dell'evidenza alla catena di custodia;

Principi della Digital Forensics

- Ci sono quattro principi basilari che bisogna sempre tenere a mente:
  1. Ridurre al minimo la perdita di dati;
  2. Annotare sempre tutto;
  3. Analizzare tutti i dati e le informazioni raccolte;
  4. Presentare quanto rinvenuto.

Principi della Digital Forensics

- Due macro fasi:
  - Incident Handling & Evidence Gathering;
  - Investigation and Analysis.
- Le fasi principali possono essere raggruppate in quattro gruppi:
  - Identificazione;
  - Acquisizione e preservazione;
  - Analisi;
  - Reportizzazione.

Due possibili scenari

- Dead System
  - Assenza di alimentazione elettrica;
  - Spegnimento del sistema
  - Hdd, floppy disk, CD/DVD;

- Live System
  - Sistema attivo;
  - Processi attivi lanciati nel sistema;
  - Connessioni di rete attive;
  - Etc…

Aree d'interesse

- La Digital Forensics si divide in tre aree:

1. Computer Forensics

2. Network Forensics

3. Mobile Forensics

Riassumendo

Quattro principi
1. Ridurre al minimo la perdita di dati;
2. Annotare sempre tutto;
3. Analizzare tutti i dati e le informazioni raccolte;
4. Presentare quanto rinvenuto.

Tre aree
1. Computer Forensics
2. Network Forensics
3. Mobile Forensics

Digital Forensics

Due scenari possibili
1. Live System
2. Dead System

Quattro fasi principali
1. Identificazione
2. Acquisizione e preservazione
3. Analisi
4. Reportizzazione

Fase 1:
Identificazione delle evidenze

Sistemi informatici coinvolti

- Siamo letteralmente invasi da:
  - Cellulare/Smartphone;
  - Computer;
  - Palmare;
  - Accesso ad Internet (oramai a banda larga sia a casa sia via mobile);
  - Notebook/Netbook ;
  - GPS;
  - Macchina Fotografica digitale.

Individuazione del dato

- Il dato digitale, per sua natura immateriale, può essere ritrovato sul campo in soli tre modi diversi:
  - Sequestrato
  - Copiato
  - Intercettato
- Qualunque altra situazione può essere ricondotta a una di queste tre.

Individuazione del dato: Sequestro

- Il dato si sequestra nel momento in cui lo si rimuove dalla *scena criminis* sul supporto (o sistema informatico) nel quale è stato trovato.

Individuazione del dato: Sequestro

- Talvolta si sequestra l'intero PC;
- Talvolta solo il supporto dove sono memorizzate le informazioni;
- È il metodo più semplice per ottenere il dato. Non richiede complesse operazioni di validazione;

Individuazione del dato: Sequestro

- Basta curare la parte “documentale”;
- Il supporto (sistema) deve essere sigillato opportunamente!!!
- Qualunque problema in fase di sequestro si tradurrà in eccezioni da parte della difesa in fase di dibattimento.

E se il sequestro non è possibile?

Individuazione del dato

- Acquisizione, diversi scenari possibili:
  - Acquisizione di un sistema “live”
  - Acquisizione di un sistema “dead”
- Intercettazione
  - Il dato non è legato ne ad un sistema di memorizzazione (e.g. hard disk), ne ad un sistema di elaborazione (e.g. computer)

Fase 2:
Acquisizione e Preservazione

Agenda

- **Acquisizione di un sistema Live**
- Acquisizione di un sistema “dead”
- Network forensics: intercettazione dei dati in transito
- Preservazione delle evidenze: gli algoritmi di Hash
- Catena di Custodia

Acquisizione Live

- L’acquisizione di un sistema live viene effettuata, come suggerisce il nome, mentre il sistema è ancora in funzione;

- Questo può dipendere da diversi fattori:
  - Tipologia di servizio offerto dal computer da acquisire;
  - Il computer è attivo al momento del sequestro.

Acquisizione Live

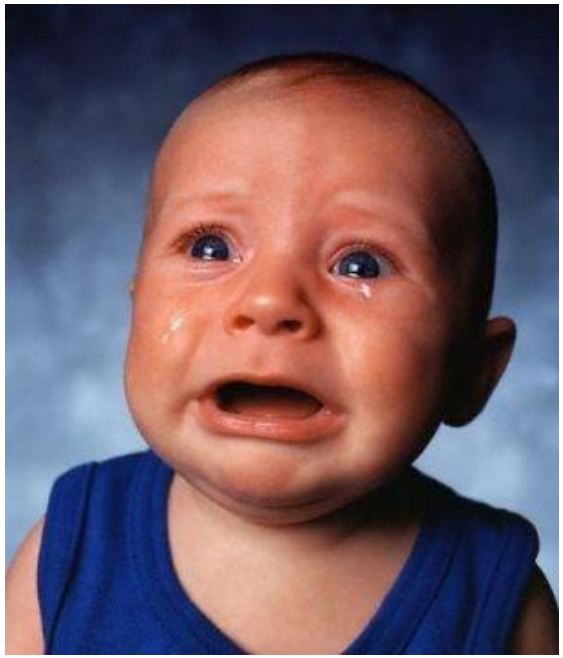
- Tuttavia quando si è di fronte al secondo scenario, spesso la prima operazione che viene fatta è quella di spegnere il computer;
- Questo comporta la perdita di numerose informazioni, le cosiddette *“evidenze volatili”*.
- Per *evidenze volatili* si intendono tutte quelle informazioni che svaniscono nel momento in cui il sistema viene spento;

Acquisizione Live

- Ordine di volatilità:
  - Memoria
  - Processi attivi
  - Connessioni attive
  - Utenti loggati
  - System Time
  - …

Acquisizione Live

- Tuttavia quando si è di fronte ad un sistema accesso, spesso la prima operazione che viene fatta è quella di spegnere il computer;



Acquisizione Live

Case Study[1]

- Nell'Ottobre 2002, sono state ritrovate nel computer Julian Green oltre 170 immagini illecite;

- Da un analisi del computer vengono ritrovati diversi trojan che accedevano a siti web "vietati" ogniqualvolta egli apriva il browser

[1] *Windows Forensics Analysis, 2nd Edition – Harlan Carvey*

Acquisizione Live

- L’anno seguente Aaron Caffrey dichiara che dei trojan presenti nel suo computer hanno permesso a terzi di lanciare gli attacchi verso altri sistemi di cui è stato accusato;

- Sebbene nessun trojan sia stato trovato durante l’analisi forense, non si può escludere quanto affermato da Caffrey.

Agenda

- Acquisizione di un sistema Live
- **Acquisizione di un sistema “dead”**
- Network forensics: intercettazione dei dati in transito
- Preservazione delle evidenze: gli algoritmi di Hash
- Catena di Custodia

Acquisizione dead

- È l'acquisizione "classica" che avviene a sistema spento;

- Diverse possibilità:
  - Live CD
    - Via rete
    - Via disco USB/eSata
  - Estrazione disco rigido
    - Macchina di acquisizione di supporto
    - Write-blocker

Acquisizione

- Confronto Live vs Dead
  - Quando conviene l'una piuttosto che l'altra
  - Vantaggi
  - Rischi

Agenda

- Acquisizione di un sistema Live
- Acquisizione di un sistema “dead”
- **Network forensics: intercettazione dei dati in transito**
- Preservazione delle evidenze: gli algoritmi di Hash
- Catena di Custodia

Network Forensics

- Quando i dati che ci interessano sono in transito su una rete, è necessario “intercettarli” in qualche modo;
- Diversi tipi di reti di trasmissione dati:
  - Reti di computer
  - Reti cellulari
  - Reti satellitari
  - Etc..

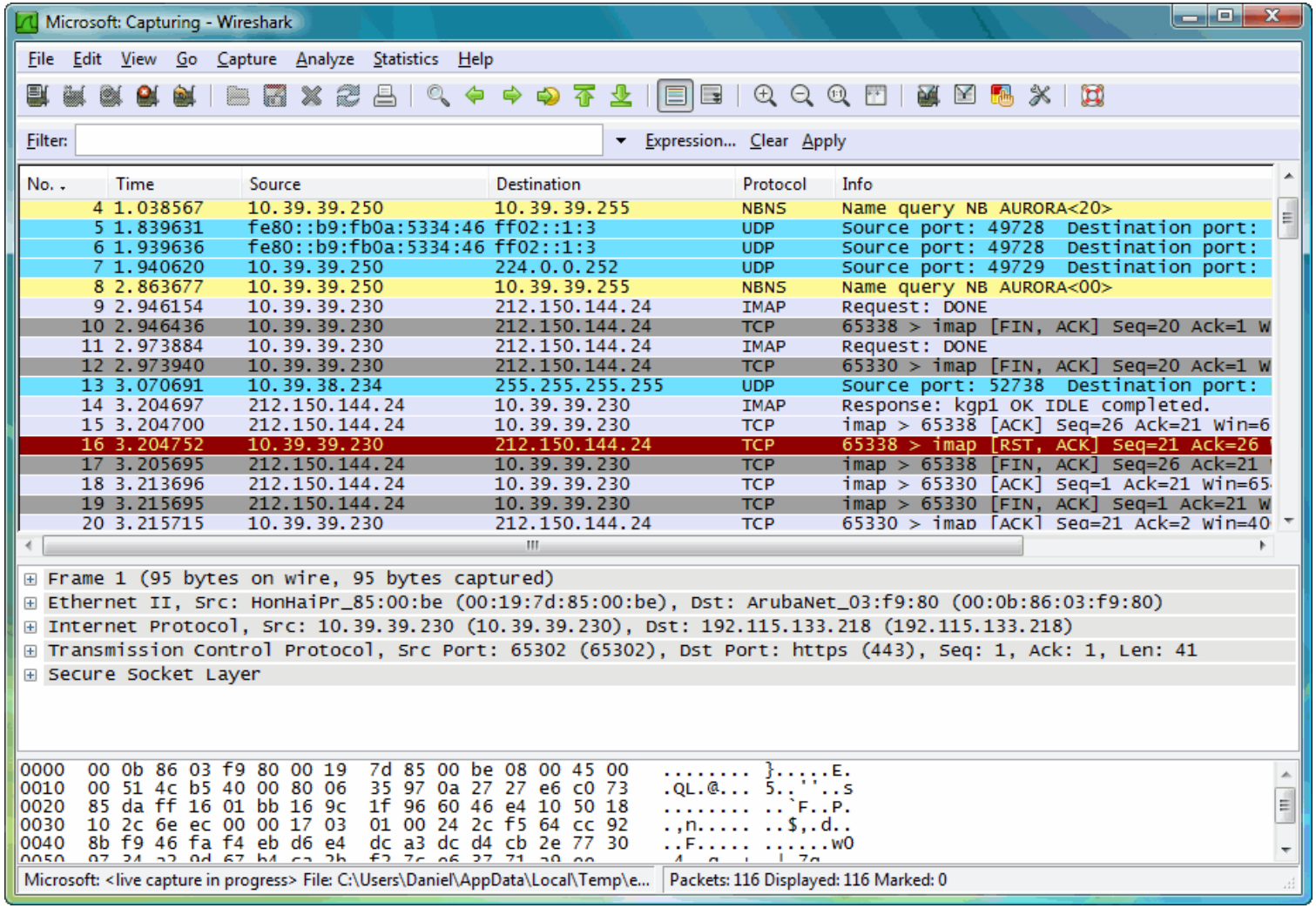
Network Forensics

- Quali dati transitano sulla rete:
  - Voce
  - Video
  - Dati
  - Navigazione Internet
  - Chat
  - Posta elettronica
  - Password
  - Etc…

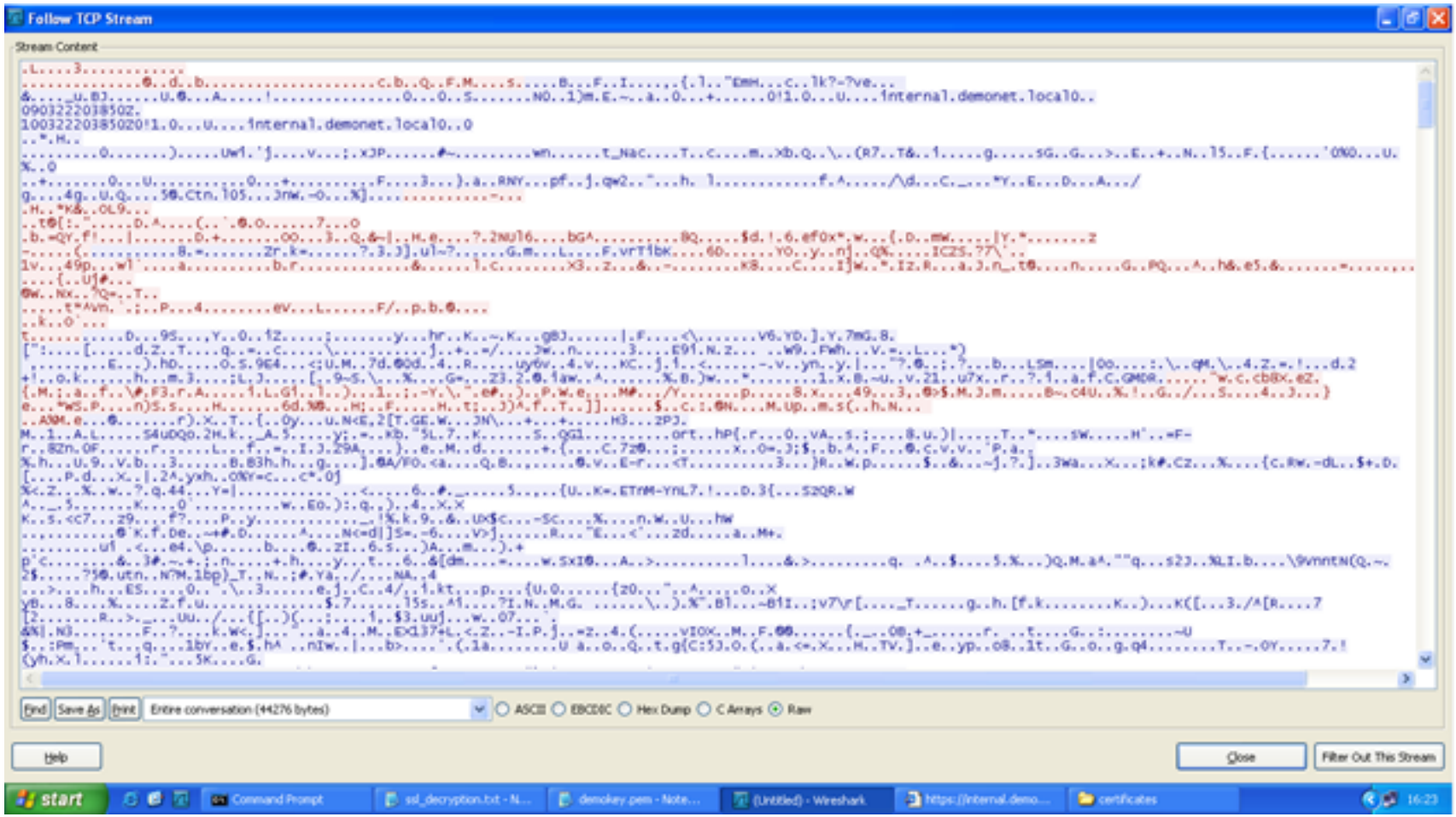
Network Forensics

- Problematiche:
  - Elevato flusso del traffico;
  - Interpretazione e decodifica del traffico (cifrato);
- Strumenti utilizzati:
  - Sonde “artigianali”;
  - IDS con regole personalizzate;
- Spesso sonda e IDS utilizzati insieme.

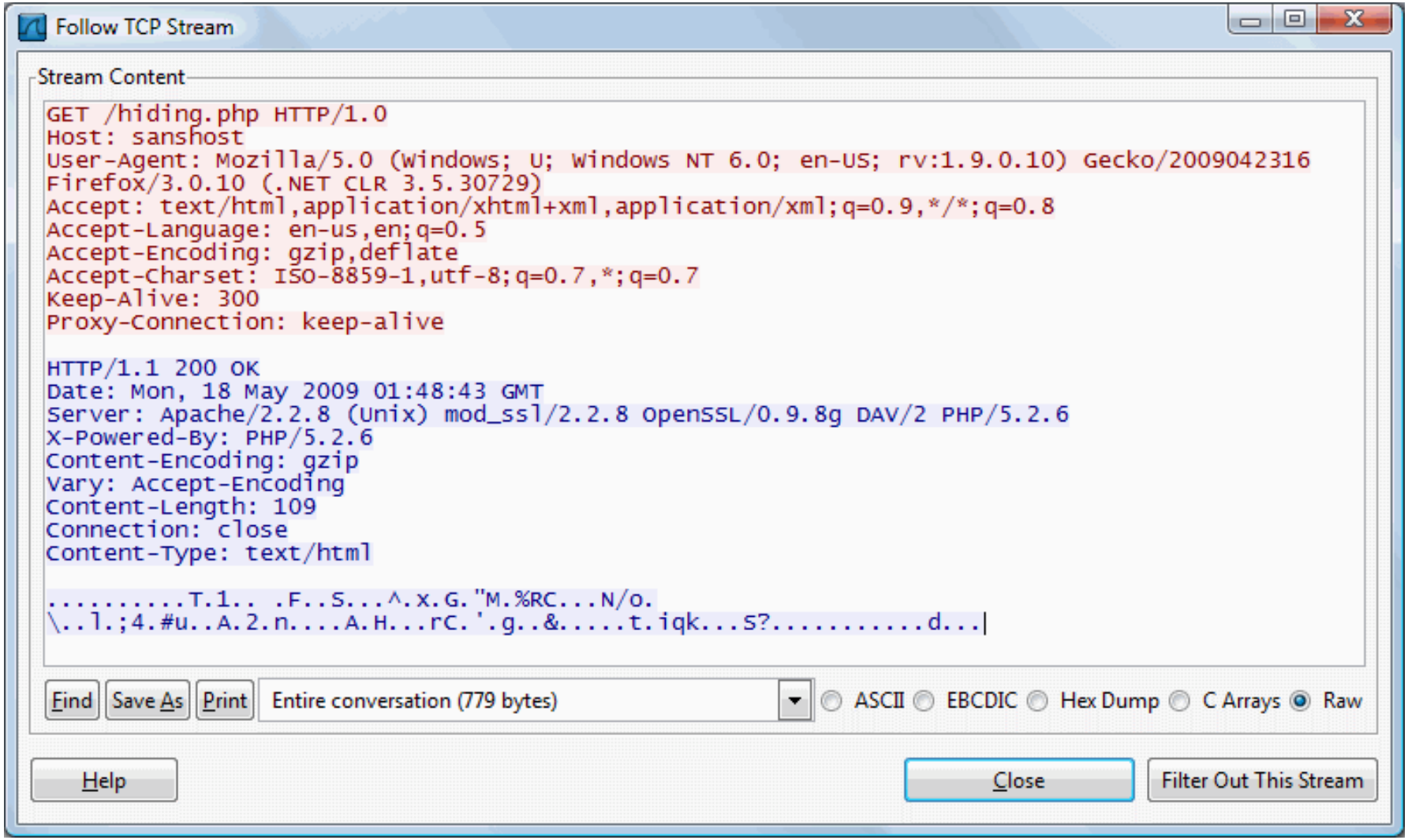
Network Forensics



Network Forensics



Network Forensics



Network Forensics

http://www.philosophyblog.com.au/images/privacy-people-eat-the-darndest-things1.jpg

Agenda

- Acquisizione di un sistema Live
- Acquisizione di un sistema “dead”
- Network forensics: intercettazione dei dati in transito
- **Preservazione delle evidenze: gli algoritmi di Hash**
- Catena di Custodia

Preservazione

- Bisogna tener presente che è necessario (ripeto necessario) utilizzare un sistema per dimostrare non solo che i dati sono stati copiati correttamente ma anche che non sarà possibile modificarli nel corso dell'indagine;
- Confrontare a mano bit per bit richiederebbe qualche decennio!!!
- Funzioni di hash… uno strumento utile.

Preservazione

- *L'hash è una funzione operante in **un solo senso** (ossia, che non può essere invertita), atta alla trasformazione di un testo di lunghezza arbitraria in una stringa di lunghezza fissa, relativamente limitata. Tale stringa rappresenta una sorta di "**impronta digitale**" del testo in chiaro, e viene detta **valore di hash**, **checksum crittografico** o **message digest**;*

Preservazione

- Proprietà utili delle funzioni di hash:
  - Irreversibile;
  - Qualunque sia la grandezza del file in input:
    - Facile da calcolare;
    - Output sempre di lunghezza fissa;
  - Probabilisticamente è molto difficile che due file di input arbitrari, per quanto simili, possano dare lo stesso valore di hash in output,

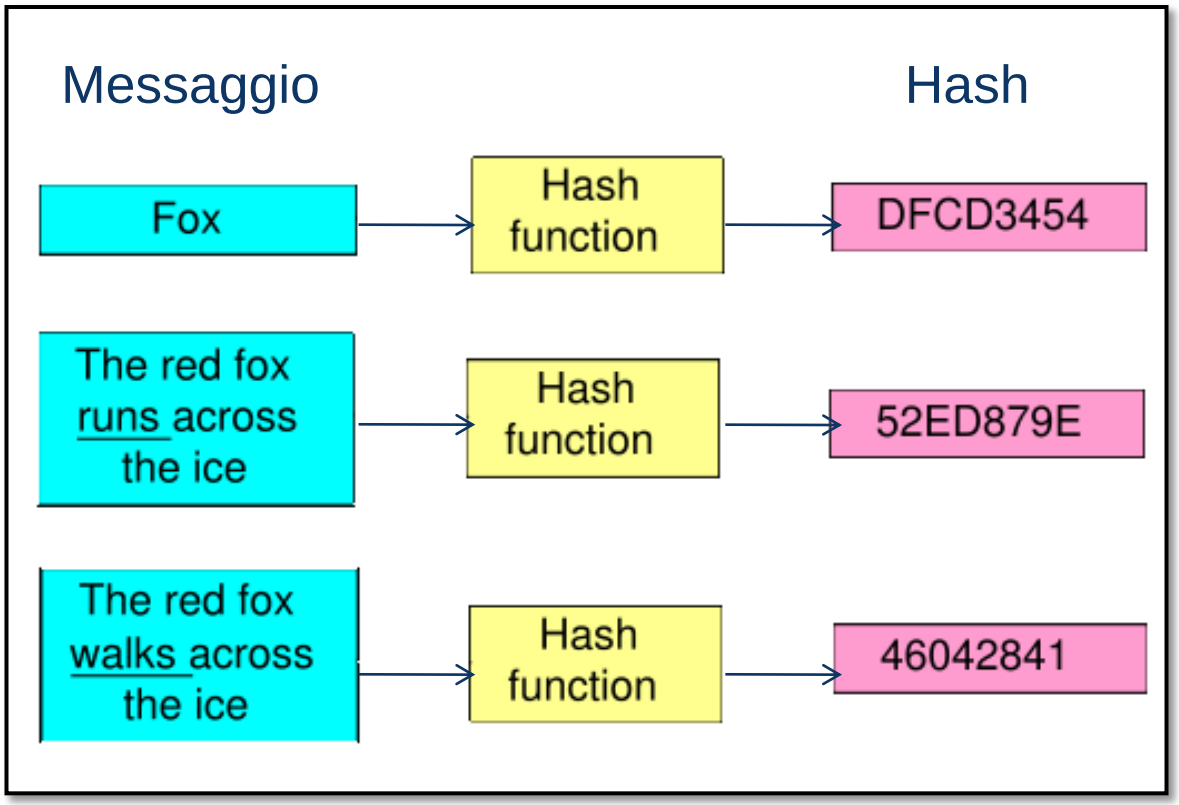
Preservazione

- *Dato che il numero dei valori in ingresso è **infinito** e il numero dei risultati in uscita è **finito**, risulta ovvia l'affermazione secondo cui "**esistono infiniti valori che danno lo stesso risultato di hash**".*

- Il problema delle collisioni.

Preservazione

- Soluzione: uso del doppio hash
  - Md5
  - Sha1



Agenda

- Acquisizione di un sistema Live
- Acquisizione di un sistema “dead”
- Network forensics: intercettazione dei dati in transito
- Preservazione delle evidenze: gli algoritmi di Hash
- **Catena di Custodia**

Catena di Custodia

- Il valore di hash non solo dimostra, al momento dell'acquisizione, che il dato è stato duplicato correttamente, ma garantisce l'integrità dello stesso nell'arco del tempo;

- A volte serve applicare una marca temporale per "congelare" l'evidenza ad una certa data;

Catena di Custodia

- La Catena di Custodia è il documento che garantisce l'integrità dell'evidenza, dall'acquisizione al termine dell'analisi;

- Tre elementi da prendere in considerazione:
  - Device originale;
  - Immagine forense dello stesso;
  - Valore di hash;

Catena di Custodia

- Nella fase di acquisizione la catena di custodia garantisce:
  - Sicurezza
  - Trasparenza

Catena di Custodia

SOURCE

01001010011
11010101010
11101010101
11110001001
1010111110

DESTINATION

01001010011
11010101010
11101010101
11110001001
1010111110

Catena di Custodia

SOURCE

01001010011
11010101010
11101010101
11110001001
1010111110

DESTINATION

01001010011
11010101010
11101010011
11110001001
1010111110

Catena di Custodia

SOURCE

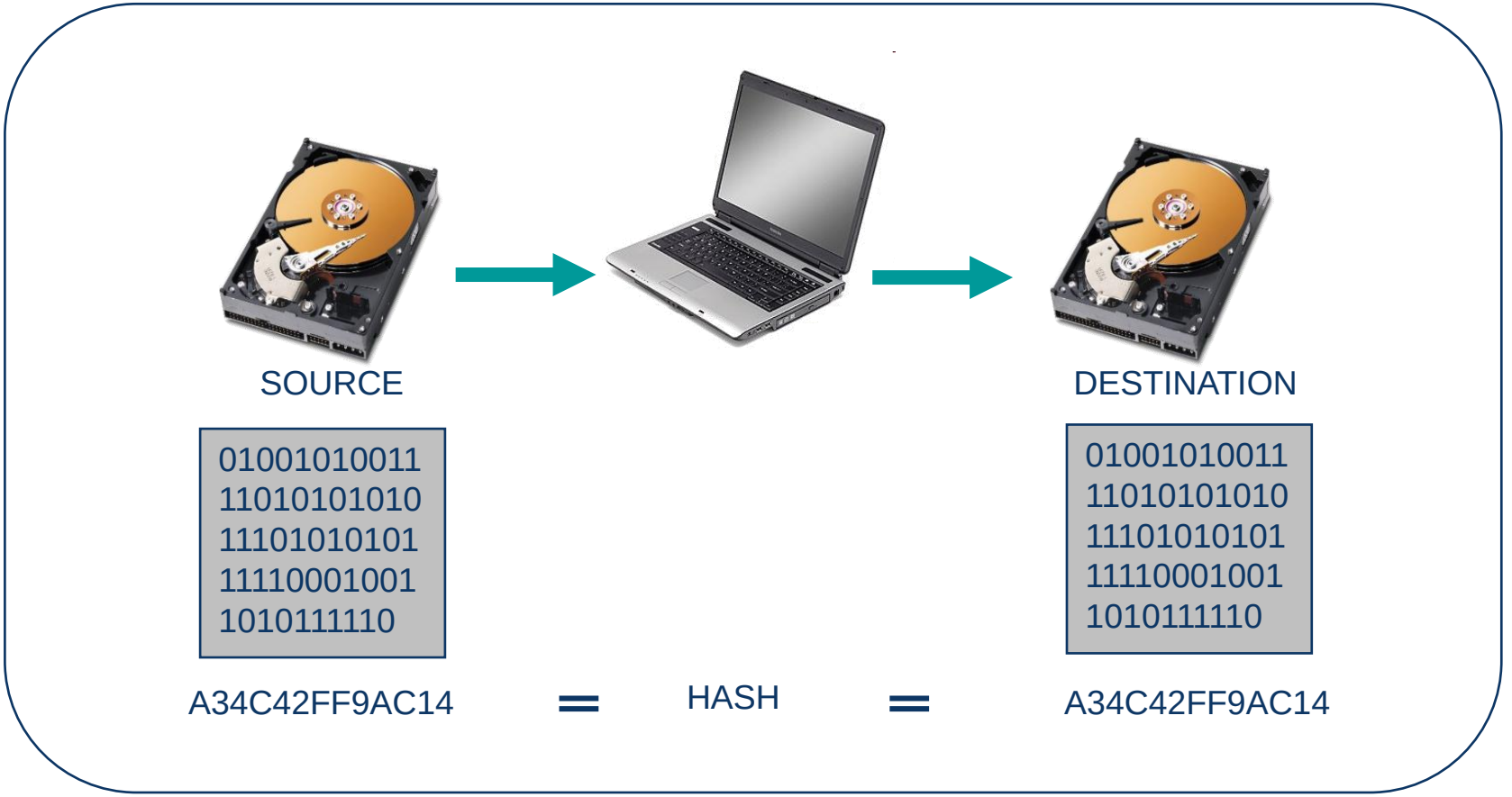
01001010011
11010101010
11101010101
11110001001
1010111110

DESTINATION

01001010011
11010101010
11101010011
11110001001
1010111110

A34C42FF9AC14 = HASH ≠ 492C97F78ADD7

Catena di Custodia



Catena di Custodia

- Nella fasi di estrazione ed analisi dei dati, può essere utilizzata per provare:
  - L'involontaria alterazione dei dati da parte dell'investigatore (errore umano);
  - Alterazione volontaria dei dati da parte di terzi;
  - Alterazione volontaria da parte dell'analista;

Catena di Custodia

PSS DIGITAL FORENSICS

Catena di Custodia

DETTAGLI DISPOSITIVO / COMPUTER		
Oggetto No: 1	Descrizione: Disco contenente copia del disco rigido del Sig. XYZ	
Manufacturer:	Modello: YYYYYYYYYY	Serial No: XXXXXXXXXXX
DETTAGLI IMMAGINE		
Data/Ora:	Creato da:	Metodo Usato: Slackware Linux
Descrizione: **Copia del disco rigido del computer XXX del Sig. XYZ**		
Nome File: **img_XYZ.img**		
Hash MD5: **abb809da2f70ea6692307d0f4325aa11**		
Hash SHA1: **b012a89da74fe2e1335cc7d854296218c3b67e75**		

Tracking No:	Data/ora	Da	A	Ragione
1	Data:	Nome/org: @PSS S.r.l. con Socio Unico	Nome/org:	
	Ora:	Firma:	Firma:	

Fase 3: Analisi

Analisi

- Come procedere correttamente?
  - Creazione Timeline;
  - Analisi dei file;
  - Ricerca Keyword;
  - Data Recovery;

- Questi quattro passaggi sono “ciclici”, continuo ad iterare man mano che trovo evidenze;

Analisi: Timeline

- Redirigere una timeline completa per poter effettuare un'analisi esaustiva:
  - MAC time di tutti i file "interessanti"
  - Quando è stato usato il sistema?

- Confrontare le eventuali timeline di dispositivi diversi.

Analisi dei File

- File ed informazioni utili:
  - Windows Registry
  - Log
  - File Metadata
  - Navigazione Internet
  - Email
  - IM e P2P
  - Recycle Bin

Analisi dei File: esempi

- Oltre ad aiutare l'investigatore durante l'analisi, la *feature* dei metadati di MS Word pone un serio problema riguardo al rischio di Information Disclosure;

- L'esempio più palese probabilmente è quello pubblicato nel Giugno 2003 da Richard M. Smith[1], riguardo ad un documento rilasciato dall'allora primo ministro inglese Tony Blair.

[1] *http://www.computerbytesman.com/privacy/blair.htm*

Analisi dei File: esempi

- Il governo britannico aveva pubblicato online, nel Febbraio 2003, un dossier su *“Iraq's security and intelligence organizations”.*

- Un docente dell’Università di Cambridge scoprì subito che gran parte del materiale presente in quel dossier era stato plagiato da degli scritti di un ricercatore americano in Iraq.

Analisi dei File: esempi

- Il clamore attorno alla vicenda ha suscitato l'interesse di diversi "curiosi" che hanno analizzato il file pubblicato;
- Tra i vari errori commessi in questa vicenda dal governo Blair, ce n'è uno che ci interessa particolarmente: il dossier pubblicato era un file MS Word;
- Vediamo quali metadati è stato possibile recuperare …

Analisi dei File: esempi

```
C:\Perl>wmd.pl g:\book2\ch5\blair.doc
--------------------
Statistics
--------------------
File = g:\book2\ch5\blair.doc
Size = 65024 bytes
Magic = 0xa5ec (Word 8.0)
Version = 193
LangID = English (US)
Document was created on Windows.
Magic Created : MS Word 97
Magic Revised : MS Word 97
--------------------
Last Author(s) Info
--------------------
1 : cic22 : C:\DOCUME~1\phamill\LOCALS~1\Temp\AutoRecovery save of Iraq -
   security.asd
2 : cic22 : C:\DOCUME~1\phamill\LOCALS~1\Temp\AutoRecovery save of Iraq -
   security.asd
3 : cic22 : C:\DOCUME~1\phamill\LOCALS~1\Temp\AutoRecovery save of Iraq -
security.asd
4 : JPratt : C:\TEMP\Iraq - security.doc
```

Analisi dei File: esempi

```
5 : JPratt : A:\Iraq - security.doc
6 : ablackshaw : C:\ABlackshaw\Iraq - security.doc
7 : ablackshaw : C:\ABlackshaw\A;Iraq - security.doc
8 : ablackshaw : A:\Iraq - security.doc
9 : MKhan : C:\TEMP\Iraq - security.doc
10 : MKhan : C:\WINNT\Profiles\mkhan\Desktop\Iraq.doc
--------------------
Summary Information
--------------------
Title : Iraq- ITS INFRASTRUCTURE OF CONCEALMENT, DECEPTION AND INTIMIDATION
Subject :
Authress : default
LastAuth : MKhan
RevNum : 4
AppName : Microsoft Word 8.0
Created : 03.02.2003, 09:31:00
Last Saved : 03.02.2003, 11:18:00
Last Printed : 30.01.2003, 21:33:00
----------------------------
Document Summary Information
----------------------------
Organization : default
```

Analisi dei File: esempi

- L'11 Gennaio 2010 viene pubblicato un post da fonte anonima, che dice di aver avuto una giornata pesante a lavoro e volerla concludere in maniera "particolare";
- Pubblica anche una foto, ma afferma che non rivelerà dove lavora…

Analisi dei File: esempi

- La foto pubblicata raffigura tre piste di (presumibilmente) cocaina;
- Qualche utente particolarmente curioso ha provato ad estrarne i metadati e…

Analisi dei File: esempi

```
Make                    : Apple
Camera Model Name       : iPhone 3GS
Orientation             : Horizontal (normal)
X Resolution            : 72
Y Resolution            : 72
…
Exif Version            : 0221
Date/Time Original      : 2010:01:11 18:05:47
Create Date             : 2010:01:11 18:05:47
…
GPS Altitude            : 49 m Above Sea Level
GPS Latitude            : 38 deg 53' 52.20" N
GPS Longitude           : 77 deg 2' 12.00" W
GPS Position            : 38 deg 53' 52.20" N, 77 deg 2' 12.00"
   W
Image Size              : 450x600
```

Analisi dei File: esempi

- Vera o falsa che sia la foto appena analizzata, dimostra comunque l'utilità e la "pericolosità" delle informazioni che possono essere contenute nei metadati…
- In ogni caso, alla corte di Obama questo lo avranno sicuramente capito ;-)

Analisi: Ricerca Keyword

- Scegliere attentamente le keyword da ricercare
  - Ridurre la possibilità di falsi positivi
- Liste infinite di keyword rallenteranno il lavoro in maniera consistente, oltre a risultare spesso ridondanti …

Analisi: Data Recovery

- Data Recovery: come e dove recuperare i dati;
  - Slack space;
  - Unpartitioned space;
  - Undeleting vs Carving;

Analisi: esempi

- “X dice al PM che nelle sei memory card NON ci sono fotografie...”
- Usando un software di “file carving” invece di un normale “undeleter” si ritrovano 453 immagini nelle sei card.
- **“Dottoressa... lei ha un problema”**

Fase 4: Reportizzazione

Reportizzazione

- Principi a cui si deve ispirare un'analisi corretta:
  - **Ripetibilità**
  - **Trasparenza**
  - **Semplicità**
  - **Disponibilità**

Reportizzazione

- **Ripetibilità**
  - Che sia corretto o no, la maggior parte degli incarichi vengono assegnati in base all'articolo 359 c.p.p.
  - E' quindi imperativo non modificare la prova durante le analisi.
  - Deve essere sempre possibile verificare che la prova non sia stata alterata.
  - Al contrario di altre discipline (balistica, genetica, biologia) è possibile duplicare il campione ad libitum (nella maggior parte dei casi). Un perito informatico che rovina una prova o è ignorante o in malafede. **Non vi sono scusanti**.

Reportizzazione

- **Trasparenza**
  - Una perizia perfetta deve permettere a chiunque conosca l'informatica di giungere allo stesso risultato;
  - Ogni passo deve essere documentato;
  - Ogni codice deve essere pubblicato in formato sorgente;
  - Si dovrebbero usare sempre formati aperti;
  - Se possibile, è sempre meglio utilizzare programmi open source.

Reportizzazione

- **Semplicità**
  - I ragionamenti devono essere semplici e lineari. Già la materia è oscura di per sè, è quindi inutile complicare inutilmente le cose.
  - Chi legge la perizia deve essere in grado di capire non solo la logica di fondo, ma soprattuto le misure tecniche adottate per arrivare ad un determinato risultato;

Reportizzazione

- **Conclusioni**
  - Non vanno sottovalutate;
  - La maggior parte di coloro che leggeranno la perizia leggeranno solo l'incipit e le conclusioni.
  - Tanto più questo vale per le figure di impronta e/o derivazione legale;
  - ***Devono essere oggettive***

Conclusioni

- Le evidenze digitali sono facilmente alterabili;
- Il mancato utilizzo di una metodologia corretta può portare alla compromissione di tali evidenze e conseguentemente alla loro inutilizzabilità;
- Non sempre è possibile effettuare una analisi non invasiva: richiedere il 360 c.p.p.

Conclusioni

- Le evidenze che cerchiamo possono essere in tantissimi posti diversi: è importante sapere dove cercare;

- Il processo di investigazione forense è composto da diverse fasi, non è solo l'analisi. Tutte le fasi sono ugualmente importanti: un errore in una qualunque di queste, potrebbe inficiare tutto il lavoro svolto.

Grazie per l'attenzione.
Domande?