Evidenza digitale ed Informatica Forense

Le tracce informatiche

- Quando si usa un dispositivo elettronico si lasciano sui dispositivi di memorizzazione ad esso collegati delle *tracce*, vale a dire artefatti dovuti all'interazione di un utente con il computer
- Queste tracce sono in genere dette *tracce informatiche* o, anche, *tracce digitali*

Le tracce informatiche

- Alcuni esempi:
  - File prodotti da applicazioni di varia natura
  - File di sistema (es. file di log)
  - Informazioni relative ai file gestite direttamente dal sistema operativo (es. data ed ora di creazione ed ultima modifica del contenuto di un file)
  - Dati trasmessi tra due o piu' computer collegati ad Internet

Immaterialita’ delle tracce digitali

- Le tracce digitali sono *immateriali*
  - non esistono come oggetto fisico, ma consistono in *sequenze di bit* memorizzate su dei *dispositivi di archiviazione dati*
- Per accedere ad una traccia digitale, occorre quindi accedere al dispositivo su cui essa e’ memorizzata

I dispositivi di memorizzazione

- *Dispositivi di memorizzazione persistenti*: non necessitano di alimentazione per mantenere i dati memorizzati
  - hard disk, penne USB, CD/DVD-ROM, nastri, schede di memoria di vario tipo (Compact Flash, Secure Digital, MMC, …)
- *Dispositivi di memorizzazione volatili*: i dati memorizzati sono persi nel momento in cui si interrompe l'alimentazione
  - Memoria RAM del computer, telefonino, palmare, ecc.

Evidenza digitale: una definizione

- “Qualsiasi informazione, con valore probatorio, che sia memorizzata o trasmessa in formato digitale “ [*Scientific Working Group on Digital Evidence, 1998*]

Da traccia ad evidenza digitale

- Affinche' una traccia digitale assuma valore probatorio , e' necessario che essa sia:
  - *Autentica*: vi e' certezza della sua provenienza
  - *Integra*: priva di alterazioni
  - *Veritiera*: ottenuta mediante una corretta interpretazione dei dati
  - *Completa*: sono stati raccolti ed interpretati tutti i dati ad essa relativi
  - *Legale*: e' stata raccolta nel rispetto delle leggi vigenti

Autenticita’ di una traccia digitale

- Determinazione certa della sua provenienza
  - Individuazione della catena causale che ha portato alla sua comparsa nel dispositivo in cui essa e’ memorizzata
- Esempio: la presenza di un determinato file in un disco puo’ essere dovuta alle operazioni compiute da un virus , o da un terzo, piuttosto che da un’azione volontaria di dato utente

Integrita’ di una traccia digitale

- Le tracce digitali sono *fragili*, cioe’ facilmente modificabili nel caso in cui i dispositivi che le contengono siano maneggiati in modo inappropriato
  - L’accensione di un computer spento comporta la scrittura e/o modifica di numerosi file sul suo disco di sistema
  - L’esplorazione del contenuto di un hard disk comporta la modifica di varie proprieta’ importanti dei file, come ad esempio l’ora e la data dell’ultimo accesso
  - Lo spegnimento di un computer determina la perdita delle evidenze contenute nella sua memoria volatile

Integrita’ di una traccia digitale

- La fragilita’ delle tracce digitali impone l’utilizzo di metodologie e strumenti in grado di garantire *in modo dimostrabile* che l’evidenza non e’ stata modificata accidentalmente o deliberatamente durante la sua conservazione ed analisi

Veridicita’ di una traccia digitale

- Una traccia digitale consiste in una o piu’ *informazioni* reperite in un dispositivo di memorizzazione
- Un computer si limita a memorizzare *dati*
  - sequenze arbitrarie di *bit* (unita’ elementare di informazione che puo’ assumere unicamente i due valori ‘0’ ed ‘1’)
- *Informazione = dato+interpretazione*
  - per essere trasformate in informazioni, e quindi tracce digitali, i dati grezzi devono essere interpretati

Veridicita’ di una traccia digitale

- Problema: data una sequenza di bit, la sua interpretazione non e’ univoca
- Ad esempio, la sequenza di bit 1111101 puo’ essere interpretata come:
  - il numero intero positivo 125
  - Il numero intero negativo -3
  - il carattere ‘}’
  - sono possibili molte altre interpretazioni

Veridicita’ di una traccia digitale

- Una corretta interpretazione richiede la conoscenza *certa* del significato del dato in questione
- che a sua volta richiede la conoscenza profonda del funzionamento del sistema informatico e delle applicazioni che lo hanno prodotto
  - notevole difficolta’ dovuta alla complessita’ dei sistemi informatici moderni

Completezza di una traccia digitale

- La corretta interpretazione di una traccia digitale puo' richiedere l'analisi di piu' informazioni ad essa relative
- Esempio: per accertare l'intenzionalita' di un utente nel detenere materiale illecito, bisogna non solo reperire il materiale, ma anche accertare che
  - Lo stesso sia presente in cartelle non di sistema (per esempio i cosiddetti "file temporanei di Internet"), e che magari sia organizzato in cartelle o sottocartelle
  - Lo stesso non sia frutto di visualizzazione di finestre di "pop-up" aperte automaticamente da un sito web durante la sua visita
  - ...

Completezza di una traccia digitale

- Anche in questo caso e’ richiesta una conoscenza profonda delle dinamiche delle varie applicazioni e delle componenti del sistema operativo, nonche’ delle loro interazioni

Informatica Forense

- L’Informatica forense studia metodologie e strumenti per la raccolta, l’interpretazione, l’analisi, e la conservazione di evidenze digitali in modo da garantirne autenticita’, integrita’, veridicita’, completezza
- La simultanea presenza di queste proprieta’ rendono l’evidenza digitale *non ripudiabile*, e quindi efficacemente utilizzabile in giudizio

Legislazione rilevante

- Sino a pochissimo tempo fa, le procedure di gestione dell'evidenza digitale non erano regolamentate in alcun modo
- La recentissima ratifica della Convenzione del Consiglio d'Europa sulla criminalita' informatica prevede per la prima volta la necessita' di adottare specifiche cautele nella gestione dell'evidenza digitale

Legislazione rilevante

- In particolare, e' previsto che nelle operazioni di perquisizione e sequestro di dati digitali ad opera della P.G.
  - siano adottate "*misure tecniche dirette ad assicurare la conservazione dei dati originali e ad impedirne l'alterazione*" (artt. 8 commi 1 e 2, 9 commi 1 e 3)
  - "*la loro acquisizione avvenga mediante copia di essi su adeguato supporto con una procedura che assicuri la conformita' dei dati acquisiti a quelli originali e la loro immodificabilita*'" (artt. 8 commi 5 ed 8, 9 comma 3)

Legislazione rilevante

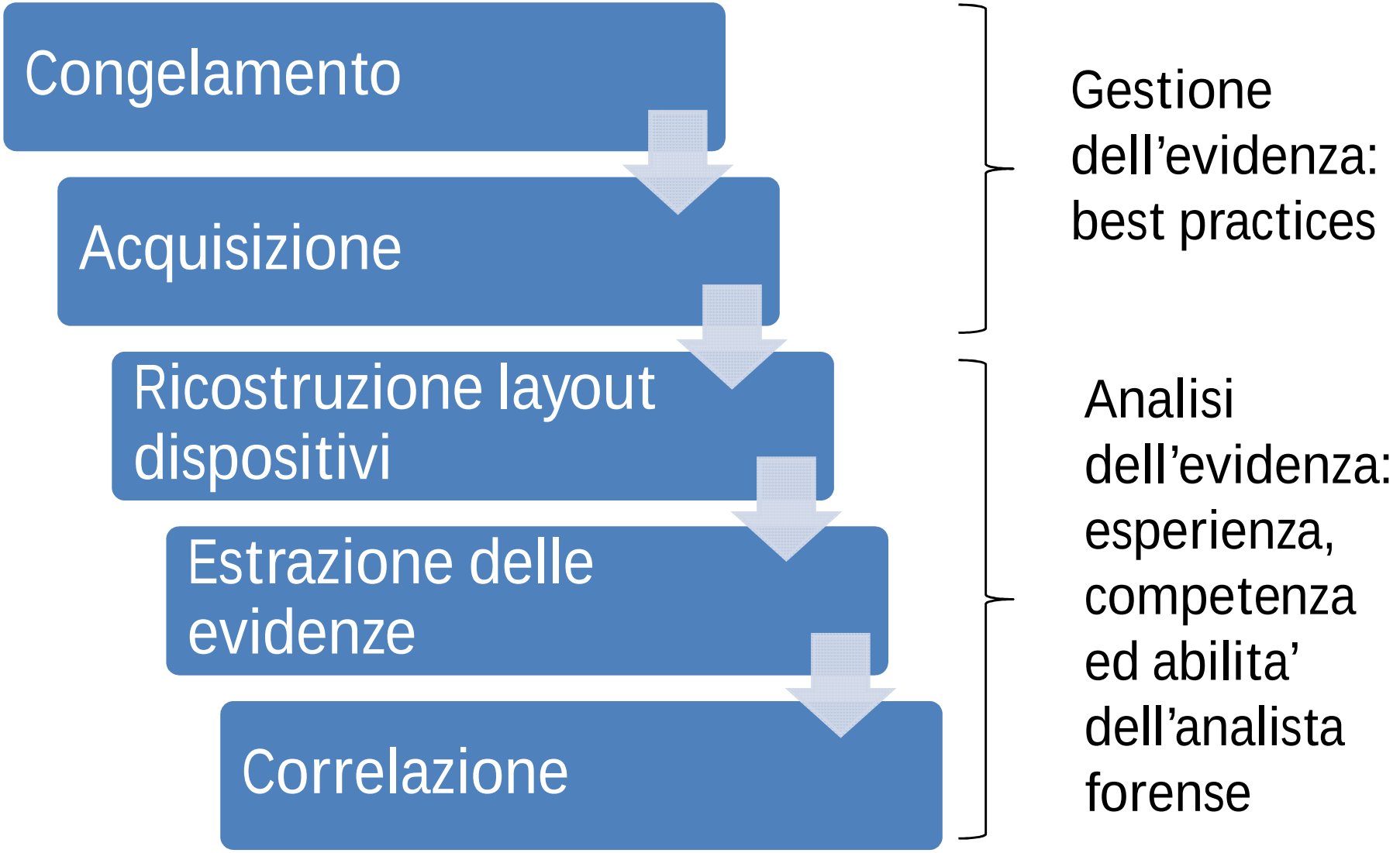
- Il legislatore non specifica pero' quali debbanno essere le procedure che permettono di rispettare le suddette disposizioni
- Una risposta in tal senso e' fornita dalle cosiddette "*Best practices*"
  - linee guida che specificano le migliori procedure per la gestione dell'evidenza digitale

Best practices

- In alcuni paesi anglosassoni (USA e UK) formalizzate da documenti emanati da organismi istituzionali
  - Association of Chief Police Officers (ACPO) in Gran Bretagna [ACPO 2007]
  - US Department of Justice – National Institute of Justice – USA [NIJ 2004] e [NIJ 2007]

Metodologie di investigazione informatica

Metodologia di indagine



Criticita’ tecnico – procedurali ed errori conseguenti

Possibilita’ di errore

- Le metodologie e gli strumenti utilizzati nell’Informatica Forense non sono perfetti, e possono dare luogo ad errori di varia natura, che possono inficiare in toto o in parte la valenza probatoria delle tracce informatiche riscontrate

Contaminazione dell'evidenza digitale

- Gli errori piu' frequentemente riscontrati sono:
  - collegamento di un disco ad un PC senza write blocker con conseguente modifica di dati e/o metadati e discrepanza tra i codici di hash
  - accensione di un PC congelato e suo collegamento ad Internet (con proseguimento di attivita' di download / file sharing)
  - ispezione di un PC acceso

Identita’ virtuale ed identita’ reale

- Spesso si tende a confondere l’identita’ virtuale di un utente con quella reale di una persona
  - le evidenze digitali permettono di individuare l’identificatore dell’utente (login name) che ha compiuto determinate azioni con un certo computer
  - questo non e’ pero’ di per se sufficiente per stabilire con certezza l’identita’ reale della persona che ha commesso un dato fatto
  - l’analista dovrebbe (anche nella sua relazione finale) non attribuire mai le attivita’ individuate sul computer ad una persona, quanto piuttosto ad un particolare utente definito nella configurazione del sistema operativo del computer

Gestione dell’informazione temporale

- Le evidenze che riguardano informazioni temporali sono quelle piu’ delicate da gestire
- Eventuali errori nella loro gestione possono inficiare ogni ricostruzione temporale delle attivita’ reperite su un computer

Gestione dell’informazione temporale

- Errore 1: non rilevare e documentare le impostazioni di data ed ora del BIOS del computer all’atto del congelamento
  - le varie informazioni temporali memorizzate dal sistema operativo nei tempi MACE o nei file di log, e dalle applicazioni nei metadati che esse producono, sono lette dalle impostazioni dell’orologio del computer
  - se l’orologio e’ impostato ad un valore diverso da quello corretto, i riferimenti temporali possono essere errati

Gestione dell’informazione temporale

- Errore 2: non controllare (ed eventualmente confermare o escludere) l’eventuale presenza di modifiche all’orologio del sistema
  - se anche le impostazioni del BIOS dovessero risultare corrette all’atto del congelamento, le stesse potrebbero essere state modificate in momenti precedenti
  - in tal caso, alcune informazioni temporali sarebbero errate, mentre altre sarebbero corrette

Gestione dell’informazione temporale

- Errore 3: errori di interpretazione e conversione dell’informazione temporale
  - Le informazioni temporali possono essere memorizzate in molti modi diversi
  - In generale, le informazioni temporali sono memorizzate come numero di unita’ temporali (secondi, minuti, millisecondi, ecc.) trascorse da un certo istante (questa quantita’ e’ detta *offset*)

Gestione dell’informazione temporale

- Differenze di unita’ di misura
  - Componenti del sistema operativo diverse, o applicazioni diverse, possono usare unita’ diverse
- Differenze di momento di riferimento
  - Il momento a partire dal quale si misura l’offset puo’ essere diverso (ad esempio, corrisponde alle 00:00 del 1/1/1970 nei sistemi Unix, ed alle 00:00 del 1/1/1601 nei sistemi Windows)
- Differenze di zona oraria
  - Applicazioni diverse possono riferire le informazioni temporali da esse memorizzate a zone orarie diverse (ad esempio, UTC o zona locale, oppure ora solare o ora legale)

Gestione dell’informazione temporale

- In sintesi, la gestione dell’informazione temporale nel corso di una investigazione digitale deve essere effettuata con la massima cura
  - e’ una grave negligenza non documentare nella relazione finale le procedure utilizzate per trattarla, che lascia la porta aperta a contestazioni successive

Gestione dell’informazione temporale

- La presenza di eventuali alterazioni puo’ essere accertata in diversi modi
- Il mancato reperimento di modifiche, invece, non permette di escludere con certezza che le stesse non si siano verificate e che le tracce siano state successivamente cancellate
  - occorre fornire il maggior numero di elementi che supportino l’ipotesi di assenza di modifiche

Non considerare il quadro d’insieme

- I computer sono sistemi complicati in cui una data traccia digitale puo’ essere dovuta a diverse cause: occorre escludere tutte le ipotesi che non spiegano l’evidenza
- Ad esempio, il ritrovamento di file illeciti su un PC puo’ non essere sufficiente se non si e’ in grado di dimomstrare che l’utente li ha scaricati intenzionalmente:
  - sono nei file temporanei di Internet o sono organizzati in cartelle/sottocartelle?
  - Sono presenti virus/trojan che possono aver trasferito il materiale?
  - Sono stati scaricati in conseguenza a visite involontare a siti web, dovute a redirezioni automatiche oppure alla comparsa di finestre di “pop up”?

Errori nell’interpretazione degli artefatti

- Una conoscenza incompleta o imprecisa del funzionamento del sistema operativo e delle applicazioni puo’ portare ad interpretare erroneamente il significato di una traccia digitale
  - decodifica dei record di un file di log
  - interpretazione delle date MACE

Errori nell’interpretazione dei dati

- Le tracce digitali sono il risultato dell’interpretazione di dati memorizzati fisicamente sui dispositivi
- Se l’interpretazione e’ errata, sono errate le conclusioni cui si giunge

Errori del software di analisi

- Un programma software contiene sempre dei bug dovuti ad errori di programmazione o ingnoranza del programmatore
  - La maggioranza dei software di analisi forense sono di tipo commerciale e closed-source, e non permettono l'ispezione del codice per accertare eventuali errori
  - L'analista dovrebbe sempre validare i propri risultati, al fine da escludere al presenza di errori

Ammissibilita’ e test scientifici

- Idealmente, l’ammissibilita’ di un dato strumento o procedura dovrebbe essere subordinata al superamento di test specifici (per es. il Test di Daubert):
  - *testing*: verifica sperimentale della procedura
  - *error rate*: la percentuale di errore deve essere nota
  - *publication*: pubblicazione della procedura su riviste/congressi peer-reviewed
  - *acceptance*: la procedura e’ generalmente accettata dalla comunita’ scientifica di riferimento

Il buon analista [Monga 2006]

- *Formazione ed aggiornamento*
  - so quel che faccio
- *Verifiche incrociate ed indipendenti*
  - lo so fare in molti modi
- *Adesione a standard d'azione internazionali*
  - molti altri fanno come me
- *Scrupolosa reportistica*
  - tutto cio' che faccio lo documento in modo che possa essere esaminato in contraddittorio

Conclusioni

- Le tecniche di Informatica Forense assumono un ruolo sempre piu' importante sia in ambito civile che penale
- Necessaria una maggior comprensione delle sue potenzialita' e problematiche da parte degli "operatori del diritto" (magistratura, avvocati, p.g.)
- Ancora troppa improvvisazione da parte di consulenti e periti che spesso mancano di una formazione specifica
- Poco applicate le regole previste dalle best practices e dalle discipline forensi classiche