# Informatica Forense ed Investigazioni informatiche

Prof. Cosimo Anglano
Dipartimento di Informatica
Laboratorio di Ricerca sull'Informatica Forense
Universita' del Piemonte Orientale, Alessandria
email: cosimo.anglano@unipmn.it
http://digitalforensics.di.unipmn.it

# Disponibilita' slides

- Queste slides sono rilasciate con licenza Creative Commons "Attribuzione-Non commerciale-Condividi allo stesso modo 2.5", il cui testo e' disponibile sul sito http://creativecommons.org/licenses/by-nc-sa/2.5/it/legalcode
- Sono disponibili sul sito http://digitalforensics.di.unipmn.it nella sezione "Pubblicazioni" del sito

# Introduzione

- Il numero di dispositivi elettronici che fanno parte della nostra vita quotidiana e' in crescita costante
  - Computer, telefonini, riproduttori MP3, palmari, navigatori GPS, …
- Conseguentemente, e' in crescita il numero di situazioni in cui le prove di determinati reati o comportamenti illeggittimi sono memorizzate in questi dispositivi

# Introduzione

- La cronaca giudiziaria piu' recente riferisce di almeno due casi in cui i dati memorizzati su un computer hanno assunto un ruolo probatorio notevole
  - Omicidio di Chiara Poggi a Garlasco
  - Omicidio di Meredith Kercher a Perugia
- In generale, si assiste ad un incremento dei reati di tipo informatico, o di tipo tradizionale commesso mediante uno strumento informatico [EUROPOL 2007]
- Sta inoltre aumentando il numero di casi in cui evidenze digitali sono usati a supporto o per confutare tesi accusatorie

# Introduzione

- L'impiego di prove digitali non e' limitato al solo ambito penale
- Sempre piu' spesso le prove documentali utilizzate in giudizi civili o giuslavoristici sono disponibili in formato digitale e devono essere estratte con le opportune metodiche dai computer in cui sono memorizzate
  - messaggi di email contenenti prove documentali di accordi tra aziende e/o privati
  - documenti attestanti  possesso illeggittimo di informazioni coperte da diritti di proprieta' intellettuale
  - tracce confermanti l'effettuazione di attivita' non permesse durante l'orario di lavoro con strumenti informatici aziendali

# Introduzione

- Domande quali:
  - Il computer e' stato utilizzato per accedere abusivamente ed inviare a terzi a documenti e/o dati riservati? Se si, a chi li ha inviati?
  - Il computer e' stato utilizzato per accedere, memorizzare, scambiare materiale illecito (es. pedopornografico)?
  - Il computer e' stato utilizzato per lanciare/coordinare un attacco informatico?
  - L'attivita registrata su un computer conferma o smentisce l'alibi di un soggetto incriminato di un dato reato non informatico?

# Introduzione

- … trovano risposta nelle attivita' di investigazioni basate sull'utilizzo di tracce informatiche
- Il termine Informatica Forense (Computer Forensics in Inglese) e' diventato di forte attualita' in conseguenza della recente cronaca giudiziaria

# Sommario

- Parte 1: evidenza digitale ed informatica forense
- Parte 2: metodologie di investigazione informatica
- Parte 3: criticita' tecnico-procedurali ed errori conseguenti

# Parte 1
# 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]

# Parte 2
# Metodologie di investigazione informatica

# Metodologia di indagine


**Gestione dell'evidenza**: best practices

**Analisi dell'evidenza**: esperienza, competenza ed abilita' dell'analista forense

# Gestione dell'evidenza digitale

# Congelamento

- I supporti originali vanno *congelati*, vale a dire non devono essere piu' collegati ad un computer senza che sia utilizzato un dispositivo che garantisca il blocco delle operazioni di scrittura
  - se i supporti non sono stati rimossi, non bisogna mai piu' avviare il computer direttamente
    - Possibile comunque usare un cosiddetto *CD avviabile* di tipo forense

# Congelamento

- I supporti originali vanno poi sigillati in modo opportuno e va gestita (documentata) la catena di custodia

# Congelamento

- All'atto del congelamento, e' di fondamentale importanza esaminare ed annotare l'ora e la data impostata nell'orologio di sistema
  - Il sistema operativo assegna una etichetta temporale ai vari file e/o eventi leggendo i valori impostati nell'orologio di sistema
  - Se l'orologio di sistema riporta valori diversi da quelli corretti, le informazioni temporali vanno opportunamente corrette a loro volta
  - Se in fase di analisi non si ha a disposizione questa informazione, non e' possibile stabilire con certezza una linea temporale degli eventi

# Congelamento

- Il congelamento dovrebbe avvenire in un momento il piu' vicino possibile alla data "dei fatti"
  - se si tratta di un hard disk collegato ad un PC che viene usato, i suoi dati possono (ed in genere sono) alterati, e non si trovano piu' nello stato in cui erano al momento del fatto
  - discorso che non sempre si applica a supporti non in linea (tipo penne USB, schede di memoria), e che non vale per i supporti non riscrivibili (CD/DVD write once)

# Congelamento

- Come si puo' "fotografare" il contenuto di un dispositivo in modo che si possa determinare se sono state introdotte delle modifiche di qualunque tipo in un momento successivo al suo congelamento?
  - sicuramente non possiamo stamparne il contenuto
- Le best practices raccomandano l'impiego di *algoritmi di hash crittografici* per calcolare un *codice hash* del dispositivo congelato

# Congelamento

- Un algoritmo di hash crittografico produce, a partire da una sequenza di bit di lunghezza e contenuto arbitrari, una codice (sequenza di caratteri) avente lunghezza prefissata che gode di alcune importanti proprieta'

# Congelamento

- *Proprieta' 1: due sequenze di input identiche danno luogo allo stesso codice hash*
  - All'atto del congelamento del dispositivo, si calcola il relativo codice hash
  - Per verificare l'assenza di alterazioni dell'originale, e' sufficiente ricalcolare il codice hash e verificare che lo stesso sia uguale a quello calcolato all'atto del congelamento

# Congelamento

- *Proprieta' 2: la probabilita' che sequenze diverse diano luogo allo stesso codice e' praticamente nulla*
  - se i valori dei codici hash calcolati all'atto del congelamento ed in un momento successivo differiscono, con altissima probabilita' si sono verificate alterazioni

# Congelamento

- La probabilita' di generare lo stesso codice hash partendo da sequenze diverse (*probabilita' di collisione*) dipende dall'algoritmo di hashing impiegato
  - MD5: produce codici lunghi 32 caratteri ed ha una probabilita' di collisione pari a $1/2^{64}$, cioe' 1 su 18.446.744.073.709.551.616 (piu' di 18 miliardi di miliardi)
  - SHA1: produce codici lunghi 40 caratteri ed ha una probabilita' di collisione pari a a $1/2^{80}$, cioe' 1 su 1.208.925.819.614.629.174.706.176 (oltre 1200 miliardi di miliardi)

# Congelamento

- Per calcolare il codice hash di un dispositivo, occorre leggerne i dati
- Per escludere in maniera assoluta la possibilita' di modificare accidentalmente I dati contenuti nel dispositivo, e' necessario utilizzare dispositivi o tecniche che bloccano le operazioni di scrittura
  - dispositivi hardware (*write blockers*)
  - possibili anche soluzioni software, da usare con grande cautela

# Acquisizione

- Per preservare l'integrita' delle evidenze digitali, le best practices raccomandano di effettuare tutte le operazioni di analisi su copie identiche dei dispositivi originali

# Acquisizione

- Al fine di garantire la completezza delle evidenze digitali, e' necessario acquisire *tutte le parti del dispositivo*, e non solo quelle utilizzate per memorizzare i file
  - Aree contenenti file cancellati e/o loro frammenti
  - Aree contenenti informazioni gestite dal sistema operativo e non visibili all'utente
- Tali copie sono dette *copie forensi* o anche *copie bit-a-bit*

# Acquisizione

- Prima di effettuare la copia, si ricalcola il codice hash dell'originale e lo si confronta con quello ottenuto all'atto del congelamento
  - permette di escludere modifiche sull'originale
- Al termine dell'acquisizione si calcola il codice hash della copia appena effettuata
  - deve essere identico a quello dell'originale in modo da attestare l'assoluta identicita' tra originale e copia

# Acquisizione

- L'operazione di acquisizione produce un *file di immagine* contente una copia di tutti i bit memorizzati nel dispositivo
  - in passato si era soliti riversare la copia direttamente su un hard disk, ottenendo quindi un *clone*, piuttosto che in un file
  - tecnica attualmente in disuso in quanto piu' scomoda da utilizzare
- Tale file viene poi interpretato ed analizzato mediante opportuni software di analisi forense

# Analisi dell'evidenza digitale

# Analisi delle evidenze digitali

- Scopo dell'analisi e' individuare le eventuali tracce digitali che permettono
  - di ricostruire le attivita' compiute mediante il computer cui il dispositivo era collegato
  - di individuare elementi probatori a favore o sfavore di una tesi accusatoria o difensiva
- Data la complessita' delle operazioni da effettuare, ed alla notevole quantita' di dati da analizzare, l'analisi viene effettuata con l'ausilio di appositi software di analisi forense

# Analisi delle evidenze digitali

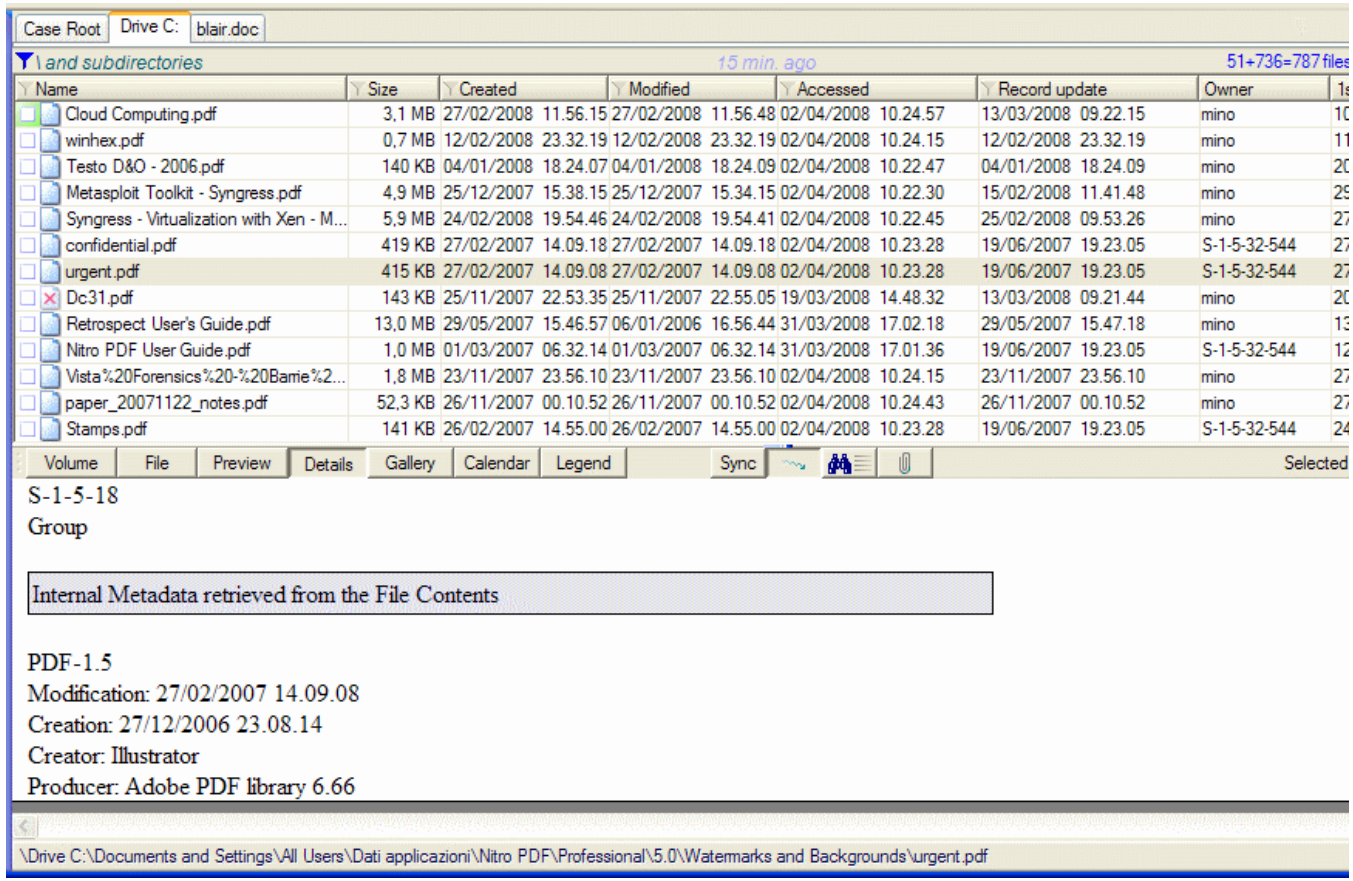
- Si interpreta il file di immagine come se fosse un dispositivo reale e si effettuano diversi tipi di analisi
- Analisi a livello del file system:
  - Estrazione di tracce informatiche prodotte dal sistema di gestione dei file (file system) o contenute nei file
- Analisi a livello del sistema operativo:
  - Estrazione degli artefatti prodotti dal sistema operativo durante il suo funzionamento (es. file di log)
- Analisi a livello delle applicazioni:
  - Estrazione di artefatti prodotti da programmi applicativi utilizzati dagli utenti

# Analisi a livello di file system: recupero di file cancellati

- L'operazione di cancellazione di un file e' solo logica
  - sono cancellati unicamente i riferimenti al file, e le aree del supporto ad esso assegnate sono rimesse a disposizione del sistema operativo
  - idati veri e propri pero' non sono cancellati a meno che non siano sovrascritti da altri file
- La persistenza dei dati sul dispositivo permette spesso di recuperare (in toto o in parte) il contenuto di file cancellati

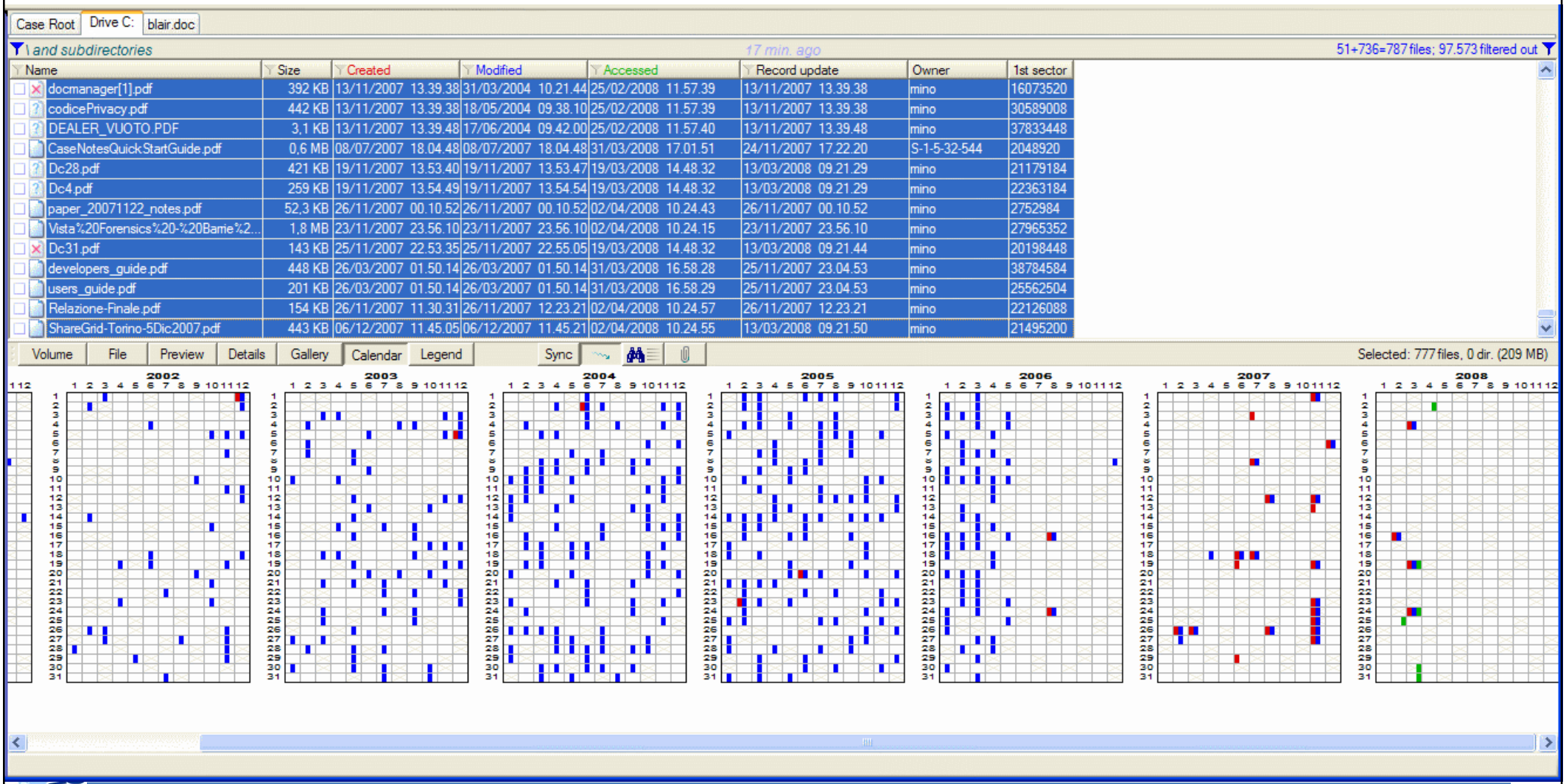
# Analisi a livello di file system: analisi delle proprieta' dei file

- Varie informazioni rilevanti ai fini probatori
  - Tempi MACE (Modified/ Accessed/Created/Entry Modified)
  - Utente proprietario di un dato file



# Analisi a livello di file system: analisi delle proprieta' dei file

- I tempi MACE permettono di ottenere una timeline delle attivita' effettuate sui file



# Analisi a livello di file system: identificazione del tipo corretto di un file

- Un espediente (ingenuo) per nascondere un file di un dato tipo consiste nel ridenominarlo modificandone l'estensione
  - In realta', ciascun tipo di file (JPEG, GIF, eseguibile, prodotto da Office, ecc.) e' caratterizzato dalla presenza – in punti ben precisi – di sequenze di valori costanti (detti *firma del file* o *file signature* in Inglese)
  - Analizzano le firme dei file, e' possibile individuarne il tipo effettivo

# Analisi a livello di sistema operativo

- Ogni sistema operativo, nel corso del suo funzionamento, registra varie informazioni in appositi file, memorizzati in posizioni note nel disco di sistema
  - file di log di varia natura
  - file di configurazione in cui sono memorizzate informazioni relative alla configurazione del sistema, agli utenti , alle applicazioni installate ed all'utilizzo delle stesse
  - "snapshot" della configurazione del sistema utilizzate per eventuali azioni di ripristino
  - In generale, sono presenti molti altre informazioni di questo tipo
- Differenze anche forti nel tipo e nella quantita' di informazioni registrate dai diversi sistemi operativi

# Analisi a livello di sistema operativo

- L'analisi a livello del sistema operativo puo', ad esempio, consentire di
  - Tracciare l'uso del computer da parte dei diversi utenti ivi definiti
  - Individuare se e quando determinati file sono stati aperti da un certo utente
  - Le periferiche che sono state collegate (ad esempio, penne USB o dischi esterni)
  - Individuare l'elenco (e spesso anche il contenuto) dei file stampati e su quale stampante
  - Identificare le reti (tradizionali o WiFi) cui il computer e' stato collegato
  - E molte altre informazioni rilevanti ai fini probatori

# Analisi a livello delle applicazioni

- In generale, un utente interagisce con un computer mediante applicativi software
- Ciascun applicativo software presenta delle caratteristiche specifiche che possono permettere di ottenere diverse informazioni utili ai fini probatori

# Analisi a livello delle applicazioni

- L'analista deve conoscere approfonditamente il funzionamento delle applicazioni da analizzare per poter individuare ed estrarre tali informazioni
- Analisi complessa a causa di
  - grandissimo numero di applicazioni esistenti, ciascuna delle quali ha delle caratteristiche assolutamente diverse dalle altre
  - mancanza di documentazione sul formato interno dei file da esse prodotti (necessario *reverse engineering*)

# Analisi a livello delle applicazioni: estrazione di contenuti “embedded”

- Estrazione di oggetti digitali contenuti nei file
  - Per nascondere contenuti illeciti, spesso li si inserisce in un file (ad esempio, foto o filmati inseriti in file Word o PDF)
  - Dall’analisi della struttura dei file “contenitori”, che dipende dal programma applicativo che li ha generati, e’ possibile individuare ed estrarre tali oggetti

# Analisi a livello delle applicazioni: *metadati* applicativi

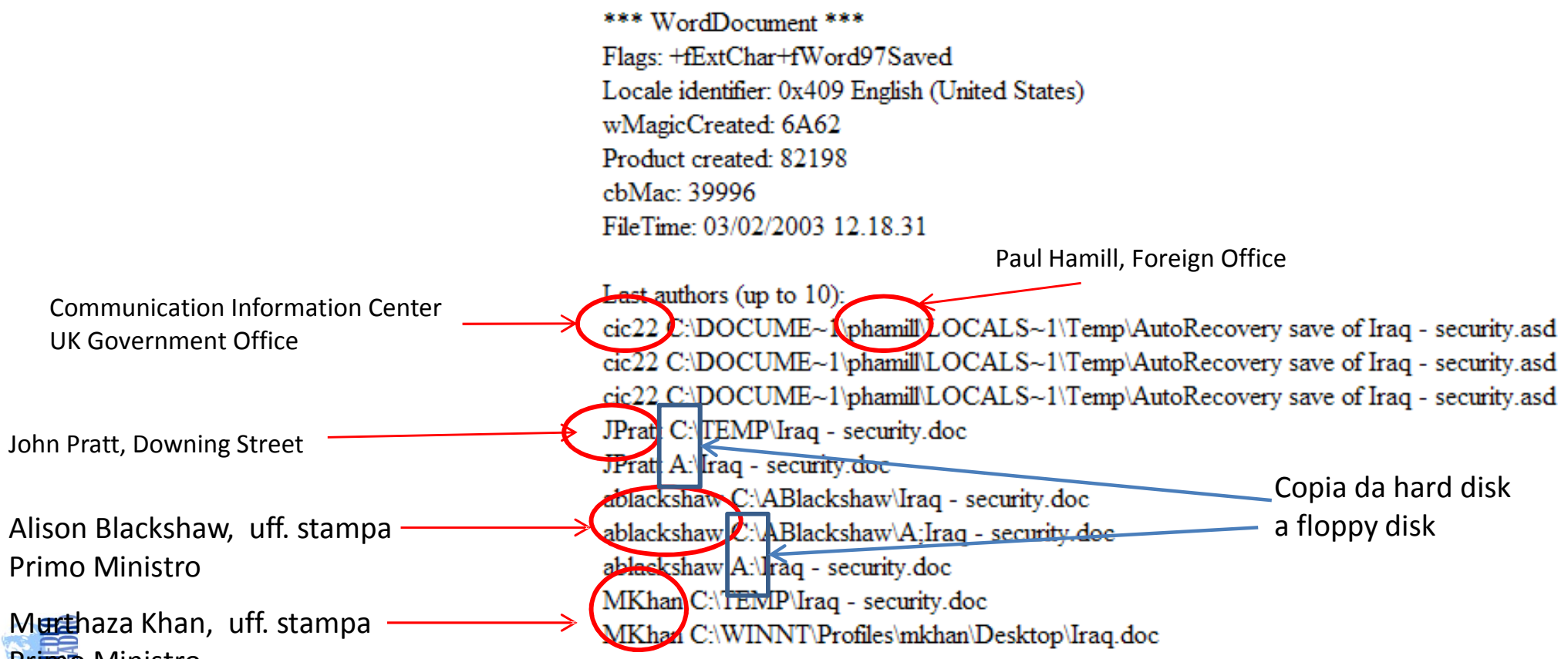
- Molte applicazioni memorizzano, spesso all'insaputa dell'utente, informazioni di varia natura (*metadati*) nei file prodotti mediante di esse
- Questi medatati non sono normalmente visibili all'utente, ma possono essere estratti dai programmi di analisi forense e possono fornire molte informazioni utili ai fini probatori

# Analisi a livello delle applicazioni: metadati applicativi – file MS Office

- Il 30/1/2003, il governo inglese pubblico' su un proprio sito web un dossier sulla struttura delle organizzazioni di intelligence e sicurezza irachene, che fu citato da Colin Powell nella sua relazione all'Assemblea delle Nazioni Unite il 5/2/2003
- Il Dr. G. Rangwala (U. Cambridge) si accorse che il dossier era stato in larghissima parte copiato da un articolo di un ricercatore del Monterey Institute of International Studies in California (e pubblicato su una rivista scientifica) senza citare la fonte

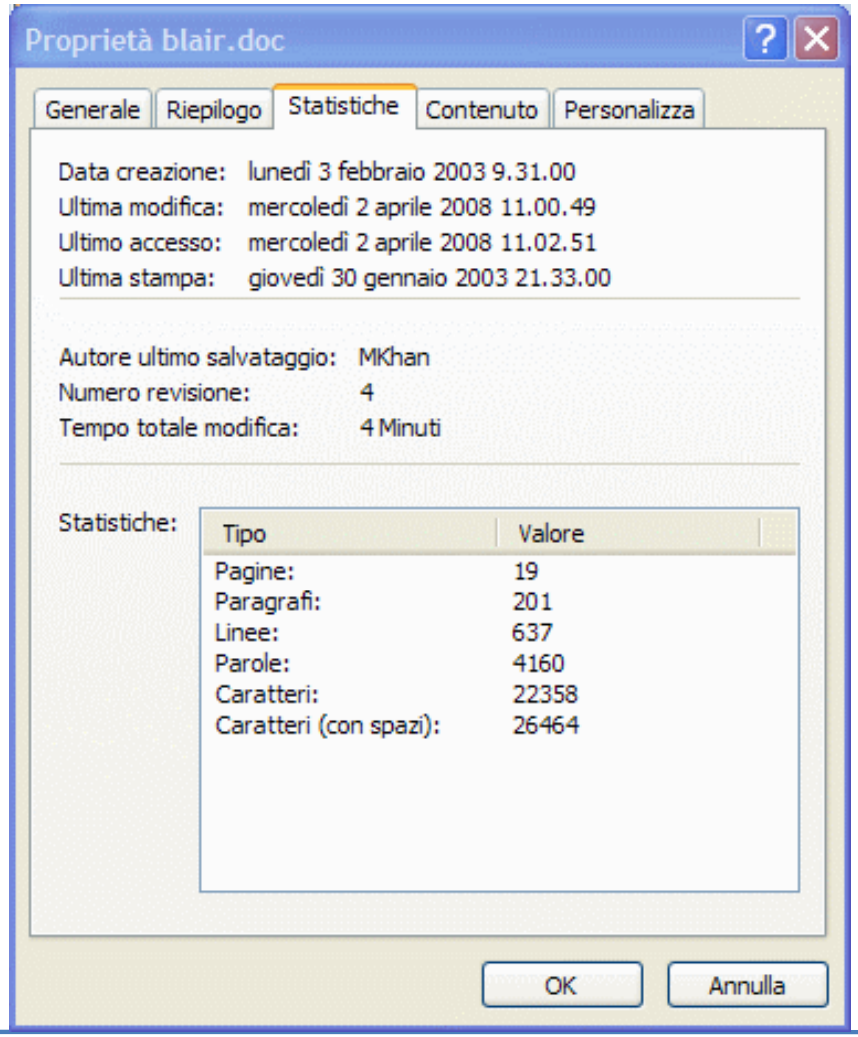
# Analisi a livello delle applicazioni: metadati applicativi – file MS Office

- L'analisi del documento Word (imprudentemente pubblicata da Downing Street) da parte di un analista informatico [BLAIR 2003] ha permesso di identificare gli autori del plagio



# Analisi a livello delle applicazioni: metadati applicativi – file MS Office

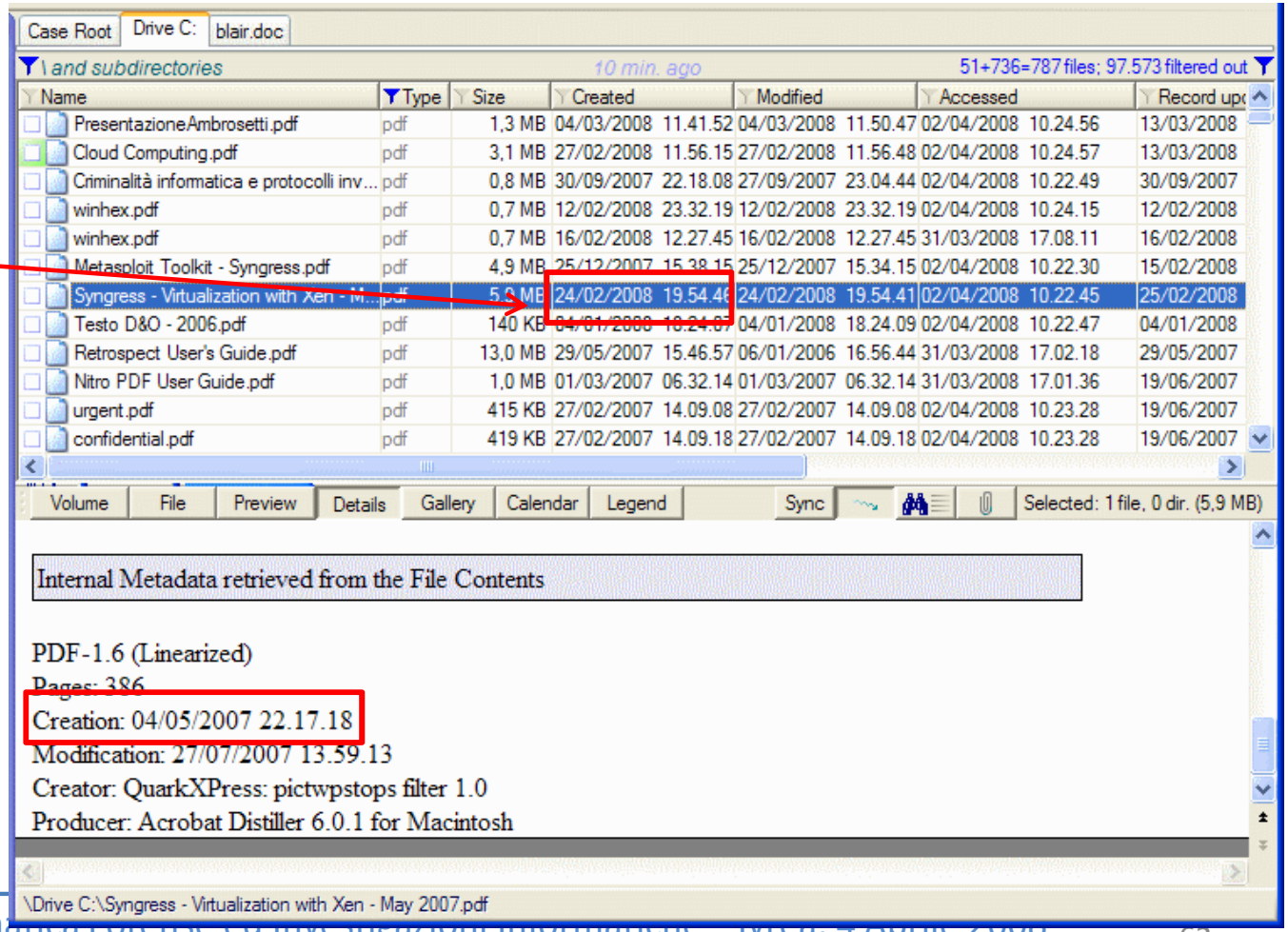
- Si noti che i metadati interni visualizzati da Word riportano l'informazione sul solo autore dell'ultimo salvataggio



# Analisi a livello delle applicazioni: metadati applicativi – file PDF

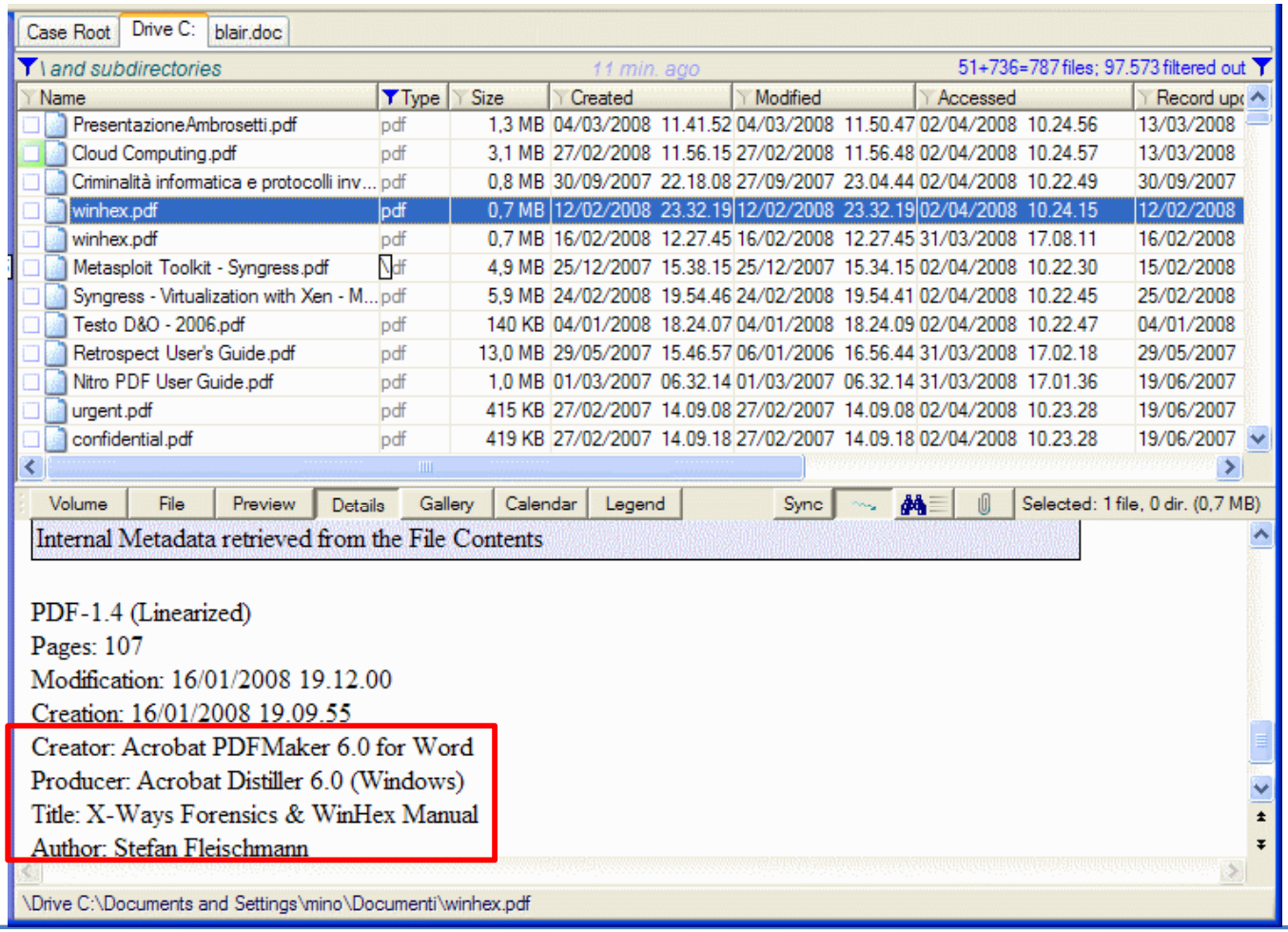
- I metadati sono presenti anche nei file PDF, ma programmi diversi producono quantita' diverse di metadati

La differenza indica che il documento e' stato creato su un altro computer e poi trasferito



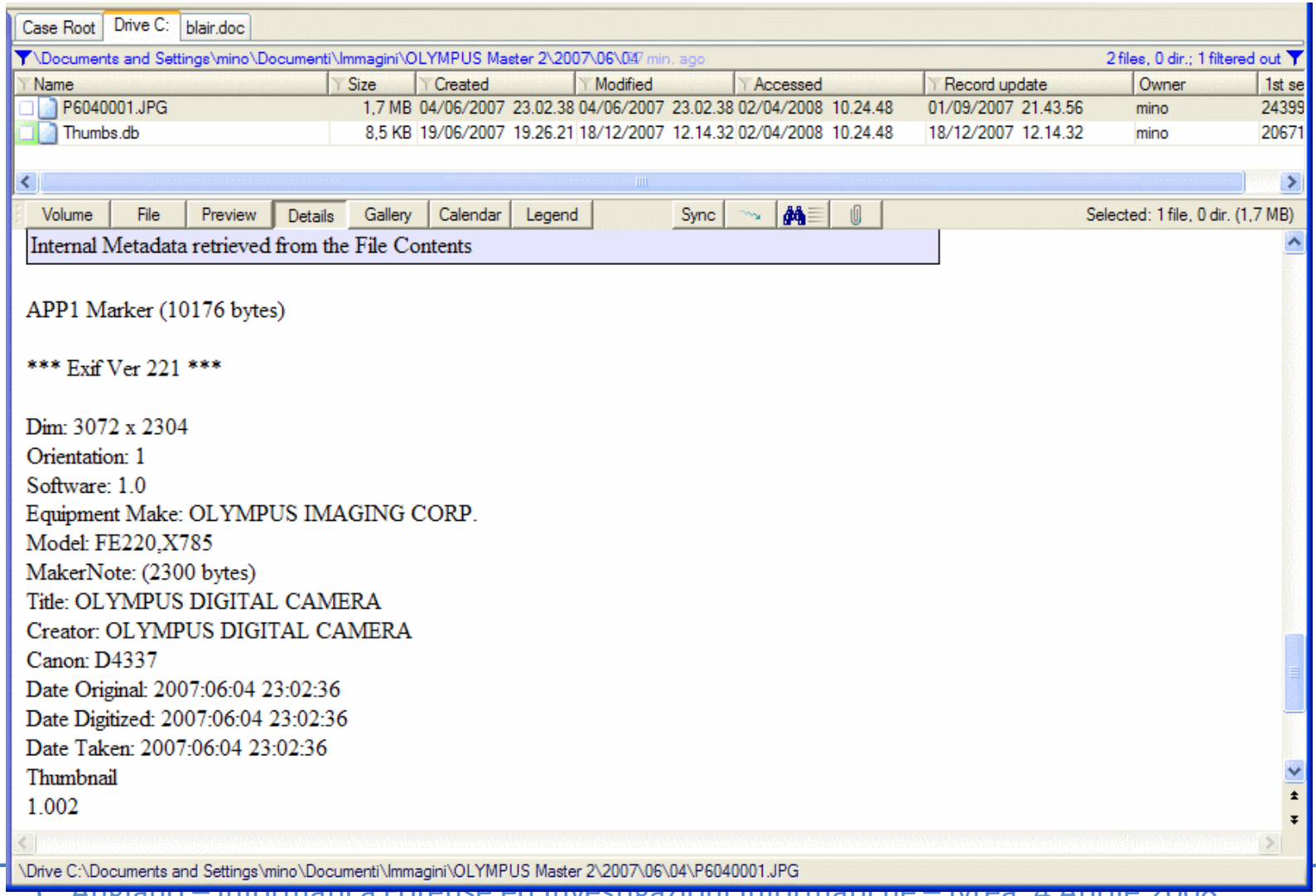
# Analisi a livello delle applicazioni: metadati applicativi – file PDF

- I metadati sono presenti anche nei file PDF, ma programmi diversi producono quantita' diverse di metadati



# Analisi a livello delle applicazioni: metadati applicativi – file grafici

- Metadati spesso utili ai fini probatori sono anche contenuti nelle immagini digitali



# Analisi a livello delle applicazioni: attivita' di navigazione su Internet

- I browser web memorizzano in appositi files varie informazioni quali
  - indirizzo dei siti visitati, e data ed ora della visita
  - ricerche effettuate su motori di ricerca
  - copie temporanee delle pagine visitate (cache)
  - cookies impostati dai siti visitati
- I file in questione, se opportunamente interpretati, permettono di ricostruire l'attivita' di navigazione degli utenti

# Analisi a livello delle applicazioni: attivita' di navigazione su Internet

NetAnalysis - Forensic Internet History Analysis

File Filter Searching Sorting Tools Reports View Column Help

Record URN: 9

| Last Visited [+0100] | Secondary Date | User | URL | Web Page Title | Search Engine Criteria |
|---|---|---|---|---|---|
| 25/08/2004 17.12.52 mer | 15/08/2004 17.53.28 dom | | http://fosi.ural.net/site.html | | |
| 25/08/2004 17.12.50 mer | 01/11/2002 17.10.26 ven | | http://fosi.ural.net/ | | |
| 25/08/2004 17.08.02 mer | 25/08/2004 16.08.02 mer | Mr. Evil | http://www.whatismyip.com | Your ip is 216.62.23.121 WhatIsMyIP.co | |
| 25/08/2004 17.08.02 mer | | mr. evil | http://www.whatismyip.com/ | | |
| 25/08/2004 17.08.02 mer | 25/08/2004 11.08.02 mer | Mr. Evil | http://www.whatismyip.com | | |
| 25/08/2004 17.07.57 mer | 25/08/2004 11.07.57 mer | Mr. Evil | Host: www.whatismyip.com | | |
| 25/08/2004 17.07.51 mer | | mr. evil | http://www.google.com/search?q=what+is+my+ip&hl=en&lr=&ie=UTF-8 | | |
| 25/08/2004 17.07.51 mer | 22/07/2004 22.05.08 gio | mr. evil | http://www.google.com/images/toolbar_promo.gif | | |
| 25/08/2004 17.07.51 mer | 22/03/2004 23.05.03 lun | mr. evil | http://www.google.com/nav_next.gif | | |
| 25/08/2004 17.07.51 mer | 22/03/2004 23.05.03 lun | mr. evil | http://www.google.com/nav_page.gif | | |
| 25/08/2004 17.07.51 mer | 22/03/2004 23.05.03 lun | mr. evil | http://www.google.com/nav_current.gif | | |
| 25/08/2004 17.07.51 mer | 22/03/2004 23.05.03 lun | mr. evil | http://www.google.com/nav_first.gif | | |
| 25/08/2004 17.07.51 mer | 24/08/2004 02.40.46 mar | mr. evil | http://www.google.com/logos/summer2004_synchro_results.gif | | |
| 25/08/2004 17.07.51 mer | 25/08/2004 16.07.51 mer | Mr. Evil | http://www.google.com/search?q=what+is+my+ip&hl=en&lr=&ie=UTF-8 | Google Search: what is my ip | what is my ip |
| 25/08/2004 17.07.51 mer | | | http://www.google.com/search?q=what+is+my+ip&hl=en&lr=&ie=UTF-8 | | what is my ip |
| 25/08/2004 17.07.51 mer | 25/08/2004 11.07.51 mer | Mr. Evil | http://www.google.com/search?q=what+is+my+ip&hl=en&lr=&ie=UTF-8 | | what is my ip |
| 25/08/2004 17.07.32 mer | | mr. evil | http://www.google.com/search?hl=en&ie=UTF-8&q=who+am+i | | |
| 25/08/2004 17.07.32 mer | | | http://www.google.com/search?hl=en&ie=UTF-8&q=who+am+i | | who am i |
| 25/08/2004 17.07.32 mer | 25/08/2004 16.07.32 mer | Mr. Evil | http://www.google.com/search?hl=en&ie=UTF-8&q=who+am+i | Google Search: who am i | who am i |
| 25/08/2004 17.07.32 mer | 25/08/2004 11.07.32 mer | Mr. Evil | http://www.google.com/search?hl=en&ie=UTF-8&q=who+am+i | | who am i |

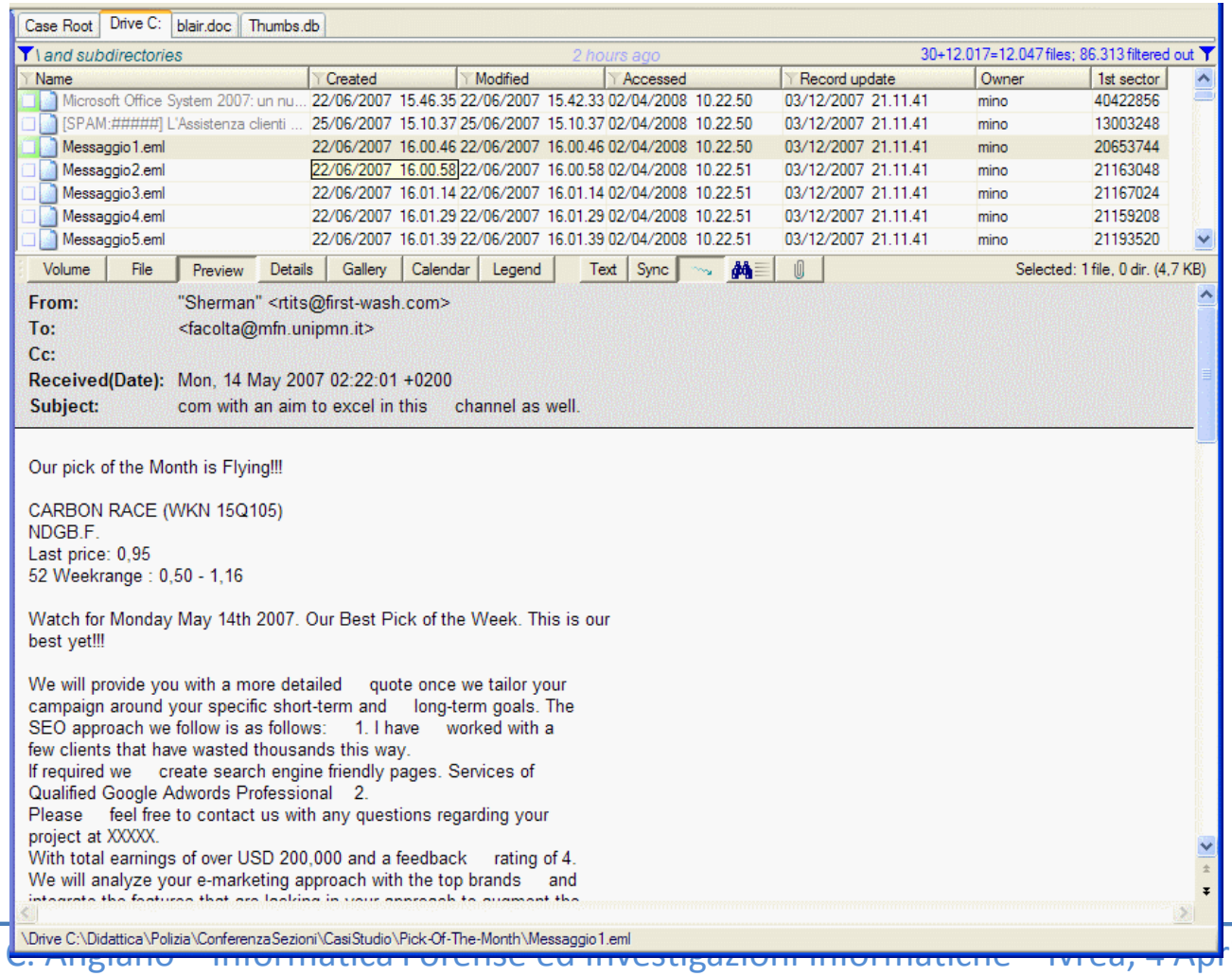
www.digital-detective.co.uk | Type: Cache | ..\Content.IE5\index.dat | Offset: 338.688 | URL Records: 945

# Analisi a livello delle applicazioni: attivita' di navigazione su Internet

- Analisi rese difficoltose da
  - presenza sul mercato di un gran numero di browser
  - differenze nel funzionamento interno dei browser e nella quantita', tipo ed ubicazione delle informazioni salvate

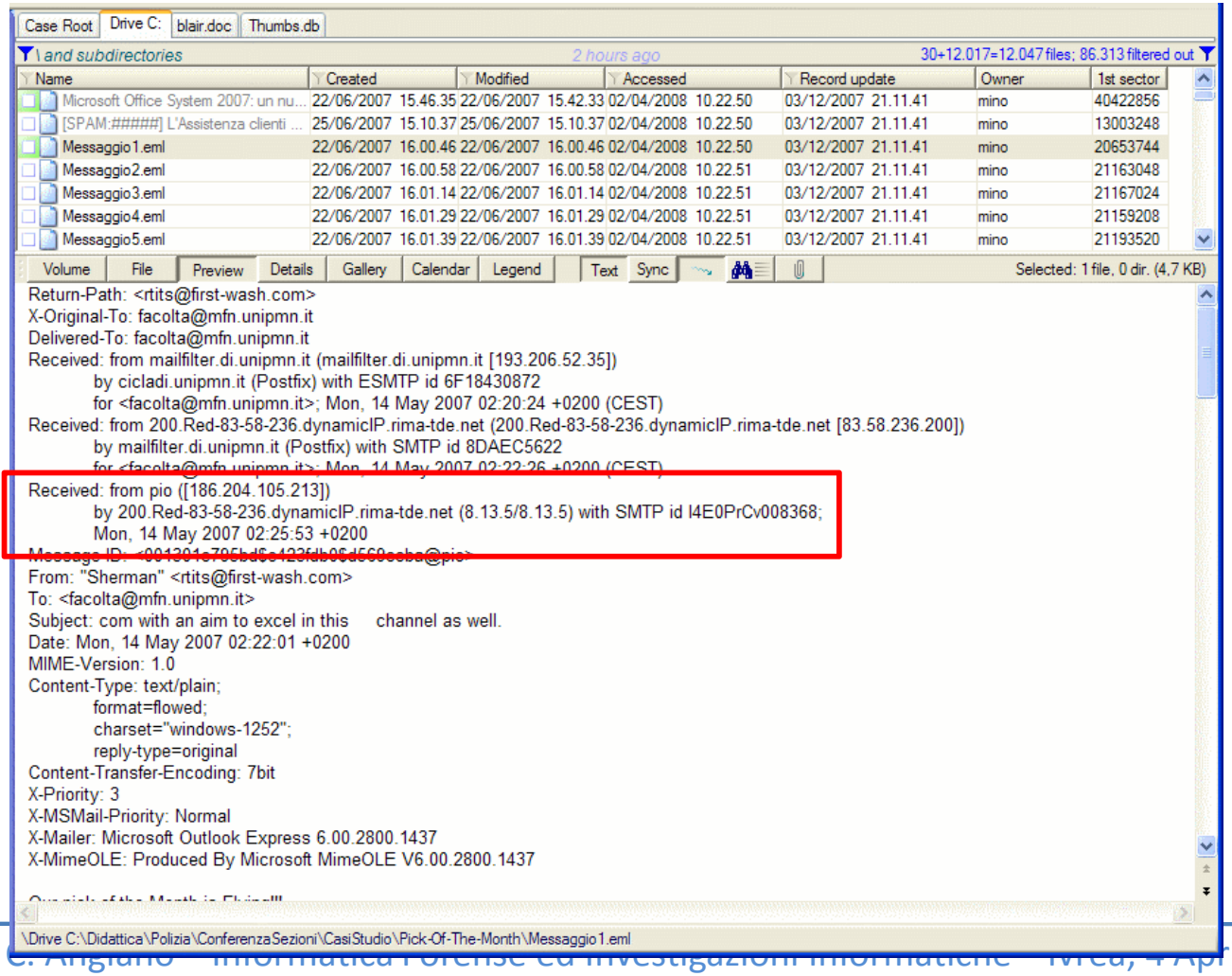
# Analisi a livello delle applicazioni: scambio di email

- Estrazione e decodifica dei messaggi di email salvati localmente dal programma di posta elettronica



# Analisi a livello delle applicazioni: scambio di email

- L'analisi delle intestazioni puo' rivelare il computer da cui il messaggio e' stato spedito e la relativa data



# Analisi a livello delle applicazioni: attivita' di Instant Messaging e Chat

- Alcuni tipi di reato sono compiuti utilizzando (anche) applicativi quali Instant Messaging e Chat
- Tali programmi memorizzano varie informazioni aventi valore probatorio sul computer su cui sono utilizzate
  - log delle conversazioni (data, ora e contenuto)
  - file scambiati
  - lista dei contatti
- Difficolta' dovute a:
  - gran numero di programmi diversi
  - uso di tecniche crittografiche da parte di qualcuno di essi

# Analisi a livello delle applicazioni: Attivita' di File Sharing

- Anche in questo caso, e' possibile recuperare ed interpretare tracce digitali lasciate da praticamente tutti i programmi per il File Sharing
- Come per gli altri tipi di applicazioni, le difficolta' principali sono dovute a
  - documentazione del formato interno dei file scarsa o inesistente
  - utilizzo di tecniche crittografiche

# Correlazione delle evidenze

- Le singole evidenze ottenute nella fase di estrazione devono essere correlate per ottenere una visione di insieme
  - arricchimento della timeline inserendo, oltre ai tempi MACE, anche le informazioni temporali provenienti dal sistema operativo e da programmi applicativi
  - individuazione di evidenze che concorrono a corroborare una determinata ipotesi investigativa

# Correlazione delle evidenze

- Per dimostrare lo svolgimento di attivita' per conto terzi a fini di lucro dell' amministratore di sistema informatico di un'azienda privata, in modo da motivare un licenziamento per giusta causa, sono state correlate le seguenti evidenze digitali:
  - presenza di email, inviate a soggetti terzi, contenenti offerte commerciali inerenti lo sviluppo di siti web
  - presenza di documenti progettuali, redatti dal soggetto in questione (come evidenziato dai metadati Word)
  - rinvenimento del codice sorgente di alcuni siti sviluppati dal soggetto
  - presenza di attivita' di navigazione verso i suddetti siti effettuata in orario di lavoro
  - attivita' di programmazione effettuata in orario di lavoro (risultata da timeline complessiva)

# Correlazione delle evidenze

- Per dimostrare l'avvenuto accesso abusivo ad un sistema informatico di un'azienda con conseguente sottrazione di informazioni riservate, sono state correlate:
  - ritrovamento dei dati riservati su una penna USB di proprieta' del soggetto
  - ritrovamento di tracce che dimostravano che tale penna era stata connessa al PC in uso al soggetto
  - ritrovamento di un'email, inviata da un soggetto terzo (successivamente imputato di concorso in accesso abusivo), che forniva le istruzioni su come connettersi in modo da bypassare le protezioni
  - ritrovamento sul PC di tracce che dimostravano l'avvenuto collegamento al server
  - ritrovamento sul server di tracce che dimostravano l'avvenuto collegamento a partire dal PC del soggetto

# L'Informatica Forense ed i nuovi media

- Negli ultimi anni il numero di dispositivi dotati di capacita' di elaborazione e storage e' cresciuto notevolmente
  - Telefoni cellulari di (pen)ultima generazione
  - Smart Phones
  - Personal Digital Assistants
  - Navigatori satellitari
  - iPod ed iPhone
  - riproduttori MP3
  - consolle per videogame (PS2/3, Xbox, Nintendo, ecc.)

# L'Informatica Forense ed i nuovi media

- Questi dispositivi pongono nuove problematiche di non facile soluzione per le quali esistono soluzioni solo parziali
  - mancanza di standard per l'accesso e l'acquisizione (telefoni cellulari, navigatori satellitari, riproduttori MP3)
  - impiego di sole memorie volatili (PDA)
  - scarsa conoscenza del funzionamento dei sistemi operativi (di tipo proprietario)
  - scarsa conoscenza del formato dei file generati da applicazioni e sistemi operativi

# L'Informatica Forense ed i nuovi media

- La comunita' scientifica ha da poco iniziato a studiare queste problematiche, e le soluzioni cominciano ad apparire solo ora
  - soluzioni parziali
  - sperimentate poco o niente in dibattimento
  - scarso coordinamento tra operatori dell'informatica forense
  - assenza di protocolli operativi condivisi

# 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

# Riferimenti bibliografici


- [ACPO 2007] "*Good Practice Guide for Computer-Based Electronic Evidence*", Maggio 2007. *http://www.acpo.police.uk*
- [NJI 2004] US Department of Justice – National Institute of Justice: *Forensic Examination of Digital Evidence: A Guide for Law Enforcement*", Aprile 2004. *http://www.ncjrs.gov/pdffiles1/nij/199408.pdf*
- [NIJ 2007] US Department of Justice – National Institute of Justice*"Investigative Uses of Technology: Devices, Tools, and Techniques",* Ottobre 2007. *http://www.ncjrs.gov/pdffiles1/nij/213030.pdf*
- [Monga 2006]. M. Monga, "L'intrinseca fragilita' delle tracce digitali", *http://www.avanzata.it/left/traccedigitali.pdf*

# Riferimenti bibliografici


- [SAMMES 2007] T. Sammes, B. Jenkinson. "*Forensic Computing: A Practitioner's Guide, Second Edition*". Springer 2007
- [GHIR 2007] A. Ghirardini, G. Faggioli. "*Computer Forensics*" , Apogeo 2007.
- [SWDGE] The Scientific Working Group on Digital Evidence Web site. *http://www.swgde.org*
- [BLAIR 2003] http://www.computerbytesman.com/privacy/blair.html