Alcune osservazioni sulla *digital evidence* – caratteristiche, ricerca, valutazione

Piras, Luca

2020

Indice

Capitolo 1

Introduzione

La sempre maggiore informatizzazione della società ha delle conseguenze interessanti per le forze dell'ordine.

I reati «classici», che possono essere compiuti anche senza il supporto della tecnologia – ad es., lo spacciatore che si accorda con i clienti tramite messaggi su WhatsApp, invece di incontrarsi di persona – beneficiano dalla diffusione capillare e l'utilizzo massiccio dei dispositivi digitali. È molto probabile che al loro interno si possano trovare informazioni utili per l'accertamento della responsabilità penale.

Inoltre, lo sviluppo della tecnologia ha permesso la nascita di una serie di reati «altamente tecnici», la cui investigazione richiede necessariamente l'uso di strumenti altrettanto sofisticati.[1]

Ad es., la L. n. 547/1993 ha introdotto per la prima volta i reati informatici (*cybercrimes*) nell'ordinamento italiano. La *Convenzione sulla criminalità*

[1] Stefano Di Pinto (2018). «La prova scientifica nel processo penale». In: *Rivista di Polizia*, pp. 909–946. URL: https://www.associazionelaic.it/wp-content/uploads/2019/03/La-prova-scientifica-nel-processo-penale.pdf, p. 913.

informatica firmata a Budapest nel 2001 tratta dei *cybercrimes* commessi via internet:[2]

> La Convenzione è il primo trattato internazionale sulle infrazioni penali commesse via internet e su altre reti informatiche, e tratta in particolare le violazioni dei diritti d'autore, la frode informatica, la pornografia infantile e le violazioni della sicurezza della rete.

Specie nel caso in cui si stiano investigando dei *cybercrimes* – ma sempre più spesso, anche nel caso dei reati tradizionali – ci si trova davanti a fonti di prova di natura **digitale**. Questo report è diviso in tre aree tematiche:

- Nella prima parte, si considerano le **caratteristiche particolari** dei dati digitali, e le loro implicazioni a livello giuridico;
- Nella seconda parte, si considera la qualificazione giuridica della ***digital evidence***, e l'adattamento dei mezzi tradizionali di ricerca delle prove materiali (ispezione, perquisizione e sequestro) ad una prova che ha natura dematerializzata;
- Nella terza parte, sulla base delle questioni analizzate nelle prime due parti, si analizzano i criteri per la **corretta valutazione** della *digital evidence* da parte del giudice, per evitare la sopravvalutazione della prova scientifica, e la «deriva tecnicistica» del processo.

[2] Consiglio d'Europa (n.d.). *Convenzione sulla criminalità informatica*. URL: `https://www.coe.int/it/web/conventions/full-list/-/conventions/treaty/185`.

Capitolo 2

Natura dei dati digitali

2.1 Significato di “digitale”

I dati digitali hanno una natura profondamente diversa rispetto a quella dei normali mezzi di prova materiali. Prima di proseguire con l’analisi, è opportuno evidenziare le loro caratteristiche tecniche, per avere una migliore comprensione delle conseguenze giuridiche.

“Digitale” deriva dalla parola inglese *digit* (cifra), che a sua volta deriva dal latino *digitus* (dito).[1] L’aggettivo viene usato per indicare che le informazioni sono memorizzate come una **sequenza discreta di numeri**.

“Digitale” è contrapposto ad “analogico”, dove le informazioni sono una **sequenza continua** di valori.

Ad es., si pensi alla musica, che consiste in **onde sonore**. Le onde sonore sono una funzione che varia in maniera continua nel tempo.

[1] Vocabolario on line Treccani (n.d.). *Digitale*. URL: http://www.treccani.it/vocabolario/digitale2/

Mediante un dispositivo chiamato **ADC** (*analog-to-digital converter*) è possibile "campionare" (*sampling*) l'onda sonora, e convertirla in una serie di misurazioni discrete. Ad es., i microfoni trasformano l'onda sonora in un segnale elettrico continuo, che poi è convertito in una sequenza di misurazioni digitali dall'ADC.

Un secondo dispositivo chiamato **DAC** (*digital-to-audio converter*) legge i campioni **digitali**, e **riproduce** l'onda sonora originale in formato **analogico**. Ad es., la presa da 3.5 mm comunemente usata per le cuffiette emette un segnale analogico.

La **frequenza di campionamento** (*sampling rate*) utilizzata per i CD audio è 44.1 kHz: ossia, ogni **secondo** di musica è convertito in **44100** valori numerici. È stata scelta perché permette di riprodurre onde sonore che hanno una frequenza massima di 22 kHz,[2] oltre la quale l'udito umano non è capace di percepire suoni.[3]

2.2 Dati, informazioni e codici

È utile operare una distinzione fra i termini "dati" e "informazioni":[4]

- **Dati** – consistono nella "rappresentazione originaria, **non interpretata** di un fatto, fenomeno o evento, effettuata attraverso simboli", e può essere analogico o digitale.

[2]Per il teorema del campionamento di Nyquist–Shannon. Per un video introduttivo all'argomento, v. Technology Connections (2018). *Nyquist–Shannon; The Backbone of Digital Sound*. URL: `https://youtu.be/pWjdWCePgvA`

[3]Per una approfondimento sul rapporto fra la capacità dell'udito umano, la scelta del *sampling rate*, e altre problematiche legate all'audio digitale, v. Chris Montgomery (2012). *24/192 Music Downloads... and why they make no sense*. URL: `https://web.archive.org/web/20200426050432/https://people.xiph.org/~xiphmont/demo/neil-young.html`.

[4]Antonio Gammarota (2016). «Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali». Tesi di dott. Università di Bologna. URL: `http://amsdottorato.unibo.it/7723/`, pp. 46–48.

- **Informazioni** – consistono in "un insieme di dati, che sono stati sottoposti ad un **processo di interpretazione** [...] che li ha resi significativi per il destinatario, e realmente importanti agli scopi prefissi".

Ad es., una **sentenza** può essere considerato un "dato", un semplice "pezzo di carta/documento PDF che contiene parole", ma la sua interpretazione porta alla comprensione delle "informazioni" in essa contenute (in particolare, la decisione ed il dispositivo).

L'insieme di regole che definisce l'**interpretazione** dei **dati digitali** è detto **codice**. Si distingue fra:

- **Codifica** – la conversione di informazioni in dati digitali (ad es., scattare una fotografia con una fotocamera digitale crea una foto in formato JPG);
- **Decodifica** – la conversione di dati digitali in informazioni (ad es., aprire la foto scattata in precedenza in un visualizzatore di immagini riproduce la foto a schermo).

2.3 Bit, byte, sistema esadecimale

In generale, i dati digitali consistono in numeri. La convenzione dominante per la scrittura dei numeri è in base dieci (*decimal*), che prevede dieci simboli, da 0 a 9. Alcuni fra i primi computer utilizzavano la base dieci,[5] tra cui il più noto è ENIAC.[6]

[5] Per una lista v. Wikipedia contributors (n.d.[c]). *Decimal computer*. URL: `https://en.wikipedia.org/wiki/Decimal_computer`.

[6] Nello specifico caso di ENIAC, il sistema decimale è stato scelto per due motivi: meno componenti – *For binary there would be over three times as many gates and power transmitting tubes in each accumulator as for decimal.* – e, principalmente, maggiore facilità di controllo sulla correttezza dei calcoli – *The primary reason was that we thought the users would be better off not having the additional complication of radix conversion every time they wanted*

La convenzione attualmente dominante per i computer è la base due (*binary*), in cui i numeri sono rappresentati usando solo due simboli, convenzionalmente indicati con 0 e 1. L'unità di base nell'informatica è un **bit** (contrazione di *binary digit*), e può rappresentare due possibili valori. Otto bit formano un **byte**, che può rappresentare $2^8 = 256$ possibili valori.[7]

Il sistema binario è semplice da utilizzare per un computer, ma ha lo svantaggio di produrre cifre molto lunghe. Il sistema di numerazione **esadecimale** (*hexadecimal*, spesso abbreviato a *hex*) utilizza sedici simboli: 0–9 (come nel sistema decimale) e A–F (corrispondenti a 10–15), e rappresenta un compromesso tra brevità (per un utente umano) e facilità di conversione in binario (per la macchina).[8]

Sistema	Rappresentazione	Lunghezza
Decimale	65535	5
Binario	0b1111111111111111	16
Esadecimale	0xFFFF	4

Le rappresentazioni binarie ed esadecimali sono convenzionalmente precedute dal prefisso `0x` per evitare ambiguità con quelle decimali.[9]

Un byte può essere rappresentato da due cifre esadecimali, da `0x00` (0) a `0xFF` (256).

to check whether the calculation was going right. John William Mauchly (1980). ≪The Eniac≫. In: *A History of Computing in the Twentieth Century*. Elsevier, pp. 541–550. URL: `https://books.google.it/books?id=AsvSBQAAQBAJ&pg=PA546`, p. 546.

[7] Oltre a valori interi, il sistema binario può rappresentare anche numeri a virgola mobile (*floating-point numbers*)), usando 32 o 64 bit. Lo standard più usato è IEEE-754. Per una introduzione, v. Steve Hollasch (2018). *IEEE Standard 754 Floating Point Numbers*. URL: `https://steve.hollasch.net/cgindex/coding/ieeefloat.html`.

[8] Ogni cifra del sistema esadecimale corrisponde ad una sequenza di quattro bit: `0x0` diventa `0000`, `0x1` diventa `0001`, e così via... fino a `0xF` che diventa `1111`. Questa corrispondenza esatta non esiste per un numero scritto in base dieci.

[9] Esiste anche un terzo tipo di rappresentazione detta ottale (*octal*), che usa i numeri da 0 a 7, ed usa il prefisso "0". Ad es., `010` corrisponde a 8. Uno dei casi in cui è utilizzato sono i permessi per i file di Unix, cfr. Dan's Tools (n.d.). *Unix Permissions Calculator – Further Information*. URL: `http://permissions-calculator.org/info/`.

Esistono strumenti chiamati *hex editors*, che permettono di e **visualizzare** qualsiasi file in formato esadecimale (*hex dump*), e **modificarli**.[10]

2.4 Memorizzazione e trasmissione dei bit

2.4.1 Memorizzazione

Per "esistere" nel mondo "fisico",[11] i bit **necessitano** di due elementi: una **base materiale** su cui essere memorizzati, ed un **mezzo trasmissivo** che permette lo spostamento dei bit fra i vari componenti del computer, e fra vari computer.

Esistono svariate basi materiali su cui memorizzare informazioni digitali.[12]

Il sistema più rudimentale è quello di marcare i singoli bit in maniera "macroscopica", ad esempio come buchi su una **scheda perforata**, o come simboli su un **foglio di carta**.[13]

I **supporti ottici** (CD, DVD, Blu-Ray...) consistono sostanzialmente nella stessa idea, ma a livello microscopico. I bit sono memorizzati sulla superficie di un disco mediante un processo di incisione con laser, e lo stesso laser (usando una potenza ridotta) viene usato per leggere i dati sul supporto.[14]

I **supporti magnetici** consistono in un substrato coperto da un materiale magnetico.[15] Le operazioni di scrittura consistono nella magnetizzazione del

[10] Per un *hex editor* utilizzabile direttamente nel browser, v. `https://hexed.it/`.

[11] Fino ad ora si è parlato in maniera puramente teorica, della natura astratta dei bit. La presente sezione tratta di come i bit "esistono" nel mondo "materiale", e le conseguenze giuridiche della loro natura.

[12] Gammarota, ≪Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali≫, pp. 54–55.

[13] Ad es., i codici a barre e codici QR, che rappresentano una sequenza di numeri o caratteri alfanumerici come linee o quadratini.

[14] Andrew S. Tanenbaum e Todd Austin (2013). *Structured computer organization*. Pearson, pp. 103–108.

[15] Ad es., nel caso di audiocassette e floppy disk, ossido di ferro.

supporto. A differenza dei supporti ottici, che tendono ad essere scrivibili una sola volta, sono riscrivibili. Includono:

- **Audiocassette** – comuni per memorizzare programmi negli anni 70 e 80;[16]
- **Floppy disk** – usati per dati e programmi negli anni 80 e 90, ma attualmente non più in produzione;[17]
- ***Linear tape*** – utilizzato principalmente per l'archiviazione a lungo termine;[18]
- **Hard disk** – di larghissimo utilizzo.

Le **memorie elettroniche** consistono in chip programmabili elettronicamente, e presentano il vantaggio di non richiedere parti mobili.[19] Includono:

- **Cartucce** (*ROM cartridge*) – usate sopratutto da vecchi computer per contenere il sistema operativo[20] e videogiochi, che sono in sola lettura;

[16]Per un esempio di lettore–scrittore di audiocassette collegato ad un computer, v. Wikipedia contributors (n.d.[b]). *Commodore Datasette*. URL: `https://en.wikipedia.org/wiki/Commodore_Datasette`.

[17]Rimangono comunque in uso: *the disks are more widely used than one might expect, especially in industrial machines, aircraft, medical devices and complex hardware systems like those used by the world's militaries.* Cfr. Liam Stack (2019). *Update Complete: U.S. Nuclear Weapons No Longer Need Floppy Disks*. URL: `https://www.nytimes.com/2019/10/24/us/nuclear-weapons-floppy-disks.html`.

[18]Rispetto agli hard disk, le cassette LT ha capacità di archiviazione superiore, ed il costo per GB minore. Un hard disk da 12 TB costa almeno il doppio di una cassetta di dimensioni equivalenti. Tuttavia, il costo dell'hardware per leggere le cassette è nell'ordine delle migliaia di euro, e lettura e scrittura sono strettamente sequenziali. Cfr. ProStorage (2017). *LTO and LTFS: The Pros and Cons of Linear Tape-Open and Linear Tape File System*. URL: `https://getprostorage.com/blog/lto-ltfs-archiving/`.

[19]Muovere un supporto che ha parti mobili può interferire con il normale funzionamento, o danneggiare il supporto stesso. Ad es., gli hard disk fermano la testina se sono mossi per evitare di graffiare i dischi al loro interno. Per una dimostrazione, v. ThioJoe (2018). *What if You SHAKE a Hard Drive WHILE It's Running?* URL: `https://youtu.be/Z3LQX9V90Vo`.

[20]*Most home computers of the 1980s stored a BASIC interpreter or operating system in ROM as other forms of non-volatile storage such as magnetic disk drives were too costly.* Cfr. `https://en.wikipedia.org/wiki/Read-only_memory#Use_for_storing_programs`.

- **Memorie SD** – largamente usate dagli smartphone (che a loro volta utilizzano memoria elettronica al loro interno);
- "**Pennette**" USB;
- **Dischi SSD** – funzionalmente equivalenti ai tradizionali hard disk, ma con il vantaggio di velocità di lettura e scrittura maggiori, e sostanziale immunità agli urti.

2.4.2 Trasmissione

Per quanto riguarda la **trasmissione** dei bit, esistono tre famiglie di metodi:[21]

- **Impulsi elettrici** – i dati binari sono convertiti in impulsi elettrici e trasmessi attraverso un cavo. Per garantire la corretta trasmissione dei bit, è necessario ridurre al minimo le interferenze nel cavo.[22]
- **Onde radio** – i dati sono trasmessi come onde radio. La trasmissione può essere a breve raggio (ad es., Bluetooth e Wi-Fi), o a lungo raggio (ad es., rete dati cellulare, segnali satellitari);[23]

[21] Gammarota, «Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali», pp. 55–56.

[22] *In networking, cable interference can refer to signal disruptions or degradations caused by electrical or electromagnetic sources. [...] With greater cable interference protection, higher transmission rates over longer distances can be supported. [...] The term alien crosstalk refers to the type of cabling interference where one cable affects other nearby cables within a bundle. Closely routed cables can create alien crosstalk [...] Disruptions also occur due to the introduction of Electromagnetic Interference (EMI) into a cable from an external source [...] Both EMI and crosstalk can be factors in attenuation or signal loss in network cabling.* Cfr. Justin Ellis (n.d.). *Cable Interference*. URL: `https://www.comms-express.com/infozone/article/cable-interference/`.

[23] Le stesse osservazioni fatte per l'interferenza nei cavi si applicano anche ai segnali radio: *radio frequency interference is the culprit. RF interference can be generated by almost any device that emits an electro-magnetic signal – from cordless phones to Bluetooth headsets, microwave ovens and even smart meters.* Cfr. David Callisch (2010). *Coping with Wi-Fi's biggest problem: interference*. URL: `https://www.networkworld.com/article/2215287/coping-with-wi-fi-s-biggest-problem-interference.html`.

- **Impulsi ottici** – l'applicazione principe è la fibra ottica nelle connessioni di rete.

2.5 Dai bit alle informazioni

2.5.1 Ambiguità intrinseca dei dati digitali

Come accennato, i dati digitali non hanno alcuno alcun "significato" intrinseco, sono solo una sequenza di valori. Il significato può essere estratto solo al momento della decodifica.

In alcuni casi, i codici possono essere **ambigui**. Ad es., la sequenza di bit `01000001` può essere interpretata in almeno tre modi:

- Come un **numero** a 8-bit – ed assume il valore di 65.
- Come un **carattere** – e secondo la codifica ASCII[24] diventa la lettera "A" maiuscola. Tuttavia, secondo la codifica EBCDIC,[25] non produce un carattere valido.
- Come un **pixel** di una immagine in bianco e nero – se si usa un sistema dove 0 è nero, 255 è bianco, ed i valori intermedi sono sfumature di grigio, corrisponde a ■.

Se dei dati sono decodificati con il **codice sbagliato**, producono informazioni inutili. Ad es., mentre è possibile aprire qualsiasi formato di file con un

[24] Acronimo di *American Standard Code (for) Information Interchange*. È una codifica dei caratteri (*character set*) a 7 bit, ossia, 127 valori corrispondono ad altrettante lettere, numeri, segni di punteggiatura, o sequenze di controllo (ritorno a capo, tabulazioni, carattere *null*, ecc...). Esistono due estensioni di ASCII a 8 bit, ISO 8859-1 e Windows-1252, che introducono supporto per alcuni caratteri accentati e segni tipografici.

[25] Acronimo di *Extended Binary Coded Decimal Interchange Code*. Standard di codifica dei caratteri proposto da IBM, ma che non ha avuto lo stesso successo di ASCII. "A" in EBDIC corrisponde al valore decimale 193.

editor di testo, qualsiasi file che non contenga soltanto testo[26] (ad es., una immagine in formato JPG) produce *mojibake*, un termine giapponese che significa "caratteri incomprensibili".[27]

Ancora, è possibile decodificare qualsiasi file come se fosse un file audio, ma l'operazione produce rumore, non informazione.[28]

In alcuni casi, l'effetto *mojibake* può essere provocato intenzionalmente a fini artistici, come nel caso del *datamoshing*.[29]

2.5.2 Indicazione del formato

Per evitare ambiguità nell'interpretazione, e garantire la **corretta decodifica** dei dati, è necessario **specificare il formato** dei dati.

Una prima possibilità consiste nell'aggiungere **alla fine del nome del file** una sequenza di lettere, dette **estensione del file**, che suggeriscono il formato.[30]

[26] Per convenzione, si distingue tra file di testo (*plaintext*) e binari (*binary*). I primi contengono solo sequenze di byte che corrispondono a caratteri visualizzabili a schermo (*printable range*) – ad es., per ASCII, solo 95 valori corrispondono a caratteri visualizzabili, da 33 a 127 – mentre i secondi possono essere composti da qualsiasi sequenza di byte. Ad es., le e-mail sono file di testo, mentre i loro allegati possono essere anche file binari. I file di testo possono essere aperti con un editor di testo, mentre i file binari richiedono editor specifici per quel formato.

[27] Alcuni sistemi di scrittura utilizzavano un sistema di codifica dei caratteri specifico – per il giapponese, "Shift JIS" – ed incompatibile con la codifica ASCII o Windows-1252. Pertanto, la decodifica di un documento scritto in Shift JIS con la codifica Windows-1252, o viceversa, produceva una sequenza di caratteri incomprensibili. In generale, l'uso di una codifica di caratteri diversa da quella utilizzata per scrivere il documento produce *mojibake*. Per un esempio visivo, v. Wikipedia contributors (n.d.[f]). *Mojibake*. URL: `https://en.wikipedia.org/wiki/Mojibake`.

[28] Hopson (2017). *Secret Binary FIle [sic] Music - MSPaint.exe and aclui.dll*. URL: `https://youtu.be/OfVERK_SreU`.

[29] Per una introduzione alle tecniche di compressione del video, e come vengono "abusate" per realizzare un *datamosh*, v. aescripts + aeplugins (2018). *Datamosh Quick Start Tutorial*. URL: `https://youtu.be/S1jIw4fufP8`.

[30] Tradizionalmente, per limitazioni tecniche ("8.3 filename"), il nome di un file era composto di 8 lettere, più 3 per l'estensione. Le estensioni di formati come HTML o JPEG

Ad es., ".TXT" per i file di testo, ".EXE" per i file eseguibili,[31] o ".PDF" per i documenti. Questo approccio presenta la limitazione di non essere affidabile, perché l'estensione può essere cambiata in maniera arbitraria semplicemente rinominando il file, e di non essere preciso, perché l'estensione non indica la versione specifica del formato utilizzato. L'estensione è al più **indicativa**, ma **non è identificativa** di un formato.

La soluzione miglior consiste nell'inserire un ***magic number***[32] all'**inizio del contenuto del file**.

In generale, non c'è uno standard condiviso: alcuni formati usano una sequenza di caratteri ASCII (ad es., le immagini GIF iniziano con `GIF89a`), altri una sequenza binaria (le immagini JPEG con `0xFFD8FF`).[33]

Per i file audio/video, la convenzione è il "FourCC" (*Four Character Code*), quattro byte che definiscono quattro lettere in ASCII[34] – ad es., `avc1` o `vp80` per due formati di video particolarmente popolari.

L'uso di un *magic number* presenta il vantaggio di essere **più affidabile**, perché modificare i primi byte è più difficile, e **più specifico**, perché è possibile inserire informazioni precise sull'esatta versione del formato del file.[35]

---

venivano accorciati in "HTM" e "JPG". I filesystem moderni non impongono più queste limitazioni, ma fatte salve alcune eccezioni – ad es., i file BitTorrent usano l'estensione ".torrent" – molti dei formati di file più diffusi continuano ad usare estensioni con 3 caratteri.

[31] Sui sistemi Linux, i file eseguibili e di testo spesso non presentano estensione.

[32] Detto anche *file signature*, da non confondere con *digital signature*. Nel primo caso, "signature" va inteso come "marchio caratteristico", mentre nel secondo caso, è la vera e propria "firma digitale", con valenza legale.

[33] Per una lista di *file signatures*, v. Wikipedia contributors (n.d.[e]). *List of file signatures*. URL: `https://en.wikipedia.org/wiki/List_of_file_signatures`.

[34] FourCC.org (2011). *What is a FOURCC?*. URL: `https://www.fourcc.org/fourcc.php`.

[35] Esistono degli strumenti per identificare la tipologia di file: su Linux si usa il comando *file* – v. `https://linux.die.net/man/1/file` – mentre per Windows esiste *TrID* – v. `https://mark0.net/soft-trid-e.html`.

Difatti, la tecnica forense del ***file carving*** si basa proprio sulla ricerca di *magic number* all'interno di un supporto. La ricerca può produrre falsi positivi, ma il vantaggio è che può essere utilizzata per recuperare file, (o quanto rimane di file) all'interno di una qualsiasi sequenza di dati.[36]

[36] *File carving is a powerful tool for recovering files and fragments of files when directory entries are corrupt or missing, as may be the case with old files that have been deleted or when performing an analysis on damaged media. [...] Most file carvers operate by looking for file headers and/or footers, and then "carving out" the blocks between these two boundaries.* Cfr. ForensicsWiki contributors (n.d.[a]). *File Carving*. URL: `https://forensicswiki.xyz/wiki/index.php?title=File_Carving`.

Capitolo 3

Caratteristiche dei dati digitali

Prima di passare all'esame della prova digitale, è opportuno soffermarsi sulle caratteristiche peculiari dei dati digitali,[1] che devono essere tenute a mente per garantire il loro migliore trattamento e valutazione.

3.1 Volatilità

3.1.1 Memorie primarie

La **memoria** è la parte del computer in cui vengono memorizzati dati in forma di bit.[2] Le memoria **primaria** è quella che contiene i dati che il processore sta utilizzando, e tende ad avere dimensioni ridotte, ma velocità

[1] Per l'elenco di riferimento, cfr. Gammarota, «Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali», pp. 57–58.

[2] Tanenbaum e Austin, *Structured computer organization*, pp. 73–74.

di accesso elevato.[3] Si divide in **registri**,[4] **cache**,[5] e **RAM**.[6]

Una caratteristica comune a tutte le memorie primarie è che sono **volatili**: una volta che l'alimentazione viene interrotta, i dati contenuti al loro interno vengono persi quasi istantaneamente,[7] e non c'è modo di recuperarli.[8]

3.1.2 Flusso di dati

La volatilità è una caratteristica anche dei bit che sono **in transito**. Finché non arrivano a destinazione, se la trasmissione viene interrotta per qualsiasi motivo (il cavo viene staccato, il segnale diventa troppo debole...), i bit si "disperdono".

3.1.3 Osservazioni

I dati catturati dalle memorie primarie sono per loro natura estremamente volatili, ed è impossibile acquisirli senza modificarli. La sola apertura di un

[3] Cfr. la gerarchia delle memorie in Tanenbaum e Austin, *Structured computer organization*, pp. 86–87.

[4] *The CPU also contains a small, high-speed memory used to store temporary results and certain control information. This memory is made up of a number of registers, each having has a certain size and function.* Cfr. ibid., p. 56.

[5] Una quantità di memoria più grande dei registri della CPU (misurata in KB e MB), ma più veloce della RAM. In generale, più una memoria è grande, più è lenta: *the bigger the cache, the better it performs, but also the slower it is to access and the more it costs.* Cfr. ibid., pp. 82–84.

[6] La memoria primaria di capacità maggiore (misurata in GB), ma che presenta la velocità di accesso più lenta: *main memories are nearly always built out of dynamic RAMs. However, this large capacity has a price: dynamic RAMs are slow (tens of nanoseconds)*. Cfr. ibid., p. 181.

[7] Gammarota, ≪Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali≫, p. 58.

[8] Ad es., le RAM devono essere "aggiornate" continuamente, altrimenti la carica contenuta al loro interno tende a dissiparsi: *the capacitors can be charged or discharged, allowing 0s and 1s to be stored. Because the electric charge tends to leak out, each bit in a dynamic RAM must be refreshed (reloaded) every few milliseconds to prevent the data from leaking away.* Cfr. Tanenbaum e Austin, *Structured computer organization*, p. 181.

programma per acquisire la RAM comporta che il programma sia caricato nella memoria stessa.

Anche non compiere alcuna operazione porta a modifiche, perché anche se a schermo non si muove nulla, il sistema operativo ed i programmi continuano comunque a lavorare e modificare la memoria.

Pertanto, data l'urgenza e l'indifferibilità dell'operazione, è inevitabile al ricorso alla disciplina degli **accertamenti urgenti** eseguiti dalla P.G.[9]

Gli ufficiali di P.G. devono acquisire il **prima possibile** le memorie volatili, prima di compiere qualsiasi altra operazione,[10] e data la **natura irripetibile** dell'atto, le operazioni devono essere **adeguatamente documentate**, per "garantire, quantomeno, il contraddittorio postumo sull'elemento di prova digitale acquisito".[11]

3.2 Deteriorabilità

3.2.1 Memorie secondarie

Le **memorie secondarie**, dette anche "**memorie di massa**", sono memorie di dimensione maggiore, costi minori, e tempi di accesso più lenti rispetto

[9] *Se vi è pericolo che le cose, le tracce e i luoghi indicati nel comma 1 si alterino o si disperdano ocomunque si modifichino e il pubblico ministero non può intervenire tempestivamente, ovvero non ha ancora assunto la direzione delle indagini, gli ufficiali di polizia giudiziaria [...].* Cfr. art. 254 co. 2 c.p.p.

[10] *Under the principle of "order of Volatility", you must first collect information that is classified as Volatile Data (the list of network connections, the list of running processes, logon sessions, and so on), which will be irretrievably lost in case the computer is powered off.* Cfr. Ayman Shaaban e Konstantin Sapronov (2016). *Practical Windows Forensics*. Packt Publishing, p. 25.

[11] Marco Torre (2015). «Aspetti giuridici e tecnici relativi al trattamento della prova digitale nel processo penale. La prova informatica nella legge 18 marzo 2008, n. 48». In: *Informatica e diritto* 24.1-2, pp. 65–104. URL: http://www.ittig.cnr.it/risorse/attivita-editoriale/rivista/indici/1-2-2015/, p. 83.

alle memorie primarie. Il loro scopo è di conservare i dati in maniera duratura, anche in assenza di alimentazione.

Tuttavia, anche questo tipo di memoria è "volatile" (in senso lato). Più specificamente, i supporti materiali sono soggetti a deterioramento, sia con l'uso che con il passare del tempo, e ciò si riflette sui bit che vi sono ancorati. Ad esempio:

- La carta ed il substrato dei nastri magnetici e dei supporti ottici sono soggetti a graduale decomposizione;
- I supporti magnetici possono smagnetizzarsi nel tempo;[12]
- Le parti in movimento degli hard disk e lettori dei supporti ottici possono subire guasti meccanici, e danneggiare i dati;[13]
- Gli SSD memorizzano informazioni come cariche elettriche, che tendono a dissiparsi nel tempo,[14] ed i componenti elettronici che le contengono sono soggeti ad usura.[15]

[12]Per una dimostrazione della smagnetizzazione di un floppy disk, v. Nostalgia Nerd (2018). *Magnets vs. Floppy Disks — Nostalgia Nerd*. URL: `https://youtu.be/Nn2R7bIzDtc`.

[13]Per gli hard disk esiste l'espressione "click of death": *Click of death [...] signals a disk drive has failed, often catastrophically. [...] If the [disk] head fails to move as expected or upon moving cannot track the disk surface correctly, the disk controller may attempt to recover from the error by returning the head to its home position and then retrying, at times causing an audible "click"*. Cfr. Wikipedia contributors (n.d.[a]). *Click of death*. URL: `https://en.wikipedia.org/wiki/Click_of_death`.

[14]*NAND [Il nome della tecnologia alla base degli SSD] can't retain data forever, or even as long as other types of media (hard drives, optical) because it stores data as tiny, trapped electrical charges. The cages that contain these charges aren't perfect—they leak [...] extremely slowly, but they leak nonetheless. [...] [T]he warmer they get, the faster they leak, and the faster operations wear them out. Cell degradation occurs whether an SSD is in use (powered/operational) or stored (unpowered/non-operational)*. Cfr. Jon L. Jacobi (2015). *Death and the unplugged SSD: How much you really need to worry about data retention*. URL: `https://www.pcworld.com/article/2921590/`.

[15]*A typical flash cell can be written only about 100,000 times before it will no longer function. The process of injecting electrons into the floating gate slowly damages it and the surrounding insulators, until it can no longer function. To increase the lifetime of SSDs, a technique called wear leveling is used to spread writes out to all flash cells in the disk.* Cfr. Tanenbaum e Austin, *Structured computer organization*, p. 99.

3.2.2 Gradi di gravità del deterioramento

La gravità del deterioramento dei bit può essere classificata in vari livelli.

Non-catastrofico

Porta alla graduale corruzione (*bit rot*) dei dati, che rimangono comunque decodificabili, ma presentano errori.[16]

Catastrofico

Il livello di corruzione è tale che i dati non sono più decodificabili, o il dispositivo non è più riconosciuto.[17]

Il deterioramento catastrofico può essere dovuto a cause naturali, o volontarie, e avvenire a tre livelli:

- **Logico** – il dispositivo perde tutti i dati, ma rimane comunque utilizzabile. Ad es., il ripristino ai dati di fabbrica (*factory reset*) di uno smartphone, o le conseguenze di un attacco malware.[18]
- **Firmware** – il *firmware* consiste nella porzione di software che è a diretto contatto con l'hardware. Se il firmware diventa corrotto, la

[16] Per una dimostrazione visiva degli effetti del bit rot, v. `https://en.wikipedia.org/wiki/Data_degradation#Visual_example`.

[17] Ad es., i programmi che avvisano l'utente che un file non è in un formato leggibile. Cfr. l'aggettivo inglese "bricked", utilizzato per definire lo stato di dispositivo elettronico che non risce ad avviarsi, od offire le funzionalità di base correttamente, e quindi è utile solo come un "mattone", o fermacarte.

[18] Il ransomware Petya corrompe due componenti del disco, il *Master Boot Record* (che contiene informazioni vitali per l'avvio del sistema operativo) e il *Master File Table* (che contiene informazioni riguardo a dove i file si trovano su disco. Il risultato è che mentre l'hard disk continua a funzionare, il sistema operativo ed i dati sono di fatto inaccessibili all'utente, perché le strutture necessarie per accedervi sono state distrutte. L'unico rimedio è reinstallare il sistema operativo. Cfr. Matt Williams (2017). *Safeguarding Data: How Ransomware Can Affect the Master Boot Record (MBR)*. URL: `https://www.faronics.com/news/blog/safeguarding-data-ransomware-can-affect-master-boot-record-mbr`.

parte hardware può dare segni di vita, ma non sarà possibile utilizzarla fino a quando il firmware viene riparato (sempre che sia possibile);[19]

- **Hardware** – l'hardware non dà segni di vita, o presenta danni che rendono impossibile il normale funzionamento. I danni all'hardware sono generalmente irreversibili,[20] e comportano l'impossibilità di accedere ai dati.[21] In questi casi, l'unica possibilità è l'uso di complesse (e costose) procedure di *data recovery*.[22]

3.2.3 Osservazioni

I dati delle memorie secondarie sono più stabili, ma il deterioramento è comunque un fenomeno da non sottovalutare.

Per precauzione, è **necessario rigettare** l'opinione della Cassazione (ad es., sent. Cass. n. 14511/2009)[23] per cui la procedura di **acquisizione** consista in un accertamento **ripetibile**.[24] Ogni operazione di acquisizione ha una probabilità – seppure minima, ma non inesistente – di modificare i dati, o danneggiare l'hardware, tanto per malfunzionamento improvviso dell'hard-

[19] Ad es., le operazioni di "rooting" o "jailbreaking", con cui si rimuovono le limitazioni imposte dal produttore sul dispositivo, vanno a modificare componenti software fondamentali per il funzionamento del dispositivo, ed in caso di malfunzionamento, è possibile che il dispositivo risulti inutilizzabile. Cfr. The iPhone Wiki contributors (n.d.). *Brick*. URL: `https://www.theiphonewiki.com/wiki/Brick`.

[20] Data la complessità ed il livello di miniaturizzazione raggiunto della tecnologia moderna, è difficile trovare un "pezzo di ricambio" che possa sostituire il componente hardware danneggiato, ed effettuare la sostituzione.

[21] Anche se l'hardware venisse riparato, rimane comunque la possibilità di danni permanenti ai dati: ad es., nel caso di una testina di un hard disk che graffia la superficie dei dischi, anche se venisse riparata, i dati situati nelle porzioni graffiate sono comunque persi.

[22] Ad es., le operazioni di data recovery sono utilizzate sulle le scatole nere degli aeroplani, che sono portate il prima possibile in laboratori specializzati per garantire la migliore riuscita del recupero dei dati. Cfr. Ontrack.com (2015). *Recovering data from black boxes*. URL: `https://www.ontrack.com/uk/blog/concepts-explained/recovering-data-black-boxes/`.

[23] Gammarota, ≪Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali≫, pp. 158–160.

[24] Cfr. art. 359 c.p.p.

ware, o per errori da parte dell'operatore (ad es., l'elettricità statica può danneggiare i componenti elettronici). Pertanto, nell'interesse della migliore conservazione dei dati, e più completa documentazione della procedura, è meglio considerare l'operazione come **irripetibile**,[25] così che possa essere supervisionata anche dal C.T. di parte.

Nel caso in cui i contenuti della memoria contengano prove a discarico, è opportuno richiedere l'incidente probatorio, perché porta all'assunzione di una prova completa.[26]

È bene fare attenzione anche al **deterioramento delle copie acquisite**. I supporti materiali che contengono la copia acquisita potrebbero non essere più funzionanti a distanza di anni.

La soluzione più semplice è di fare più copie di backup dei supporti materiali nel tempo. La soluzione ideale sarebbe di dotare i tribunali di un sistema di *storage* per l'archiviazione a lungo termine di dati digitali, con meccanismi che garantiscano l'integrità dei dati, in maniera simile a sistemi cloud come Dropbox o Drive.

In un certo senso, l'"archivio riservato" per le intercettazioni rappresenta quasi un antecedente per un ipotetico sistema di archiviazione dei dati digitali:

> I verbali e le registrazioni, e ogni altro atto ad esse relativo, sono conservati integralmente in apposito archivio riservato presso l'ufficio del pubblico ministero che ha richiesto ed eseguito le intercettazioni, e sono coperti da segreto.[27]

Volendo, sarebbe possibile per i tribunali appoggiarsi a strutture di privati già esistenti. Sempre nell'ambito delle intercettazioni, esiste già un precedente

[25] Cfr. art. 360 c.p.p.
[26] Cfr. art. 360 co. 4 c.p.p. e art. 392 e ss. c.p.p.
[27] Cfr. art. 269 co. 1 c.p.p.

per l'uso di strutture private nell'ambito di operazioni informatiche:

> Quando si procede a intercettazione di comunicazioni informatiche o telematiche, il pubblico ministero può disporre che le operazioni siano compiute anche mediante impianti appartenenti a privati.[28]

Ancora, le disposizioni sul sequestro permettono che la cosa sequestrata possa essere conservata presso terzi:

> Le cose sequestrate sono affidate in custodia alla cancelleria o alla segreteria. Quando ciò non è possibile o non è opportuno, l'autorità giudiziaria dispone che la custodia avvenga in luogo diverso, determinandone il modo e nominando un altro custode, idoneo a norma dell'articolo 120. [...] Quando la custodia riguarda dati, informazioni o programmi informatici, il custode è altresì avvertito dell'obbligo di impedirne l'alterazione o l'accesso da parte di terzi, salva, in quest'ultimo caso, diversa disposizione dell'autorità giudiziaria.[29]

In ogni caso, ogni soluzione di storage a lungo termine deve essere bilanciata con la tutela della privacy, e la distruzione sicura dei dati una volta che non sono più necessari.

3.3 Modificabilità

3.3.1 Irreversibilità delle modifiche

Ogni memoria consiste in una serie di **celle**, che possono contenere **valori**, e che sono contrassegnate da un **indirizzo**. Per convenzione, il primo

[28] Cfr. art. 268 co. 3-bis c.p.p.
[29] Cfr. art. 259 c.p.p.

indirizzo è 0, poi seguito da 1, 2, 3... L'indirizzo può indicare i singoli bit della memoria, oppure un byte, oppure blocchi di byte.[30]

Le operazioni di **lettura** richiedono solo l'indirizzo da cui leggere, mentre quelle di **scrittura** richiedono il valore da assegnare a quell'indirizzo.

Ogni operazione di scrittura è **irreversibile** – in altre parole, una volta scritto un certo valore ad un certo indirizzo, non c'è modo di recuperare il valore memorizzato in precedenza a quello stesso indirizzo.[31]

3.3.2 *Media sanitization*

La caratteristica dell'irreversibilità è rilevante per la cancellazione sicura dei dati (*media sanitization*). Nel relativo standard NIST sono definiti tre livelli, uno più sicuro e distruttivo del precedente.[32]

Clear

Il primo livello opera una degradazione al solo livello "logico", ed usa solo comandi **software**:

> One method to sanitize media is to use software or hardware products to overwrite user-addressable storage space on the

[30] Tanenbaum e Austin, *Structured computer organization*, pp. 74–75.

[31] Tecniche che permettono di recuperare i bit sovrascritti esistono, ma richiedono apparecchiature specializzate. *Faced with techniques such as MFM [magnetic force microscopy], truly deleting data from magnetic media is very difficult. The problem lies in the fact that when data is written to the medium, the write head sets the polarity of most, but not all, of the magnetic domains. [...] The recovery of at least one or two layers of overwritten data isn't too hard to perform [...].* Cfr. Peter Gutmann (1996). ≪Secure Deletion of Data from Magnetic and Solid-State Memory≫. In: *Sixth USENIX Security Symposium Proceedings*. URL: http://softpres.org/cache/SecureDeletionOfDataFromMagneticAndSolidStateMemory.pdf.

[32] Richard Kissel et al. (2014). *Guidelines for media sanitization*. US Department of Commerce, National Institute of Standards e Technology. URL: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-88r1.pdf, pp. 24–25.

> media with non-sensitive data, using the standard read and write commands for the device.[33]

La modifica consiste in una **sovrascrittura** dei dati, in maniera analoga al passare un pennarello indelebile su un foglio di carta.

Può avere ad oggetto sia **singoli file**,[34] che l'**intero disco**.[35]

Purge

Il secondo livello opera sempre a livello logico, ma usa comandi del **firmware** per aggirare le restrizioni imposte dal software:

> Some methods of purging [...] include overwrite, block erase, and Cryptographic Erase, through the use of dedicated, standardized device sanitize commands that apply media-specific techniques to bypass the abstraction inherent in typical read and write commands.[36]

Per **firmware** si intende:

> A specific class of computer software that provides the low-level control for a device's specific hardware. Firmware can [...]

[33] Kissel et al., *Guidelines for media sanitization*, p. 24.

[34] *Secure delete applications overwrite a deleted file's on-disk data using techiques that are shown to make disk data unrecoverable, even using recovery technology that can read patterns in magnetic media that reveal weakly deleted files. [...] What is more tricky is securely deleting Windows NT/2K compressed, encrypted and sparse files, and securely cleansing disk free spaces.* Cfr. Mark Russinovich (2018). *SDelete v2.02*. URL: `https://docs.microsoft.com/en-us/sysinternals/downloads/sdelete`.

[35] Il comando *dd* presente sui sistemi GNU/Linux viene scherzosamente soprannominato *disk destroyer* perché un errore di battitura può portare alla completa sovrascrittura dei contenuti di un disco. L'esecuzione del comando `dd if=/dev/zero of=/dev/sda` sovrascrive ogni byte sul disco con zeri.

[36] Kissel et al., *Guidelines for media sanitization*, p. 24.

> provide a standardized operating environment for more complex device software[...].[37]

I vari componenti hardware presentano ognuno caratteristiche di fabbricazione diverse. Il software **non è, e non deve** esserne a conoscenza. L'intermediario fra hardware e software consiste nell'**astrazione** fornita dal **firmware**. Ad esempio, negli SSD l'algoritmo di *wear leveling* è definito a livello di firmware, e serve a distribuisce le scritture sul disco in maniera uniforme per aumentarne la vita utile.

Di fatto, solo il firmware è a conoscenza di quali indirizzi di celle "virtuali" corrispondono a quali celle "materiali". Più scritture a livello software allo stesso indirizzo "virtuale" risultano in scritture di bit su posizioni "materiali" diverse. Il software non sa, né può sapere "dove" (nel senso materiale) si trovano i dati, dato che rimane prerogativa del firmware.

Pertanto, anche se si elimina o sovrascrive un file con software di cancellazione sicura, a causa del *wear leveling* è possibile che rimangano tracce. L'unico modo per garantire la distruzione effettiva dei dati è di inviare un **comando di cancellazione sicura** (`SECURITY_ERASE`) **direttamente al firmware**:[38]

Destroy

L'ultimo livello opera a livello fisico. Consiste nella distruzione fisica dell'hardware, in modo che i dati da eliminare non possano essere recuperati:

> The device is not considered Destroyed unless Target Data retrieval is infeasible using state of the art laboratory techniques.[39]

[37] Wikipedia contributors (n.d.[d]). *Firmware*. URL: `https://en.wikipedia.org/wiki/Firmware`.

[38] *When a Secure Erase is issued against a SSD drive all its cells will be marked as empty [...]. This will erase all your data, and will not be recoverable by even data recovery services.* Cfr. Kernel.org Wiki contributors (n.d.). *ATA Secure Erase*. URL: `https://ata.wiki.kernel.org/index.php/ATA_Secure_Erase`.

[39] Kissel et al., *Guidelines for media sanitization*, p. 25.

3.3.3 *Information security* e *access control*

In primo luogo, **non ogni modifica è possibile**. Qui assume particolare importanza la disciplina dell'***information security*** (abbreviata in *infosec*):

> Information security ensures the confidentiality, availability and integrity of information. Information security involves the application and management of appropriate controls that involves consideration of a wide range of threats [...][40]

L'implementazione di **controlli appropriati** permette di proteggere le informazioni da modifiche indesiderate. Uno di questi controlli è la definizione di regole per il controllo degli accessi (*access control*), che serve a prevenire accessi non autorizzati.[41]

Una implementazione di controllo degli accessi consiste nella creazione di **account utente**, ognuno dotato di **determinati privilegi**, e protetto da **password**. L'account "amministratore" ha pieni privilegi, e può compiere qualsiasi operazione, mentre gli account "utente" hanno privilegi limitati: ad es., non possono accedere a cartelle e file di cui un altro utente è "proprietario", o non possono eseguire determinati comandi.

La stessa logica è applicabile anche ai siti internet: il "deep web" consiste in contenuti non pubblicamente disponibili, o perché è necessario avere un **account** (*login wall*), o perché è necessario **pagare** (*paywall*), o perché anche avendo un account, non si dispone delle necessarie **autorizzazioni** (ad es., i controlli per la privacy sui social media che permettono di impostare un account o determinati contenuti come "privati", e quindi visibili sono dagli account che il proprietario decide di autorizzare).

[40] International Organization for Standardization (ISO) (2018). *Information technology — Security techniques — Information security management systems — Overview and vocabular*. URL: https://www.iso.org/standard/73906.html, p. 12.

[41] Georg Disterer (2013). «ISO/IEC 27000, 27001 and 27002 for Information Security Management». In: *Journal of Information Security* 4.2, pp. 92–100. URL: http://dx.doi.org/10.4236/jis.2013.42011, p. 96.

3.3.4 Anonimità

Fungibilità dei bit

I bit sono anonimi nel senso che sono **perfettamente fungibili**. In altre parole, "un bit vale l'altro", perché i **singoli bit** "non hanno memoria", ma contengono solo dati.[42]

Se si prendono una o più posizioni di memoria, è impossibile sapere quando sono state cambiate l'ultima volta, nel contesto di quale operazione, o persino da quale operatore umano.

Incertezza intrinseca

Chiunque può eseguire un comando, e modificare i bit di una memoria secondaria in maniera (quasi, tendenzialmente) arbitraria (sempre nei limiti dell'*access control*). Così come nei reati tradizionali, il **risultato** è sempre ben visibile, ma determinare il **processo** che ha portato a quel risultato è difficile, anche per l'intrinseca deteriorabilità e modificabilità dei dati.

A dimostrazione della incertezza specifico ambito dei *cybercrimes*, l'identificazione dei responsabili di un crimine commesso su internet è realizzata con la ricerca dell'indirizzo IP. In teoria, l'indirizzo IP "corrisponde" ad una persona fisica. Tuttavia, in pratica, questa corrispondenza è molto approssimativa. Esistono vari elementi che interferiscono con questa equivalenza:[43]

- Un utente può mascherare il proprio indirizzo IP usando servizi di proxy (VPN, Tor[44]...);

[42] L'affermazione è solo apparentemente paradossale: dato che il bit consiste nella minima unità di memoria, non può "avere memoria" di sé stesso. Per conservare informazioni riguardanti un bit, è necessario usare altri bit.

[43] Marco Pittiruti (2015). ≪Le indagini informatiche nel processo penale≫. Tesi di dott. Università degli Studi Roma Tre. URL: http://hdl.handle.net/2307/5026, p. 20.

[44] V. https://www.torproject.org/

- Gli indirizzi IP ad uso domestico sono assegnati dinamicamente, cosi ché un singolo indirizzo IP è associato a più utenti nel corso del tempo;
- Gli indirizzi IP da soli spesso non corrispondono allo specifico dispositivo, ma piuttosto al router a cui il dispositivo è connesso;
- In ogni caso, non c'è necessariamente corrispondenza tra l'utenza del router, il proprietario del dispositivo, e la persona indagata o imputata, perché nulla vieta di usare un dispositivo altrui, che è collegato ad una router che è intesto ad altri, per commettere un crimine informatico.

Correlazione dei dati

Per comprendere l'autore, la data, l'origine, la destinazione, ecc... di un dati informatici – tutti elementi di fondamentale importanza per l'accertamento dei reati – è necessario considerarli nel **contesto digitale** più ampio possibile. Più elementi si hanno a disposizione, e più la ricostruzione dei fatti può essere precisa.

Ad es., se un determinato file si trova nella cartella "Downloads", è ragionevole controllare la *cache* o la cronologia del browser per verificare da quale sito quel file sia stato scaricato, a che giorno e ora; per verificare che uso ne è stato fatto, si può eseguire una ricerca per il nome di quel file, che potrebbe comparire nei file di *log* di un altro programma; se non si trovano elementi che dimostrano che quel file è stato scaricato, allora si può presumere che sia stato aggiunto da qualcuno, ecc...

Inoltre, i dati digitali non devono essere presi in isolamento, ma devono essere correlati anche alla **realtà materiale**. Anche nel caso di *cybercrimes* "puri", la *digital evidence* è sicuramente la principale, ma mai l'unica fonte di prova disponibile.[45]

[45] *Solo una correlazione di più elementi, non per forza di natura esclusivamente digitale, permette di contestualizzare l'evidenza informatica, facendole assumere il ruolo ben più importante di fonte di prova.* Cfr. Maurizio Tonnellotto (2014). ≪Evidenza informatica,

In particolare, il punto più delicato è proprio l'identificazione della persona fisica responsabile delle azioni:

> Conoscere le abitudini, il livello di competenze informatiche, le caratteristiche del soggetto, permetterà una più corretta e quantomai completa correlazione degli elementi anche non di natura digitale in possesso agli investigatori, al fine di poter addivenire ad un quadro ampio ed esaustivo del fatto investigato.[46]

Si pensi ad un ipotetico caso di accesso non autorizzato a sistema informatico, ed il titolare dell'indirizzo IP da cui ha avuto origine l'attacco è un pensionato di età avanzata, che vive da solo, e usa il computer solo sporadicamente. È irragionevole pensare che il pensionato sia il responsabile dell'attacco, mentre è molto più probabile che sia a sua volta una vittima, e che il suo computer sia stato infettato da malware controllato a distanza (*trojan*).

3.3.5 Metadati

Definizione di metadati

I dati digitali non sono completamente anonimi. Un elemento utile è dato dai **metadati**, ossia **dati** che contengono **informazioni** riguardo **altri dati**.

Alcuni tipi di file hanno degli standard appositi. Ad es., le immagini digitali in formato JPEG usano lo standard EXIF:

> Specifically, digital camera pictures may contain an Extended File Information (EXIF) header, which saves information about the camera that took the picture. [...] By reviewing EXIF headers,

computer forensics e best practices». In: *Rivista di Criminologia, Vittimologia e Sicurezza* 8.2, pp. 68–103. URL: http://www.vittimologia.it/rivista/articolo_tonellotto_2014-02.pdf, p. 71.

[46] Tonnellotto, «Evidenza informatica, computer forensics e best practices», p. 88.

> some valuable information can be recovered. [...] The date/time the picture was taken will not change, even if the file is copied to another medium. The EXIF header also shows the camera make and model.[47]

Header

Più in generale, molti tipi di file hanno un ***header***, dati situati all'inizio del file. In particolare per le e-mail, l'*header* contiene informazioni riguardo alla mail stessa, ed in particolare, riguardo al mittente, ed ai passaggi che ha seguito per arrivare al destinatario:

> All email messages contain a header, located at the top of the email. The header contains the source of an email in the "From" line, while in the "Received" lines, the header lists every point the email passed through on its journey, along with the date and time. The message header provides an audit trail of every machine the email has passed through.[48]

File di log

Ancora, moltissimi software producono **file di log**. Un "log" è una traccia delle operazioni compiute dal software, spesso accompagnata da riferimenti temporali. Nascono per l'esigenza degli sviluppatori di avere riferimenti per diagnosticare la causa del malfunzionamento del software. Tuttavia, sono estremamente utili anche per l'accertamento dei *cybercrimes*. I file di log possono:

[47] Paul Alvarez (2004). «Using extended file information (EXIF) file headers in digital evidence analysis». In: *International Journal of Digital Evidence* 2.3. URL: https://www.utica.edu/academic/institutes/ecii/publications/articles/A0B1F944-FF4E-4788-E75541A7418DAE24.pdf.

[48] Natarajan Meghanathan, Sumanth Reddy Allam e Loretta A. Moore (2009). «Tools and techniques for network forensics». In: *International Journal of Network Security & Its Applications (IJNSA)* 1.1, pp. 14–25. URL: https://arxiv.org/abs/1004.0570, p. 15.

- Servire per **correlare indizi** – se un file è menzionato all'interno di un programma regolarmente usato dall'utente, è probabile che l'utente abbia volontariamente scaricato quel file;
- Contenere tracce utili per risalire all'**identità del soggetto** che ha causato il danno – si pensi ai file di log dei server, che contengono anche indirizzi IP;
- Indicare **quali operazioni** sono state commesse – e quindi aiutare gli investigatori nella ricostruzione della dinamica di un incidente.

Attributi

Un ultimo tipo di metadati, comune a tutti i file, consiste negli **attributi** di un file. Alcuni attributi comuni a praticamente qualsiasi file *filesystem* sono le date di **creazione**, ultima **apertura** ed ultima **modifica** dei file, altri elementi utili per gli investigatori, specie per ricostruire una **linea temporale** (*timeline*) di quali file sono stati aperti, e quando.[49]

3.3.6 Osservazioni

L'estrema facilità con cui è possibile modificare i dati digitali deve indurre il giudice a praticare una certa misura di **scetticismo** – che non degeneri nella paranoia! – nei confronti dell'autenticità dei dati digitali. Il ruolo del perito e dei CC.TT. del P.M. e delle parti è di dimostrare che i dati sono **affidabili**, adducendo correlazioni all'interno dei dati stessi. Il ruolo del giudice è di valutare l'affidabilità delle conclusioni tratte dagli operatori tecnici, e di considerarle insieme agli altri elementi di prova non-tecnici raccolti.

[49] *Uno dei modi migliori per presentare i dati è quello di utilizzare una timeline table che, tra le altre cose, consentirà di mostrare la data e l'ora dell'accesso di quel determinato file da parte dell'utente.* Cfr. Giuseppe Vaciago (2011). ≪Digital evidence: profili tecnico-giuridiche e garanzie dell'imputato≫. Tesi di dott. Università degli Studi di Milano-Bicocca. URL: `https://boa.unimib.it/handle/10281/20472`, p. 113.

3.4 Analisi

3.4.1 Metodi di analisi

Per quanto siano fragili e manipolabili, i documenti digitali sono anche **facilmente analizzabili**. Il vantaggio è che esistono moltissimi strumenti per analizzare i dati, e l'analisi può essere fatta anche su enormi quantità di dati.

Per alcuni esempi di analisi di dati digitali:

- **Alterazioni cromatiche** – in Photoshop, si può modificare la tinta di un'immagine e creare una mozzarella blu; ma al tempo stesso, sempre in Photoshop, analizzando la distribuzione del colore dell'immagine, è facile notare un picco innaturale del colore blu, che è chiaro segno di contraffazione;[50]
- **Fotomontaggi** – è possibile combinare insieme più immagini creando un fotomontaggio, che può essere screditato ricercando le immagini originali (*reverse image search*), oppure usando la tecnica dell'*error level analysis*, che permette di verificare se, ed in quali aree, una foto è stata modificata digitalmente;[51]
- **Data carving** – Quando un file viene eliminato, la regione del supporto che lo contiene non viene immediatamente azzerata, ma semplicemente contrassegnata come "libera". Le tecniche di *data carving* permettono di recuperare questi file, o quanto ne rimane, nel caso in cui la regione sia stata parzialmente sovrascritta con altri dati;

[50] Sebastiano Battiato, Giovanni Maria Farinella e Giovanni Puglisi (2011). *Image/video forensics: casi di studio*. URL: `https://www.academia.edu/2867895/Image_Video_Forensics_Casi_di_Studio`.

[51] Cfr. in particolare le pp. 18–20, che mostrano alcuni esempi di immagini fotoritoccate, e come il metodo di analisi mette in evidenza quali zone sono state modificate. Neal Krawetz (2007). *A picture's worth...* URL: `http://hackerfactor.com/papers/bh-usa-07-krawetz-wp.pdf`.

- **Analisi della memoria** – ad es., Volatility permette di analizzare le acquisizioni della memoria volatile di vari sistemi operativi;[52]
- **Analisi di traffico di internet** – ad es., Wireshark può catturare il traffico di internet, o analizzare traffico già catturato, e ricostruire quali richieste e file sono stati scambiati tra due sistemi.[53]
- ***Hash databases*** – gli *hash databases* consistono in una collezione di *hash* di file di un certo tipo (ad es., contenuti pedopornografici). È possibile analizzare i contenuti di un disco, per verificare se l'hash di alcuno dei file contenuti in quel disco corrisponde ad un elemento nel database.[54]

3.4.2 Osservazioni

Così come si deve valutare l'affidabilità dei dati digitali, si deve valutare anche l'**affidabilità degli strumenti e metodi di analisi** che i CC.TT. impiegano. Le operazioni di analisi devono essere **documentate**, in modo che siano riproducibili, e **motivate**, per spiegare lo scopo dell'operazione. Inoltre, è preferibile usare software *open source*, il cui codice sorgente è disponibile. Ciò presenta il vantaggio di documentare anche l'esatto **funzionamento** del programma di analisi utilizzato.

[52] *After a successful memory acquisition process, the investigator will have a single dump file that contains the full memory. [...] Each operating system has a different memory structure [...] volatility can understand the correctdata structures of the image under investigation and apply the right analysis and parsing tools.* Cfr. Shaaban e Sapronov, *Practical Windows Forensics*, pp. 235–236.

[53] *WireShark [...] also plays an invaluable role in cases of network traffic analysis in investigations of incidents.* Cfr. ibid., p. 253.

[54] Michele Ferrazzano (2014). «Indagini forensi in tema di scambio di file pedopornografici mediante software di file sharing a mezzo peer-to-peer». Tesi di dott. Università di Bologna. URL: `http://amsdottorato.unibo.it/6697/`, p. 153.

3.5 Verifica dell'integrità

3.5.1 *Error correcting code*

L'**algoritmo di correzione degli errori** (*error correction code*) (abbreviato in ECC) permette di **individuare e correggere** automaticamente gli errori contenuti in una sequenza di bit, basandosi sui bit "corretti" che rimangono.[55]

3.5.2 Hash

Data una qualsiasi sequenza di bit, è possibile estrarne una "impronta digitale" (*digest*) mediante una **funzione di *hash*** (*hash function*). Il *digest* è a sua volta una sequenza di bit, con le seguenti caratteristiche:[56]

- La **stessa** sequenza in entrata produce **sempre** la stessa sequenza bit in uscita;
- Indipendentemente dalla lunghezza della sequenza in entrata, la sequenza in uscita ha lunghezza fissa;[57]
- Una variazione **piccola** in una coppia di sequenze in **entrata** deve produrre una variazione **grande** nelle coppie in **uscita** ("effetto valanga").[58]

[55] *Computer memories occasionally make errors due to voltage spikes on the power line, cosmic rays, or other causes. To guard against such errors, some memories use error-detecting or error-correcting codes. When these codes are used, extra bits are added to each memory word in a special way. When a word is read out of memory, the extra bits are checked to see if an error has occurred.* Cfr. Tanenbaum e Austin, *Structured computer organization*, p. 78.

[56] Rajeev Sobti e G. Geetha (2012). ≪Cryptographic hash functions: a review≫. In: *International Journal of Computer Science Issues (IJCSI)* 9.2, pp. 461–479, p. 462.

[57] La lunghezza varia da funzione a funzione. Le funzioni MD5 e SHA-1 producono un *hash* di 128 e 160 *bits*, rispettivamente.

[58] *Strong Avalanche effect represents a property when small change in input result in a significant change in message digests. Completeness represents a property when each input bit affects all output bits. Strict Avalanche Criterion combines both the avalanche effect and the completeness and represent a property when a change in one bit of input results in*

Il seguente esempio dimostra le differenze fra i *digest* di due sequenze di bit che differiscono per un solo bit:[59]

ASCII	Binario	*Hash* MD5
B	01000010	9D5ED678FE57BCCA610140957AFAB571
C	01000011	0D61F8370CAD1D412F80B84D143E1257
Bit alterati:	1 (su 8; 12.5%)	66 (su 128; 51.5%)

Prese insieme le proprietà delle funzioni di *hash* sono utili per verificare l'integrità di un file dopo una copia, o nel tempo: se il *digest* è **diverso**, il file è stato **certamente** alterato. Al contrario, se corrisponde, il file **non è stato alterato**.[60]

Tuttavia, può accadere che due sequenze di bit producano lo stesso *digest*, fenomeno detto collisione (*hash collision*).

Le collisioni possono avvenire per **cause naturali**. Le funzioni di *hash* accettano input di qualsiasi lunghezza, ma il *digest* ha una lunghezza limitata. Pertanto, è **sempre** possibile che due sequenze producano lo stesso risultato. Si immagini una funzione di *hash* che produce un *digest* lungo un solo bit: dati **tre** input, ci sarà necessariamente una collisione, perché un singolo bit può rappresentare solo **due** valori.

Con la funzione di hash CRC-32, le sequenze di caratteri `plumless` e `buckeroo` producono entrambe il digest `4DDB0C25`.[61]

changing every bit of the output (message digest) with a probability of [50%]. Sobti e Geetha, ≪Cryptographic hash functions: a review≫, p. 466.

[59] Valori calcolati con `https://passwordsgenerator.net/md5-hash-generator/`. La variazione di un solo bit nella sequenza di input comporta la variazione del 51.5% nel *digest*, valore molto vicino alla media del 50% prevista dall'*avalanche effect*.

[60] *The usage of Hash Functions for Message Authentications and ensuring message integrity has surged because majority of hash functions are faster than block ciphers in software implementation and these software implementations are readily and freely available.* Sobti e Geetha, ≪Cryptographic hash functions: a review≫, p. 463.

[61] La collisione è verificabile usando `https://crccalc.com/`. I valori sono citati in Jeff Preshing (2011). *Hash Collision Probabilities*. URL: `https://preshing.com/20110504/hash-collision-probabilities/`.

Tuttavia, **in pratica**, le funzioni di *hash* producono *digest* sufficientemente lunghi da evitare collisioni, anche con un numero molto più elevato di input. Ad es., SHA-1 produce un *hash* di 160 bit. Dato che ogni bit può assumere due valori, ci sono 2^{160} ($= 1.56 \times 10^{48}$) possibili *digest*.

La probabilità che **due** sequenze prese individualmente producano lo stesso hash è assolutamente trascurabile: anche con un numero di valori nell'ordine di grandezza di 10^{15}, con una buona funzione di hash, la probabilità di una collisione è 1 su 10^{18}.[62]

Altrimenti, le collisioni possono essere il prodotto di un **attacco intenzionale**. È possibile trovare due valori in input che producono lo stesso output. La prima collisione intenzionale per la funzione MD5 è stata scoperta nel 2005, e volendo, è possibile generare collisioni MD5 anche partendo da un file qualsiasi.[63]

Il secondo caso ha una maggiore rilevanza pratica, ma il problema delle collisioni può essere aggirato usando **due o più** funzioni di hash. Anche se un algoritmo produce una collisione, l'altro produrrà valori diversi. Attualmente si usano MD5 e SHA-1, ma sarebbe opportuno iniziare a considerare funzioni di hash più moderne.[64]

[62] Preshing, *Hash Collision Probabilities*.

[63] Per una dimostrazione, cfr. Nat McHugh (2015). *Create your own MD5 collisions*. URL: `https://natmchugh.blogspot.com/2015/02/create-your-own-md5-collisions.html`.

[64] *Last week, the Scientific Working Group on Digital Evidence published a draft document [...] where it accepts the use of MD5 and SHA-1 in digital forensics applications: "While SWGDE promotes the adoption of SHA2 and SHA3 by vendors and practitioners, the MD5 and SHA1 algorithms remain acceptable for integrity verification and file identification applications in digital forensics. [...]" This is technically correct [...] Still, it's really bad form to accept these algorithms for any purpose.* Cfr. Bruce Schneier (2018). *MD5 and SHA-1 Still Used in 2018*. URL: `https://www.schneier.com/blog/archives/2018/12/md5_and_sha-1_s.html`.

3.5.3 Osservazioni

Garantire l'integrità della copia nel tempo è fondamentale. Oltre alle tecnologie **standard** e di largo uso (ad es., MD5 e SHA-1), è utile iniziare ad adottare anche **standard moderni**, tanto per ragioni di *future-proofing*,[65] che per evitare di mettere "tutte le uova in un paniere".[66]

Nel caso in cui gli hash non siano uguali, è necessario fare una operazione di confronto fra l'originale e la copia, per controllare esattamente dove differiscono. Se i dati che sono oggetto di accertamento risultano alterati, la copia va distrutta, ed eventuali analisi compiute sulla copia "corrotta" vanno ripetute. Se la corruzione non riguarda i dati oggetto dell'indagine, è possibile, ma fortemente sconsigliato – salvo i casi di estrema urgenza, o mancanza di altre copie di backup – di continuare ad usare quella copia.

3.6 Riproducibilità

3.6.1 Caratteristiche delle copie

Nonostante la loro intrinseca fragilità, i dati digitali possono essere riprodotti in maniera **perfetta, illimitata**, e su **qualsiasi tipo di supporto**.

La ragione è semplice da dimostrare: i dati digitali consistono solo ed unicamente in una sequenza di bit, quindi **replicando** la sequenza di bit, si ottiene una **copia perfetta ed indistinguibile** dell'originale. Invece, con i "dati" analogici è pressoché impossibile ottenere una copia perfetta: ogni esemplare sarà sempre leggermente diverso.

[65] Eventualmente gli standard attuali saranno considerati non più sicuri, e quindi deprecati. Anticipare i tempi è utile, così che nel domani in cui le indicazioni cambieranno, si era già pronti ieri.

[66] Come già indicato, più funzioni di hash si usano, e più il risultato è affidabile, perché si riduce il rischio di collisioni.

Astrattamente, è possibile creare un numero **illimitato** di **copie perfette**, ma concretamente, esistono dei limiti materiali:

- **Spazio di archiviazione** – lo spazio di archiviazione è limitato dalle possibilità economiche. Anche se il costo delle memorie di archiviazione è diminuito considerevolmente, conservare grandi quantità di dati non è economico.
- **Errori** – durante l'operazione di copia, possono avvenire degli errori. Ad es., se un masterizzatore viene colpito mentre sta leggendo o scrivendo un disco, si può subire una perdita di dati; se la connessione viene persa durante la trasmissione di un file, può risultare incompleto; ecc. . .

Infine, le sequenze di bit possono essere copiate su **qualsiasi tipo** di supporto. Come già visto, anche se le caratteristiche a livello hardware sono diverse, il firmware produce un'astrazione, per cui ogni supporto può essere visto come una semplice sequenza di celle.

3.6.2 Osservazioni

L'operazione di riproduzione è il **momento più critico** per l'acquisizione di dati digitali. È necessario **calcolare l'hash del supporto prima e dopo** dell'operazione di copia, per verificare che non siano avvenuti errori. Devono essere presi accorgimenti affinché il disco da copiare sia in **modalità sola lettura**, per evitare di modificare i suoi contenuti. È preferibile usare soluzioni sia a livello **software** che a livello **hardware**. **Più precauzioni** vengono prese, e più i dati sono affidabili.

La riproduzione è utile anche per contrastare il deterioramento dei supporti materiali, mediante la creazione di copie di backup: nel caso di deteriora-

mento anche catastrofico di una copia, esistono altre copie.[67] È opportuno monitorare l'integrità delle copie di backup sia dopo la loro acquisizione, che nel tempo, per evitare di perdere copie del supporto originale.

[67] La *rule of three* è una regola indicativa su quanti backup avere e come gestirli: si devono avere almeno tre copie, su due tipi di supporti diversi, e almeno una delle tre copie deve essere conservata fuori sede. Cfr. Frazer Lloyd-Davies (2018). *The backup rule of three is a simple way to remember backup best practice.* URL: `https://www.acronyms.co.uk/blog/backup-rule-of-three/`.

Capitolo 4

Digital evidence e documento informatico

4.1 Definizione di *digital evidence*

La *digital evidence* è il nome dato alle **fonti di prova di natura digitale**. Ne esistono varie definizioni, ognuna delle quali ne cattura alcune caratteristiche.

In primo luogo, devono essere dati o informazioni digitali che siano **rilevanti** per una **investigazione**:

> Information or data, stored or transmitted in binary form which has been determined, through the process of analysis, to be relevant to the investigation.[1]

Può sembrare scontato, ma è bene precisare che la rilevanza è definita come la capacità di **provare fatti** relativi alla **responsabilità penale**:

[1] Sezione 3.5 in International Organization for Standardization (ISO) (2015). *Information technology — Security techniques — Guidelines for the analysis and interpretation of digital evidence*. URL: `https://www.iso.org/standard/44406.html`.

> Any data stored or transmitted using a computer that support or refute a theory of how an offence occurred or that address critical elements of the offence such as intent or alibi.[2]

In altre parole, gli atti processuali in forma digitale, anche se "rilevanti" per l'investigazione, non fanno parte della *digital evidence*, perché non hanno valore probatorio.[3]

È possibile individuare due tipi di *digital evidence*:[4]

- ***Computer-derivated*** – dati digitali che costituiscono l'**oggetto** delle investigazioni, e quindi sono **estratti** dal computer. L'atto di indagine corrispondente è il **rilievo**;[5]
- ***Computer-generated*** – dati digitali che sono creati **in funzione** delle investigazioni mediante l'uso di strumenti digitali. A seconda del tipo di attività svolte dall'operatore, possono consistere sia in rilievi[6] che in **accertamenti**.[7]

[2] Casey, cit. in Severino Murgia (n.d.). ≪Prova informatica e processo penale≫. Tesi di dott. Università degli Studi di Pavia. URL: https://iris.unipv.it/retrieve/handle/11571/1203330/184892/Severino%20Murgia%20-%20Prova%20informatica%20e%20processo%20penale.pdf, p. 57.

[3] *Il pericolo principale che deve essere scongiurato è quello di una imprudente confusione tra prova documentale e atto processuale.* Cfr. ibid., p. 59.

[4] Luparia, cit. in ibid., p. 60.

[5] *Si intende con "rilievo" quell'atto avente natura meccanica in cui l'apporto dell'uomo non richiede particolari competenze [...].* Gli operatori non hanno un margine discrezionale nello svolgimento delle operazioni, dato che esistono strumenti e procedure standardizzate (*best practices*). Per la defizione di "rilievo", cfr. Michele Pennisi (2015). ≪Rilievi ed accertamenti di polizia giudiziaria. I problemi esegetici posti dalla normativa vigente e gli sviluppi dottrinali e giurisprudenziali in materia≫. Tesi di laurea mag. Università di Bologna. URL: http://www.giurisprudenzapenale.com/wp-content/uploads/2017/09/Tesi-Pennisi.pdf, p. 36.

[6] Si pensi al caso di fotografie e videoriprese effettuate durante una ispezione con fotocamere e videocamere digitali: non c'è un margine discrezionale di apprezzamento da parte dell'operatore.

[7] *[Si intende] con "accertamento" [quell'atto] in cui l'esperto riversa un apporto critico-valutativo.* Qui rientra tutta l'attività svolta dai consulenti, che fa uso di computer per analizzare i dati raccolti. L'elemento soggettivo consiste nella scelta dei mezzi tecnici, nelle attività svolte, e nella argomentazione e conclusioni che vengono rese nella loro relazione.

La *digital evidence* è strettamente legata alla disciplina della ***digital forensics***, che può essere vista da due punti di vista:

- **Sostanziale** – come "l'insieme di attività finalizzate alla risoluzione dei casi connessi alla criminalità informatica";[8]
- **Processuale** – come lo studio del "valore che un certo dato informatico può avere in ambito giuridico [...] [e] la capacità di resistenza di un dato informatico alle contestazioni delle altre parti processuali".[9]

4.2 Documento informatico

Esistono più definizioni di documento informatico.[10] Tuttavia, ai fini della presente trattazione, interessa analizzare la nozione di "documento" ai soli fini del procedimento penale. L'art. 234 co. 1 c.p.p. recita:

> È consentita l'acquisizione di scritti o di altri documenti che rappresentano fatti, persone o cose mediante la fotografia, la cinematografia, la fonografia o qualsiasi altro mezzo.

Ad un primo sguardo, la definizione non esclude la possibilità di usare dati digitali.

- "Di scritti o di altri documenti" – non sono ulteriormente qualificati con espressioni che precludono l'uso di documenti non materiali, che in ogni casso possono bene rientrare nella categoria "altri documenti". Con spicco tecnocratico, si potrebbe anche argomentare che le due

Cfr. Pennisi, ≪Rilievi ed accertamenti di polizia giudiziaria. I problemi esegetici posti dalla normativa vigente e gli sviluppi dottrinali e giurisprudenziali in materia≫, p. 36.

[8]Maioli, cit. in Murgia, ≪Prova informatica e processo penale≫, p. 58.

[9]Ziccardi, cit. in ibid., p. 58.

[10]Per una rassegna, cfr. Gammarota, ≪Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali≫, pp. 175–177.

espressioni riproducono la distinzione fra file di testo ("scritti") e file binari ("altri documenti");

- "La fotografia, la cinematografia, la fonografia" – esistono anche in versione digitale;
- "Qualsiasi altro mezzo" – può essere inteso come una allusione tanto alla ***digital evidence* in sé** (*computer-derivated*), quanto alle possibilità di rappresentazione di "fatti, persone o cose" che sono **rese possibili** dai mezzi digitali (*computer-generated*) – ad es., si pensi alla fotogrammetria[11] e modellazione 3D, oppure alle simulazioni digitali di fenomeni reali[12]).

Pertanto, si può affermare che la *digital evidence* può essere vista come una *species* del *genus* "documento", più precisamente, un "documento informatico".

4.3 Confronto con il paradigma tradizionale

L'ipotesi è confermata se si confronta il paradigma dei documenti tradizionali con le caratteristiche offerte dai documenti digitali.[13] Anticipando le conclusioni, pur con qualche adattamento dovuto alla loro particolare

[11] La fotogrammetria consiste nel compiere misurazioni di oggetti utilizzando solo delle fotografie come riferimento. È utile per ricostruire cose, persone e luoghi in versione digitale. Esistono anche applicazioni per smartphone che consentono l'acquisizione della forma di un oggetto semplicemente riprendendolo da più angoli. Per un'introduzione, cfr. Crompwell (2018). *Intro To Photogrammetry*. URL: `https://youtu.be/3EENC9rFWhc`.

[12] Un esempio di *computer-generated evidence* consiste nell'esperimento giudiziale "virtuale", che permette di creare una versione virtuale di "fenomeni [...] che sarebbe impossibile replicare nella realtà, come ad esempio la rovina del fianco di una montagna in un bacino idrico delimitato da una diga". Altri vantaggi includono la possibilità di ripetere l'esperimento un numero potenzialmente illimitato di volte, e di poter impostare i parametri della simulazione in maniera precisa. Cfr. Gammarota, «Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali», p. 197.

[13] Più precisamente, le caratteristiche indicate da Tonini, cit. in ibid., p. 178.

natura dematerializzata, i dati digitali sono pienamente capaci di integrare le varie caratteristiche dei documenti tradizionali.

4.4 Fatto rappresentato

4.4.1 Fatti rappresentabili

Il "fatto rappresentato" consiste in:

> Atti, persone, cose, indicate dall'art. 234, ma anche pensieri, ovvero tutto ciò che può costituire oggetto di prova, ovvero un accadimento naturalistico o un atto umano, quale una dichiarazione.[14]

I **fatti rappresentabili dai documenti digitali** possono essere divisi in due tipi:

- **Fatti nativamente digitali** – fatti che consistevano in dati digitali fin dal momento della loro creazione;
- **Fatti digitalizzati** – fatti inizialmente "materiali", poi convertiti in dati digitali in un secondo momento (processo detto **digitalizzazione**),[15] come ad esempio la scansione di un foglio di carta in un documento PDF.

4.4.2 Dominio digitale e dominio materiale

La distinzione può sembrare artificiosa, ma è fondata sulla diversa natura dei due tipi di fatti.

[14] Gammarota, ≪Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali≫, p. 178.

[15] *Si definisce digitalizzazione il processo di trasformazione di un dato analogico in dato digitale, ovvero la sua rappresentazione mediante la codifica binaria.* Cfr. ibid., p. 48.

Finché **si rimane nel dominio digitale**, le operazioni sono (almeno tendenzialmente) **senza perdita di informazioni** (*lossless*). Il fatto digitale – sia nativo che digitalizzato – **consiste interamente** in una **sequenza di bit**. Pertanto, la sua **acquisizione** consiste nella semplice duplicazione dei dati digitali di cui è composto, e crea un **clone perfetto**.

D'altro canto, ogni volta che **si passa** dal dominio digitale a quello materiale, o viceversa, è **impossibile evitare** una certa quantità di **degradazione** (*lossy*). Al più, la copia digitale rappresenta una **mera approssimazione** del fenomeno materiale.

Così come il dipinto di una pipa non è una pipa, la **stampa** di una e-mail **non è** una **e-mail**, perché la prima è un foglio di carta, mentre la seconda è una sequenza di bit. Viceversa, la scansione in PDF di un foglio di carta non è un foglio di carta: il primo è composto di pixel, e ha una risoluzione limitata, mentre il secondo è composto di fibre ed inchiostro, e non ha alcuna nozione di "risoluzione" e "pixel", ma al più di "dimensioni" e precisione dello strumento usato per la stampa.

Colloquialmente, si può dire che il dipinto di una pipa sia una "pipa", ma da un punto di vista **ontologico** ("cos'è") e soprattutto **epistemologico** ("che informazioni ci può dare"), c'è una **differenza incolmabile**: la pipa-dipinto non può essere usata per fumare tabacco, né si possono svolgere sulla pipa-dipinto le analisi che si potrebbero svolgere sulla pipa-oggetto.

Analogamente, la stampa di una e-mail comporta la serie di una perdita di informazioni. Viene riprodotta solo l'**apparenza** del documento digitale, ma non la sua **essenza**, ed è solo la seconda che può essere oggetto di analisi tecniche. È estremamente facile **falsificare** l'apparenza di un documento digitale, e la rilevazione di modifiche può essere fatta solo analizzando la **sequenza di bit** di cui è composto.

Ad es., chiunque può usare lo strumento "ispeziona elemento" per modifi-

care i contenuti di una pagina web, e poi stampare la pagina, o fare uno screenshot. La maniera corretta di acquisire una pagina web richiede la cattura dei dati di cui è composta, così come vengono inviati dal server: quella è la vera "essenza" della pagina web. Quanto appare a schermo è solo un'apparenza.[16]

Ogni digitalizzazione o "materializzazione" consiste nella **perdita di caratteristiche, dati e metadati** che sono inerenti alla natura digitale o materiale di un supporto, e non possono essere accuratamente rappresentati sull'altro tipo di supporto. Purtroppo c'è una tendenza ad accontentarsi delle mere "apparenze", o per accorciare i tempi, o per la mancata conoscenza delle problematiche, e si scoraggia l'uso di analisi tecniche che vadano a verificare l'"essenza".[17]

La stampa di una e-mail non è una e-mail, è solo un pezzo di carta. Già è difficile determinare se una e-mail è stata alterata, anche analizzando la sua rappresentazione binaria, ma se si aggiunge l'ulteriore ostacolo di avere solo una rappresentazione, una mera "apparenza" dei dati, e non la loro "essenza", allora quell'elemento è assolutamente inaffidabile.[18]

4.4.3 Adeguatezza e completezza della documentazione

Tuttavia, non sempre è possibile rappresentare i dati digitali in maniera perfetta. In alcuni casi è possibile solo ottenere una **rappresentazione adeguata**, che sia il più vicina possibile ai dati originali.

[16] Per una guida che dimostra la semplicità di creare falsi, cfr. Jack Busch (2011). *Friday Fun: Use Chrome to Create Fake Screen Captures*. URL: `https://www.groovypost.com/howto/howto/friday-fun-chrome-create-fake-screen-captures/`.

[17] *Si può quindi affermare che la giurisprudenza291 è generalmente incline "a limitarsi" ad un esame "sensoriale" della rappresentazione dei mezzi di prova documentale a contenuto informatico [...].* Cfr. Gammarota, ≪Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali≫, p. 184.

[18] Per una serie di esempi, cfr. ibid., pp. 187–188.

Un primo caso riguarda risorse che non sono nella immediata materiale disponibilità degli investigatori, come siti internet, o file conservati su servizi cloud. L'"essenza" di quei dati si trova su un server remoto, che può facilmente trovarsi fuori dal territorio nazionale.

Anche mettendo da parte le problematiche giuridiche, in ogni caso è difficile ottenere una copia esatta dei dati per ragioni tecniche. Un caso comune è l'impiego di crittografia.[19]

In questi casi è necessario affidarsi a quanto è possibile rinvenire sul computer, o visionare mediante i programmi installati.

Un altro caso riguarda l'uso di *thumbnails* (miniature). Un *thumbnail* consiste in una versione a risoluzione ridotta di un'immagine. Sono spesso generate come anteprime per i file in una cartella. Consistono in una mera apparenza del file, non nella sua essenza. Ma in almeno un caso le *thumbnail* sono state considerate come equivalenti all'immagine originale, che nel frattempo era stata cancellata.[20]

Un punto a favore dell'affidabilità *thumbnails* è che sono generate in maniera automatica, e quindi – fatta salva l'inevitabile perdita di qualità dovuta alla risoluzione ridotta – rappresentano lo stesso fatto dell'immagine originale.

[19] Ad es., Protonmail.com protegge i dati degli utenti con sistemi crittografici, e avvisa gli utenti che se perdono la password, diventa impossibile recuperare i dati, anche per il fornitore dei servizi. *Messages are stored on ProtonMail servers in encrypted format. They are also transmitted in encrypted format [...] ProtonMail's zero access architecture means that your data is encrypted in a way that makes it inaccessible to us. [...] If you forget your password, we cannot recover your data.* Cfr. `https://protonmail.com/security-details`.

[20] *Detective Timothy Luckie testified to the results of the full forensic analysis of Romm's hard drive. Luckie confirmed that all of the child pornography on Romm's computer had been deleted. The vast majority of the images Luckie found had been deleted from Romm's internet cache. [...] Luckie's analysis also showed that Romm had enlarged a few smaller "thumbnail" images in the internet cache. [...] In short, given the indicia that Romm exercised control over the images in his cache, there was sufficient evidence for the jury to find that Romm committed the act of knowing possession.* Cfr. United States Court of Appeals, Ninth Circuit (2006). *United States v. Romm*. URL: `https://caselaw.findlaw.com/us-9th-circuit/1231820.html`.

Pertanto, in alcuni casi è necessario accontentarsi delle "apparenze" che è possibile visionare dal sistema. **Laddove è possibile**, è sempre preferibile procedere ad una acquisizione dei **bit-essenza**; ma **in ogni caso**, per esigenze di completezza, è sempre e comunque utile catturare anche l'**apparenza**.

Le due operazioni sono complementari. L'**acquisizione** dei bit-essenza cattura i "**dati**", e cerca di avvicinarsi con meno mediazioni possibile alla **esatta** sequenza di *bit* che descrivono **quel fatto** – quasi una "verità digitale", parallela alla "realtà storica".

La documentazione delle **apparenza** ha la funzione dei "**metadati**", perché fornisce informazioni utili riguardo alle **modalità seguite** per l'acquisizione, e documenta tutte le **circostanze anomale** – violazioni delle *best practices*, malfunzionamenti dell'apparecchiatura, ecc... – avvenute durante le operazioni. Deve essere eseguita con modalità che rappresentino il fatto in **maniera adeguata** – ad es., la videoripresa della ispezione di un supporto, o di un sito web, sia di quanto accade sullo **schermo**, che dell'**operatore** che sta eseguendo le operazioni.

4.5 Rappresentazione

La rappresentazione è:

> La ricostruzione di un fatto mediante un equivalente fatto di immagini, parole, suoni, in modo tale da renderlo conoscibile ad altre persone.[21]

La rappresentazione digitale avviene sempre è in formato binario. Esistono una serie di regole che determinano **come** quel codice binario deve essere interpretato per produrre un certo risultato.

[21] Gammarota, ≪Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali≫, p. 178.

Esistono numerosi formati per svariati tipi di rappresentazioni. A titolo solo illustrativo, "la fotografia, la cinematografia, la fonografia"[22] nel mondo digitale corrispondono a:

- **Immagini** – JPG, PNG e GIF sono formati con supporto pressoché universale. Spesso gli editor di immagini usano formati proprietari, come PSD (Photoshop).
- **Video** – MP4,[23] MKV e AVI sono i formati più diffusi.
- **Audio** – MP3 e AAC sono utilizzati per le registrazioni vocali.

Anche i **supporti materiali** possono essere rappresentati in un **documento digitale**, questo tipo di file è chiamato **immagine** di un disco (*disk image*). Ai fini dell'informatica forense, esistono tre formati di immagine:

- Formato **RAW** (o DD o IMG)[24] consiste **esclusivamente** nel **contenuto** del supporto, copiato bit per bit, mentre eventuali metadati riguardanti il supporto e l'operazione sono conservati in un file a parte.
- Formato **EWF** – presenta varie funzionalità, come l'aggiunta di metadati relativi al caso, la compressione dei contenuti del disco, la verifica dell'integrità dei contenuti dell'immagine.[25]

  Esistono anche altri tipi di formati che supportano i metadati, ma l'EWF è l'*industry standard*, ed il più largamente supportato.[26]

[22] Cfr. art. 234 co. 1 c.p.p.

[23] Apple usa il formato M4V, sostanzialmente identico all'MP4, ma che può contenere anche meccanismi per la protezione digitale dei diritti (*DRM*).

[24] ForensicsWiki contributors (n.d.[b]). *Raw Image Format*. URL: `https://forensicswiki.xyz/wiki/index.php?title=Raw_Image_Format`.

[25] Sally Vandeven (2014). *Forensic Images: For Your Viewing Pleasure*. URL: `https://www.sans.org/reading-room/whitepapers/forensics/paper/35447`.

[26] Nonostante il formato EWF sia usato da un software commerciale – EnCase – esiste una libreria *open source* – "libewf" – che semplifica l'adozione del formato da parte di altri programmi. Cfr. `https://github.com/libyal/libewf`.

- **Dischi virtuali** – sono il formato di disco utilizzato dalle macchine virtuali. Le immagini forensi (i primi due tipi) possono essere convertite in un disco virtuale affinché possano essere avviate in una macchina virtuale.[27]

4.6 Incorporamento

L'incorporamento è:

> L'operazione con la quale la rappresentazione viene fissata su una base materiale.[28]

L'incorporamento informatico ha una caratteristica particolare, perché può creare documenti dematerializzati, ed è possibile memorizzare ed accedere ai dati digitali "in maniera indipendente dal supporto sul quale [sono contenuti]".[29]

Si può anche affermare che l'operazione di incorporamento – specie sulle memorie secondarie – è il momento in cui il dato digitale assume una maggiore materialità, perché diventa meno volatile.

Concretamente, le operazioni di incorporamento dipendono dal tipo di base materiale – o meglio, hardware – su cui si vogliono memorizzare i dati. Come già accennato, esistono tre grandi famiglie di memorie, quelle ottiche, magnetiche ed elettriche. Ognuna presenta una modalità di memorizzazione dei dati a sé.[30]

Le operazioni di lettura e scrittura richiedono la cooperazione di vari livelli:

[27] Per una guida, cfr. PassMark Software (n.d.). *Booting a forensics image on a Virtual Machine*. URL: `https://www.osforensics.com/faqs-and-tutorials/booting-image-virtual-machine.html`.

[28] Gammarota, ≪Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali≫, p. 178.

[29] Murgia, ≪Prova informatica e processo penale≫, p. 69.

[30] Cfr. sezione 2.4.1.

- **Utente** – l'utente è il "motore immobile" aristotelico, la prima causa di ogni modifica, perché ogni operazione digitale – anche nel caso di operazioni automatiche – è in ultima analisi riconducibile ad una operazione umana, con cui il sistema è stato avviato o programmato;
- **Applicazioni** – le applicazioni mostrano all'utente direttamente informazioni già decodificate, e permettono di manipolare i dati memorizzati;
- **Sistema operativo** – tutte le operazioni di lettura da/scrittura verso memoria richieste dalle applicazioni sono gestite dal sistema operativo, che presenta un sistema standardizzato, ed astrae i dettagli specifici;
- **Driver** – più precisamente, il sistema operativo fa riferimento alle istruzioni contenute nei driver per comunicare correttamente con l'hardware, ed inviare i comandi corretti;[31]
- **Firmware** – il firmware è a diretto contatto con l'hardware. Riceve i comandi inviati dal driver, e li mette in atto sull'hardware. [32]
- **Hardware** – l'hardware è l'ultimo livello, ed è il componente "materiale" che viene effettivamente modificato, dal punto di vista materiale.

Si può affermare che driver e firmware rappresentino il **ponte** che collega il mondo software (dematerializzato), ed il mondo hardware (materiale),

[31] *In the most fundamental sense, a driver is a software component that lets the operating system and a device communicate with each other. For example, suppose an application needs to read some data from a device. The application calls a function implemented by the operating system, and the operating system calls a function implemented by the driver. The driver, which was written by the same company that designed and manufactured the device, knows how to communicate with the device hardware to get the data. After the driver gets the data from the device, it returns the data to the operating system, which returns it to the application.* Cfr. Microsoft Docs Contributors (2017). *What is a driver?* URL: `https://docs.microsoft.com/en-us/windows-hardware/drivers/gettingstarted/what-is-a-driver-`.

[32] Per maggori informazioni sul firmware, cfr. il paragrafo intitolato "*Purge*" nella sezione 3.3.2

e che siano il vero cuore dell'operazione di incorporamento, perché definiscono in maniera esatta – in particolare il firmware – come i dati digitali devono essere rappresentati sulla base materiale.

4.7 Base materiale

Così come i documenti, anche i bit **richiedono una base materiale** su cui essere fissati. È possibile affermare che i documenti informatici, pur non avendo corporalità propria, **non sono immateriali**, ma bensì **dematerializzati**, perché "esistono" all'interno di un supporto fisico, e come le procedure di *data sanitization* dimostrano, intervenendo sul fisico il dato è irrimediabilmente perso.

Tuttavia, nel caso di dati digitali caratterizzati dalla **volatilità** – memorie primarie e bit in trasmissione – è più difficile parlare di materialità, dato che non c'è una memorizzazione stabile: nel primo caso, la memorizzazione è temporanea e volatile, mentre nel secondo caso non si può nemmeno parlare di memorizzazione, ma solo di trasporto. Pertanto, in questi casi è possibile affermare che i bit sono "**più immateriali che dematerializzati**", perché richiedono comunque un supporto per essere trasmessi, ma la loro esistenza nel mondo materiale è estremamente **effimera**.[33]

[33] *I bit teletrasmessi, a prescindere dalla tecnologia e dalla tecnica utilizzata, sono disgunti [sic] da qualunque supporto materiale e quindi sono immateriali.* Cfr. Gammarota, ≪Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali≫, p. 56.

Capitolo 5

Adattamento dei mezzi di ricerca della prova alla *digital evidence*

5.1 Introduzione

La Convenzione di Budapest è stata emanata dal Consiglio Europeo nel 2001, con l'intento di "di permettere una collaborazione ed una legislazione coordinate fra gli Stati membri in grado di affrontare in maniera incisiva la criminalità informatica".[1]

In Italia, è stata recepita con la L. n. 48/2008, che ha introdotto una serie di principi per avvicinare il c.p.p. alle *best practices* della comunità scientifica, ed evitare l'uso (e abuso) di "pratiche lassiste" nell'acquisizione di dati da dispositivi informatici.[2]

I due principi-cardine, contenuti agli articoli 8 e 9, sono:

[1] Maria Concetta De Vivo e Giovanna Ricci (2012). ≪Diritto, crimini e tecnologie≫. In: *Informatica e diritto* 21.2, pp. 27–114. URL: http://www.ittig.cnr.it/risorse/attivita-editoriale/rivista/indici/2-2012/, p. 92.

[2] Torre, ≪Aspetti giuridici e tecnici relativi al trattamento della prova digitale nel processo penale. La prova informatica nella legge 18 marzo 2008, n. 48≫, pp. 68–69.

- **Conservare i dati originali** – le procedure di acquisizione non devono modificare i dati originali, per garantire "la possibilità, per le parti processuali, di riferirsi e di confrontarsi con i dati originali" anche a distanza di tempo, anche nel caso di analisi future.[3]

- **Creare copie conformi all'originale** – la copia deve essere, e deve rimanere, conforme all'originale. Se non riproduce fedelmente i contenuti dell'originale, **bit per bit**, la sua intera ragion d'essere viene meno. A dimostrazione dell'importanza di una duplicazione senza errori, per passare da "2020" a "2028" è sufficiente cambiare **un solo** bit, il terzultimo, da `...0110000` a `...011100`.

Significativa è la menzione di dati informatici nell'articolo rubricato "Cose deperibili.", al seguito di cose "che possono alterarsi" o "di difficile custodia":

> L'autorità giudiziaria fa estrarre copia dei documenti e fa eseguire fotografie o altre riproduzioni delle cose sequestrate che possono alterarsi o che sono di difficile custodia, le unisce agli atti e fa custodire in cancelleria o segreteria gli originali dei documenti, disponendo, quanto alle cose, in conformità dell'articolo 259. Quando si tratta di dati, di informazioni o di programmi informatici, la copia deve essere realizzata su adeguati supporti, mediante procedura che assicuri la conformità della copia all'originale e la sua immodificabilità;[4]

[3] Ziccardi, cit. in Tonnellotto, ≪Evidenza informatica, computer forensics e best practices≫, p. 77.

[4] Cfr. art. 260 co. 2 c.p.p.

5.2 Sanzione per la violazione delle *best practices*

Mentre il legislatore prende pieno atto della fragilità dei dati digitali, ed indica il risultato da perseguire nella loro gestione,[5] allo stesso tempo non prevede alcuna sanzione nel caso in cui le operazioni eseguite non rispettino i criteri richiesti dal codice.

La dottrina è divisa in base a come si debba procedere:[6]

- Alcuni ritengono che la violazione incida solo sulla **valutazione della prova**, che rimane pertanto utilizzabile, ma poco attendibile. Più è grave la violazione, più è compromessa l'affidabilità;
- Altri preferiscono la soluzione più radicale della **non ammissibilità**,[7] e sostengono che la violazione delle garanzie risulta in un "materiale inquinato capace di adulterare la ricostruzione penale",[8] che deve essere estromesso senza mezze misure.

Si potrebbe argomentare che la seconda soluzione sia più coerente con l'intento del legislatore. Anche se la legge non lo menziona, è evidente che dati digitali male acquisiti siano intrinsecamente inaffidabili, e quindi inutili. Basta alterare un solo bit per cambiare una data, quindi ogni operazione che non sia eseguita con tutti i dovuti accorgimenti va considerata come inammissibile.

[5] Il codice non specifica le modalità concrete, ma l'omissione è ragionevole, dato che la disciplina è in continua evoluzione. È pertanto intuibile un riferimento implicito alle ***best practices***, come standard, linee guida, e la più autorevole dottrina. Un possibile elemento di dibattito in giudizio, e valutazione da parte del giudice, consiste proprio nell'affidabilità delle modalità di acquisizione che sono state usate in concreto.

[6] De Vivo e Ricci, ≪Diritto, crimini e tecnologie≫, pp. 88–89.

[7] Tonnellotto, ≪Evidenza informatica, computer forensics e best practices≫, p. 79.

[8] Cit. di Luparia in De Vivo e Ricci, ≪Diritto, crimini e tecnologie≫, p. 89.

Tuttavia, la prima soluzione ha il pregio di essere più flessibile. Tutte le prove vengono raccolte, anche se le *best practices* non sono state pienamente rispettate, e non ci si deve preoccupare di avere escluso prove che invece potrebbero risultare comunque utili. Fatto salvo il caso di errori macroscopici,[9] valutare l'affidabilità della *digital evidence* a priori non è facile.[10]

I nodi verranno al pettine durante il dibattimento. Sarà compito dei CC.TT. di parte e del perito di fornire al giudice gli elementi per la valutazione dell'affidabilità della prova, e fornire il proprio parere tecnico.[11] La *digital evidence* è solo una parte del quadro probatorio, ed il giudice è libero di valutare se ignorare completamente la prova, o darle un peso ridotto – motivando adeguatamente la sua decisione.

In particolare, i CC.TT. devono stimolare il giudice verso un giudizio critico, contestando le modalità di assunzione usate dalla controparte, ed enfatizzando le proprie. Il rischio maggiore è che il giudice si affidi ciecamente alla prova scientifica, sopravvalutandone la capacità rappresentativa, proprio perché "scientifica".

5.3 Affidabilità dei mezzi di acquisizione

Una questione interessante riguarda l'affidabilità dei mezzi di acquisizione. L'operatore può utilizzare o software commerciali, o software *open source*.

[9] Ad es., un computer spento viene acceso, la copia di singoli file avviene con il loro trascinamento da una cartella ad un'altra, non si prende l'hash della copia... – tutte operazioni tali da compromettere con sicurezza, "oltre il ragionevole dubbio", l'affidabilità della prova.

[10] La valutazione preliminare dell'ammissibilità è difficile: ad esempio, una catena di custodia sommaria fino a che punto rende implicitamente inaffidabile la prova? Così come è difficile valutare ad anteriori cosa acquisire sulla scena del crimine, e si prende tutto, è meglio ammettere tutto, e valutare dopo

[11] Raffaella Brighi e Cesare Maioli (2015). ≪Un cambio di paradigma nelle scienze forensi. Dall'armonizzazione tecnico-giuridica a una nuova cornice epistemologica≫. In: *Informatica e diritto* 24.1-2, pp. 217–234. URL: http://www.ittig.cnr.it/risorse/attivita-editoriale/rivista/indici/1-2-2015/, p. 233.

La seconda scelta è preferibile, perché si può controllare l'esatto funzionamento del programma andando a leggere il codice sorgente. Nel primo caso, dato che i sorgenti non sono disponibili, si deve necessariamente fare affidamento sull'integrità del software.[12]

È possibile utilizzare software commerciali senza comprare licenza: le "crack" sono modifiche al programma che rimuovono o modificano i componenti del programma che verificano lo stato della registrazione.

Tuttavia, sono strumenti illeciti, non supportati dallo sviluppatore del software, e che possono potenzialmente introdurre comportamenti imprevedibili:

> Crack tools are detected as malware or viruses because, by definition, they are. Their specific purpose is to modify programs and files so that they don't work as designed. They delete verification files, modify registration status and do whatever they can to make their target not work as intended.[13]

Tuttavia, la Cassazione ritiene che una prova acquisita con software che ha subito modifiche illegittime ("non licenziato") non abbia "alcun riflesso sotto il profilo dell'illegittimità della prova stessa".[14]

È una posizione non giustificabile: una prova acquisita con software non licenziato deve essere assolutamente considerata come inaffidabile *in res ipsa*, perché comporta affidamento a modifiche introdotte da terze parti rispetto allo sviluppatore del software, e non documentate.

Sarà dovere del C.T. segnalare le problematiche di questa irregolarità – ed in generale, criticare le possibili carenze degli strumenti di acquisizione della

[12] Pittiruti, ≪Le indagini informatiche nel processo penale≫, p. 35.

[13] Cfr. `https://superuser.com/a/1096815`. Per alcuni studi sulla frequenza della presenza di malware in software piratato, cfr. `https://security.stackexchange.com/a/135650`.

[14] Pittiruti, ≪Le indagini informatiche nel processo penale≫, p. 36.

controparte – al giudice.

5.4 Ispezioni informatiche

Le ispezioni sono il primo mezzo di ricerca della prova menzionati nel c.p.p.:

> L'ispezione delle persone, dei luoghi e delle cose è disposta con decreto motivato quando occorre **accertare le tracce e gli altri effetti materiali** del reato.[15]

La L. n. 48/2008 ha esplicitamente previsto che anche i sistemi informatici e telematici possano essere sottoposti ad ispezione:

> L'autorità giudiziaria può disporre **rilievi segnaletici, descrittivi e fotografici e ogni altra operazione tecnica**, anche in relazione a **sistemi informatici o telematici**, adottando misure tecniche dirette ad assicurare la conservazione dei dati originali e ad impedirne l'alterazione.[16]

Sono possibili due interpretazioni dell'ambito dell'attività ispettiva:

- **Riduttiva** – l'ispezione deve limitarsi alla mera osservazione del sistema informatico, nelle sue componenti hardware, senza potervi interagire. Si può solo prendere nota dello stato dell'apparecchiatura, se è presente uno schermo, cosa è visibile, ma non si può "toccare" nulla.[17] Questa interpretazione riprende la nozione classica dell'ispezione, "chi effettua un'ispezione usa gli occhi, mentre chi compie una perquisizione utilizza le mani".[18]

[15] Cfr. art. 244 co. 1 c.p.p.
[16] Cfr. art. 244 co. 2 c.p.p.
[17] Pittiruti, ≪Le indagini informatiche nel processo penale≫, p. 42.
[18] Murgia, ≪Prova informatica e processo penale≫, p. 147.

- **Espansiva** – l'ispezione può essere estesa anche ai contenuti "immateriali" del dispositivo: pertanto, l'operatore può anche interagire con il dispositivo, prendere visione dei programmi aperti, e dei contenuti dei dispositivi di memoria collegati.[19]

La prima opzione è troppo limitante, e quindi è da scartare. La seconda interpretazione è preferibile, ma si pone il rischio di sconfinare nella perquisizione.[20]

5.5 Perquisizioni informatiche

La perquisizione consiste nella una ricerca di elementi specifici:

> Quando vi è fondato motivo di ritenere che **dati, informazioni, programmi informatici o tracce comunque pertinenti al reato** si trovino in un sistema informatico o telematico, ancorché protetto da misure di sicurezza, ne è disposta la perquisizione, adottando misure tecniche dirette ad assicurare la conservazione dei dati originali e ad impedirne l'alterazione.[21]

Il "fondato motivo" consiste nella "sufficiente probabilità, e non la mera possibilità" che i sistemi contengano dati utili.[22] Nel caso dei *cybercrimes* la presenza di dati è *ipso facto*.

Una volta trovate, le cose (dati) sono sequestrate:

> Le cose rinvenute a seguito della perquisizione sono sottoposte a sequestro con l'osservanza delle prescrizioni degli articoli 259 e

[19] Pittiruti, ≪Le indagini informatiche nel processo penale≫, p. 43.
[20] Cfr. la sezione 5.6.
[21] Cfr. art. 247 co. 1-bis c.p.p.
[22] Gammarota, ≪Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali≫, p. 81.

260.[23]

5.6 Distinzione fra ispezioni e perquisizioni informatiche

I due istituti **tendono a sovrapporsi** nel mondo digitale, data la natura dematerializzata del loro oggetto.[24] È opportuno definire un *test* che permetta di distinguere se un determinato atto è una ispezione o perquisizione.[25]

5.6.1 Definizioni tradizionali

Una prima possibilità consiste nel recuperare le **definizioni tradizionali**: si ha perquisizione solo se si va alla ricerca di elementi specifici, altrimenti è ispezione. L'utilità di questa definizione è limitata dal fatto che quando l'oggetto dell'indagine sono dati digitali, è difficile stabilire se l'operatore sta andando dalla ricerca di elementi specifici o meno.

In entrambi i casi, deve aprire file e cartelle, operazione che è affine al rovistare in un armadio od un cassetto.[26] Che si stia rovistando alla cieca fa poca differenza, perché si è comunque più vicini alla perquisizione.

In particolare, si pensi all'uso di comandi per la ricerca di file: il loro uso comporta lo sconfinamento nella perquisizione? E se l'operatore li sta usando solo per accertare la presenza di determinati file? Da questo punto di vista, la soluzione che lascia meno ambiguità è quella di escludere ogni accesso ai dati.

[23] Cfr. art. 252 c.p.p.

[24] *Tale sovrapposizione si accentua nella prassi, ove ispezione e perquisizione vengono compiute con le medesime modalità operative.* Cfr. Pittiruti, ≪Le indagini informatiche nel processo penale≫, pp. 43–45.

[25] Murgia, ≪Prova informatica e processo penale≫, pp. 147–148.

[26] Tanto che "file" è il termine inglese per "archivio".

Pertanto, l'ispezione consiste solo nella visualizzazione di programmi e informazioni già aperti e visibili a schermo,[27] e della cattura della memoria volatile.[28] Più precisamente, l'operatore può muovere le finestre già aperte, e visionarne il contenuto. Non deve interagire con i programmi, perché ciò può portare a modificazioni delle memorie primarie e secondarie.

Le uniche eccezioni sono l'apertura di programmi strettamente necessari per la cattura dei dati, o altre interazioni che sono imprescindibili ai fini dell'ispezione. Idealmente, per ogni decisione presa se ne deve dare motivazione in tempo reale, e l'operazione deve essere registrata, così che le parti abbiano più elementi di valutazione in sede di dibattimento.

5.6.2 Misure di sicurezza

Una seconda possibilità è di tracciare il confine sulla base di **presenza di misure di sicurezza**. Si può affermare che si ha perquisizione solo nel caso in cui il sistema sia protetto da "misure di sicurezza",[29] e ispezione negli altri casi.

Tuttavia, questa interpretazione pare valorizzare eccessivamente il riferimento alle misure di sicurezza. È molto più probabile che il legislatore intendesse **legittimare esplicitamente** la possibilità di forzare misure di sicurezza nel contesto di una perquisizione, e non viceversa – di definire la perquisizione **solo** nel contesto della presenza di misure di sicurezza.

[27] In ultima analisi, questa interpretazione riprende il senso tradizionale dell'ispezione perché non va alla ricerca di "cose nascoste", ma si limita a "fotografare" e prendere cognizione di quanto è già visibile.

[28] Operazione che va compiuta il prima possibile, per evitare la dispersione di tracce.

[29] Cfr. art. 247 co. 1-bis c.p.p.

5.6.3 Sistema acceso o spento

Una terza possibilità si basa su una semplice distinzione oggettiva: i sistemi **trovati accesi** o **disponibili online**, come i servizi di *cloud storage* o *cloud computing* sono sempre soggetto di ispezione, i sistemi **trovati spenti** sono sempre soggetto di perquisizione.[30]

Questa distinzione ha il pregio della semplicità, e coerenza con le *best pratices*. Le operazioni su un sistema acceso non possono essere invasive, per evitare di modificare i dati. Invece, un sistema trovato spento può essere disassemblato, i supporti che contengono dati sono collegati a dispositivi hardware che li rendono disponibili in sola lettura, e gli operatori possono liberamente procedere all'esplorazione del supporto.

L'unico elemento dubbio consiste nel fatto che fra i **dati disponibili online**, ispezione e perquisizione **tornano a sovrapporsi**. Tuttavia, rimane il fatto che i dati ricevuti da un server remoto non sono l'"essenza", bensì possono essere stati modificati prima del transito,[31] e quindi consistere solo in una "apparenza" dei dati originari. Pertanto, **pare ragionevole** parlare di una "**ispezione**", per due ordini di motivi:

- La "perquisizione" dovrebbe riguardare i dati originali residenti sul server, e non quanto il server invia;[32]

[30] Cfr. le sezioni 3 e 4 in Stefano Aterno (2014). ≪Digital forensics (investigazioni informatiche)≫. In: *Digesto delle Discipline Penalistiche, VIII Aggiornamento*. UTET Giuridica.

[31] Si fa riferimento alle "pagine dinamiche", in cui c'è una distinzione fra i dati usati per generare la pagina, e la pagina stessa. *Il contenuto finale viene determinato solo quando il visitatore richiede la pagina al server Web. Poiché il contenuto finale della pagina varia da richiesta a richiesta in base alle azioni eseguite dal visitatore, questo tipo di pagina viene definito pagina dinamica. [...] Ad esempio, una pagina dinamica può indicare al server applicazioni di estrarre dei dati da un database e inserirli nel codice HTML della pagina.* Cfr. Adobe (2017). *Applicazioni Web*. URL: `https://helpx.adobe.com/it/dreamweaver/using/web-applications.html`.

[32] Cfr. la sezione 4.4.3 per le problematiche relative al reperimento di dati sui server remoti.

- Il "sequestro" di siti internet è difficile da realizzare, data la facilità con cui è possibile copiare i dati che compongono un sito su un altro server. Al più, è possibile bloccare l'acceso ad un determinato sito,[33] ma esistono modalità per aggirare questi blocchi.[34]

5.6.4 Acquisizione forense

Una quarta possibilità è quella di **valorizzare** il momento della **acquisizione forense**. Attualmente, le operazioni di ispezione e perquisizione tendono a sovrapporsi proprio perché realizzate nello stesso modo, l'acquisizione di una copia e la sua analisi.[35]

Pertanto, si può affermare che si ha **perquisizione** se l'operatore **acquisisce** dei dati (copia della RAM, copia dei supporti, cattura di pacchetti...), ma in assenza di copia, finché l'operatore si limita alla **sola visualizzazione**, si ha **ispezione**. Il fondamento di questa distinzione è fatto che il **sequestro** (la naturale conseguenza della perquisizione) consiste nell'acquisire una **copia** dei dati.

Ovviamente, questa interpretazione non deve essere intesa nel senso per cui l'ispezione non vada eseguita secondo le *best practices*. Data la fragilità dei dati, è sempre necessario fare attenzione con qualsiasi operazione che

[33] Ad es., il sito "Project Gutenberg" è stato bloccato in Italia perché sospettato – erroneamente – di distribuire materiale coperto da copyright. *Ora, è possibile che sul sito del Progetto Gutenberg fossero state caricate erroneamente opere che in Italia sono ancora protette dal diritto d'autore, ma qualunque indagine [...] avrebbe potuto distinguere tra uno dei progetti più ambiziosi e rinomati di tutto l'Internet, una vera e propria istituzione fondata nel 1971, e "downmagaz punto com."* Cfr. Alessandro Massone (2020). *La Procura di Roma ha bloccato l'accesso a Project Gutenberg, la più grande biblioteca di internet*. URL: `https://thesubmarine.it/2020/05/25/procura-roma-bloccato-accesso-project-gutenberg/`.

[34] La modalità più semplice consiste nel riaprire lo stesso sito sotto un dominio diverso.

[35] *In entrambi i casi [ispezione e perquisizione], come accennato, si realizza un duplicato del sistema informatico o telematico (copia forense). Creato, quindi, un ulteriore duplicato, su quest'ultimo vengono svolte le opportune operazioni tecniche.* Cfr. Pittiruti, ≪Le indagini informatiche nel processo penale≫, p. 45.

coinvolga l'originale. Anzi, secondo questa impostazione, è proprio il fatto che l'ispezione non comporta un'acquisizione, che impone l'uso di **maggiori** accorgimenti – si sta rischiando di modificare il sistema, senza nemmeno acquisirne una copia.

5.7 Rinvenimento di altre notizie di reato

Data l'enorme quantità di dati personali contenuti negli hard disk, e a disposizione degli investigatori, c'è il rischio che la ricerca della prova diventi un "indebito strumento di reperimento di diverse notizie di reato".[36]

La giurisprudenza richiede che il decreto di perquisizione sia **fondato** e **sufficientemente specifico** per evitare "perquisizioni esplorative", e si dovrebbe preferire lo strumento della richiesta di consegna (art. 248 c.p.p.). Tuttavia, rimane il problema che nella maggioranza dei casi, tali elementi vengono trovati casualmente, anche durante le ispezioni. Pertanto, sarebbe opportuno disciplinare esplicitamente le procedure da seguire in caso di ritrovamento di elementi relativi a reati estranei a quello in corso di investigazione.[37]

5.8 Richiesta di consegna

Se la perquisizione riguarda la ricerca di una cosa determinata, l'autorità giudiziaria può invitare il possessore a consegnarla, per evitare atti invasivi della sfera delle persone, salvo si ritenga utile procedere comunque alla perquisizione per completezza.[38] La norma può essere interpretata anche per includere i **dati informatici**. È ragionevole supporre che l'acquisizione avvenga nel contesto di un incidente probatorio, per evitare successive

[36] Pittiruti, ≪Le indagini informatiche nel processo penale≫, p. 45.
[37] Ibid., pp. 44–45.
[38] Cfr. art. 248 co. 1 c.p.p.

contestazioni da parte del P.M. nel caso di acquisizioni fatte direttamente dal soggetto indagato.

Il problema fondamentale è che un singolo file è intrinsecamente meno attendibile di una acquisizione completa del supporto che contiene quel file, perché l'acquisizione contiene più elementi per valutare la sua attendibilità. È fin troppo semplice modificare un singolo file in maniera che sia favorevole a sé, e consegnando solo quel file, è difficile valutare se è stato modificato.

Pertanto, è ragionevole pensare che il P.M. farà pressione per procedere al sequestro dell'intero dispositivo, perché "utile [...] per la completezza delle indagini".[39]

L'A.G. può anche esaminare "dati, informazioni e programmi informatici" presso "banche", e procedere a sequestro in caso di rifiuto.[40]

È bene notare che l'espressione "banche" continua a riferirsi esclusivamente agli istituti di credito, e non alle "banche dati", che invece rientrano nell'ambito dell'espressione "sistemi informatici e telematici". Questa interpretazione estensiva era stata utilizzata da una procura per giustificare l'accesso ad una banca dati di una compagnia aerea, con fini peraltro meramente esplorativi.[41]

[39] Cfr. art. 248 co. 1 c.p.p.

[40] Cfr. art. 248 co. 2 c.p.p.

[41] *Pertanto, è da escludere un preventivo ed indefinito monitoraggio del sistema predetto in attesa dell'eventuale e futura comparsa del dato da acquisire a base delle indagini: si verrebbe altrimenti ad integrare un nuovo ed anomalo strumento di ricerca della prova, con finalità nettamente esplorative, di mera investigazione (paragonabile alle intercettazioni), che nulla ha a che fare con la perquisizione. [...] La locuzione [...] laddove richiama le "banche" [...] non può che riferirsi [...] solo agli istituti di credito [...]. Nulla consente di dilatare estensivamente l'accezione di "banche" fino a comprendere le "banche-dati" presenti, per giunta in continuo aggiornamento automatico, presso qualsiasi altro ente o struttura privata o pubblica, tanto più che, come acutamente rilevato dall'Ordinanza impugnata, il termine banca-dati, omologo della corrispondente espressione inglese "data-base", non risulta mai adoperato dall'ordinamento giuridico italiano il quale, laddove ha inteso riferirsi ad un centro di raccolta ed gestione di dati informatici, ha impiegato la diversa specifica dizione di "sistema informatico o telematico" [...].* Cfr. la sent. Cass. n. 19618/2012, richiamata in

5.9 Sequestro di dati informatici

5.9.1 Oggetto del sequestro informatico

Sequestro di dati informatici

Il **sequestro** ha come oggetto:

> [Il] corpo del reato e delle cose pertinenti al reato necessarie per l'accertamento dei fatti.[42]

Il **corpo del reato** è definito come:

> Le cose sulle quali o mediante le quali il reato è stato commesso nonché le cose che ne costituiscono il prodotto, il profitto o il prezzo.[43]

Il fatto che il sequestro può riguardare **dati digitali** non è menzionato esplicitamente nell'art. 253 c.p.p., ma ciò è largamente desumibile da altri articoli:

- Se si possono eseguire perquisizioni aventi ad oggetto dati informatici[44] allora è naturale che si possano sequestrare dati informatici, perché le "cose" ritrovate a seguito di perquisizione sono sottoposte a sequestro;[45]
- È possibile sequestrare "oggetti di corrispondenza, anche se inoltrati per via telematica";[46]

Pittiruti, ≪Le indagini informatiche nel processo penale≫, pp. 46–47.

[42] Cfr. art. 253 co. 1 c.p.p.
[43] Cfr. art. 253 co. 2 c.p.p.
[44] Cfr. art. 247 co. 1-bis c.p.p.
[45] Cfr. art. 252 c.p.p.
[46] Cfr. art. 254 co. 1 c.p.p.

- È prevista una particolare procedura di sequestro per i dati detenuti presso fornitori di servizi;[47]
- I dati digitali sono esplicitamente menzionati nell'articolo relativo alla conservazione di cose che possono alterarsi;[48]

Sequestro di criptovalute

In particolare, il riferimento al "profitto" e "prezzo" del reato sono interessanti, perché le criptovalute (Bitcoin, Ethereum, ecc...) sono spesso usate in attività criminali.[49]

Il meccanismo di funzionamento si basa sul salvataggio di informazioni private in un "portafoglio" (*wallet*). Ad ogni *wallet* è associato un indirizzo (*address*), che lo identifica in maniera univoca, ed è composto da una sequenza di lettere e numeri.[50]

Mentre le transazioni sono pubbliche, e facilmente consultabili mediante i *blockchain explorers*,[51] gli **utenti** che compiono queste transazioni rimangono tendenzialmente anonimi, proprio grazie al fatto che gli indirizzi non sono "trasparenti", e non permettono di risalire al *wallet*, né – tanto meno – all'identità del suo utilizzatore.

Tuttavia, **l'anonimato non è perfetto**. Dato che le transazioni sono pubbliche, è possibile analizzarle e cercare correlazioni, oppure è possibile

[47] Cfr. art. 254-bis co. 1 c.p.p. e la sezione 5.9.6.

[48] Cfr. art. 260 co. 2 c.p.p.

[49] Uno dei casi che ha avuto maggiore importanza è stato il mercato nero on-line chiamato "Silk Road". *In February of 2011, Ross William Ulbricht, who went by the nom de guerre of "Dread Pirate Roberts," founded the site "Silk Road." Ulbricht [...] dreamt of an online marketplace where people would be able to buy and sell narcotics and other illicit items, without governmental interference.* Cfr. David Adler (2018). *Silk Road: The Dark Side of Cryptocurrency*. URL: `https://news.law.fordham.edu/jcfl/2018/02/21/silk-road-the-dark-side-of-cryptocurrency/`.

[50] Ad es., `1BvBMSEYstWetqTFn5Au4m4GFg7xJaNVN2`. Esempio preso da Bitcoin Wiki contributors (n.d.). *Address*. URL: `https://en.bitcoin.it/wiki/Address`.

[51] Ad es., v. `https://www.blockchain.com/btc/unconfirmed-transactions`.

rinvenire un *wallet* durante le perquisizioni di un computer. Nel primo caso, è possibile richiedere il sequestro rivolgendosi alle società che gestiscono gli scambi (*exchanges*),[52] mentre nel secondo è sufficiente acquisire una **copia** del *wallet*, e poi trasferire i fondi contenuti su un *wallet* controllato dalle autorità giudiziarie.

5.9.2 Impugnazione del sequestro di dati digitali

Il sequestro può essere impugnato:

> Contro il decreto di sequestro l'imputato, la persona alla quale le cose sono state sequestrate e quella che avrebbe diritto alla loro restituzione possono proporre richiesta di riesame, anche nel merito, a norma dell'articolo 324.[53]

Il **presupposto per l'impugnazione** del sequestro è l'esistenza di un **vincolo** – più precisamente, il **sequestro (tradizionale)** produce una compressione **del diritto di proprietà**. Tuttavia, nel caso di dati digitali, il diritto di proprietà si ferma al supporto su cui i dati sono memorizzati. I dati digitali sono dematerializzati, e quindi non sono "cose", perché carenti di corporalità, e quindi non possono formare oggetto di diritti, se non per mezzo di *fictio* – come ad es., l'equiparazione dell'energia elettrica ad un bene materiale, affinché si potesse sanzionare penalmente il suo furto.[54] Per i dati digitali, questa equiparazione manca.

[52] *L'assicuratore ha deciso di denunciare l'attacco alle autorità giudiziarie, ed ha ingaggiato una società di analisi blockchain (Chainalysis) per rintracciare i bitcoin pagati come riscatto. In questo modo si è scoperto che [...] la maggior parte, ovvero 96 BTC, erano stati trasferiti ad un indirizzo di Bitfinex. Per questo motivo l'Alta Corte ha emesso l'ordine di sequestro nei confronti dell'exchange [...].* Cfr. Marco Cavicchioli (2020). *Bitfinex: sequestro di 96 bitcoin. Una ulteriore prova del non anonimato di BTC*. URL: `https://cryptonomist.ch/2020/01/28/bitfinex-sequestro-96-bitcoin/`.

[53] Cfr. art. 257 co. 1 c.p.p.

[54] Gammarota, ≪Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali≫, pp. 59–60.

Nel caso del sequestro informatico, la prassi consiste nell'**acquisire il contenuto del supporto, e poi restituirlo**.[55] Specie nel caso in cui i dati informatici sono considerati come un tipo di documento,[56] la possibilità di **restituire l'originale** dopo averne estratto una copia è **espressamente prevista**:

> L'autorità giudiziaria può fare estrarre copia degli atti e dei documenti sequestrati, restituendo gli originali [...].[57]

Ancora, ed in maniera più specifica:

> L'autorità giudiziaria fa estrarre copia dei documenti e fa eseguire fotografie o altre riproduzioni delle cose sequestrate che possono alterarsi o che sono di difficile custodia [...]. Quando si tratta di dati, di informazioni o di programmi informatici, la copia deve essere realizzata su adeguati supporti, mediante procedura che assicuri la conformità della copia all'originale e la sua immodificabilità.[58]

Pertanto, la Corte tradizionalmente **negava la possibilità di impugnare il sequestro** di dati digitali, affermando che l'**interesse a riavere la "cosa" terminava** nel momento in cui il **supporto** su cui erano memorizzati i dati **veniva restituito**, anche se i dati rimanevano in mano agli inquirenti. Affinché si potesse procedere al dissequestro, si doveva dimostrare che "il valore dell'informazione risiede nell'esclusività del suo controllo".[59]

[55] Laura Bartoli (2016). «La catena di custodia del materiale informatico: soluzioni a confronto». In: *Anales de la faculdad de derecho – Universidad de La Laguna* 33, pp. 145–162. URL: https://www.ull.es/revistas/index.php/derecho/article/view/86, p. 2.

[56] Cfr. le considerazioni nella sezione 4.2.

[57] Cfr. art. 258 co. 1 c.p.p.

[58] Cfr. art. 260 co. 2 c.p.p.

[59] Laura Bartoli (2018). «Sequestro di dati a fini probatori: soluzioni provvisorie a incomprensioni durature». In: *Archivio penale* 1. URL: http://www.archiviopenale.it/sequestro-di-dati-a-fini-probatori-soluzioni-provvisorie-a-incomprensioni-durature/articoli/15338, p. 5.

Negli anni, l'evoluzione giurisprudenziale ha portato al riconoscimento di una netta distinzione fra "supporto" ed "informazione". Entrambe possono essere sottoposte al vincolo del sequestro.[60]

Attualmente, se un supporto viene acquisito mediante **copia forense** (*bitstream image*), la Cassazione permette l'impugnazione della duplicazione in sé, proprio perché il **patrimonio informativo** ha un suo **riconoscimento autonomo**, e la presenza di una copia viene considerata *ipso facto* una **compressione del godimento di quel patrimonio**.[61]

Un limite alla restituzione del supporto, anche a seguito dell'acquisizione di copia, è dato nel caso in cui ciò comporti l'aggravamento del reato, o permetta la commisione di nuovi reati:[62]

> Non si fa luogo alla restituzione e il sequestro è mantenuto ai fini preventivi quando il giudice provvede a norma dell'articolo 321.[63]
>
> Quando vi è pericolo che la libera disponibilità di una cosa pertinente al reato possa aggravare o protrarre le conseguenze di esso ovvero agevolare la commissione di altri reati, a richiesta del pubblico ministero il giudice competente a pronunciarsi nel merito ne dispone il sequestro con decreto motivato.[64]

[60] *Sul tema, la legge è disseminata d'indicazioni che le sezioni unite hanno pazientemente spigolato, rintracciando nei singoli frammenti una direzione talmente univoca da concludere che le informazioni sono "pacificamente" autonome dal dispositivo in cui alloggiano. Informazione e supporto sono dunque entità distinte, che possono essere indipendentemente vincolate dagl'inquirenti [...].* Cfr. Bartoli, ≪La catena di custodia del materiale informatico: soluzioni a confronto≫, pp. 6–7.

[61] *Accanto al diritto di proprietà, si accorda tutela anche all'esclusivo godimento del patrimonio informativo. [...] Se è stata formata un'immagine forense, saremo davanti a un vincolo vero e proprio che merita d'essere sottoposto a una verifica; si seguiranno anche qui le regole di un normale riesame.* Cfr. Bartoli, ≪Sequestro di dati a fini probatori: soluzioni provvisorie a incomprensioni durature≫, pp. 13–14.

[62] Pittiruti, ≪Le indagini informatiche nel processo penale≫, p. 53.

[63] art. 262 co. 3 c.p.p.

[64] art. 321 co. 1 c.p.p.

5.9.3 Distinzione fra sequestro, e ispezione e perquisizione

A questo punto però sorge una questione. Se il sequestro consiste nella **copia di dati**, e l'ispezione e la perquisizione sono anche loro compiute mediante l'acquisizione forense di dati – per soddisfare le garanzie richieste dal codice, di non modificare i dati originali, e di ottenere copie conformi all'originale – come si possono distinguere le tre operazioni?

La natura dematerializzata dei dati digitali porta ad una perversione degli schemi tradizionali, in cui non solo è difficile distinguere tra i tre istituti, perché vengono attuati tendenzialmente con le stesse modalità, ma in cui anche l'ordine stesso delle operazioni è invertito: "prima si sequestra tutto e poi si perquisisce", perché "la bit stream image è la procedura più sicura per quanto riguarda la conservazione dell'originale, e la sua tutela è posta dal codice come una priorità".[65]

Si è già discusso delle possibili modalità di distinzione fra ispezione e perquisizioni.[66]

Data la naturale continuità fra perquisizione e sequestro, trovare una distinzione è più difficile.

Una possibilità consiste nel riprendere il modello delle "**udienze stralcio**" previsto per le intercettazioni. Secondo questo schema, prima si procede effettivamente a sequestro, e poi a perquisizione. Una volta compiuta a posteriori la selezione dei materiali utili – e di fatto è difficile fare altrimenti, per ridurre al minimo le interazioni con i supporti materiali – il resto dei dati distrutto.[67]

[65] Bartoli, ≪Sequestro di dati a fini probatori: soluzioni provvisorie a incomprensioni durature≫, p. 17.

[66] Cfr. la sezione 5.6.

[67] Bartoli, ≪Sequestro di dati a fini probatori: soluzioni provvisorie a incomprensioni durature≫, p. 17.

La soluzione sacrifica l'ordine naturale delle operazioni, ma ha il pregio di rispettare il principio di proporzionalità. Un limite tuttavia sussiste nel fatto che richiede la collaborazione delle parti, e quindi può compromettere lo svolgimento delle indagini.

Una seconda possibilità consiste nella valorizzazione di **elementi formali**, come l'apposizione di sigilli:

> Le cose sequestrate si assicurano con il sigillo dell'ufficio giudiziario e con le sottoscrizioni dell'autorità giudiziaria e dell'ausiliario che la assiste ovvero, in relazione alla natura delle cose, con altro mezzo, anche di carattere elettronico o informatico, idoneo a indicare il vincolo imposto a fini di giustizia.[68]

Pertanto, fino al momento in cui si appone il sigillo, si ha una perquisizione, mentre dal momento in cui il sigillo viene apposto – anche se si tratta di una operazione puramente formale – si ha un sequestro. La menzione di "carattere informatico" rimanda alla nozione di **firma digitale**[69] e **marca temporale**,[70] che insieme danno piena certezza in merito a tre elementi:

- **Chi** ha partecipato all'apposizione del sigillo;
- **Quando** il sigillo è stato apposto;
- **L'integrità** dei dati che sono stato oggetto di sequestro.

[68] Cfr. art. 260 co. 1 c.p.p.

[69] *La Firma Digitale è l'equivalente informatico di una tradizionale firma autografa apposta su carta [...] garantisce l'identità del sottoscrittore [...] assicura che il documento non sia stato modificato dopo la sottoscrizione [...] attribuisce piena validità legale al documento firmato.* Cfr. Aruba.it (n.d.[a]). *Cos'è la Firma Digitale*. URL: `https://www.pec.it/firma-digitale.aspx`.

[70] *La Marca Temporale è un servizio che permette di associare data e ora certe e legalmente valide ad un documento informatico, consentendo quindi di associare una validazione temporale opponibile a terzi.* Cfr. Aruba.it (n.d.[b]). *Cos'è la Marca Temporale*. URL: `https://www.pec.it/marche-temporali.aspx`.

In generale, sembra difficile trovare una soluzione. Il problema fondamentale consiste nell'adattamento di misure pensate per cose, oggetti dotati di una propria corporalità, a dati informatici, la cui caratteristica principale è proprio quella di essere dematerializzati. È come cercare di quadrare il cerchio usando solo la riga ed il compasso: una operazione impossibile, che può essere risolta solo con strumenti nuovi e specializzati.

5.9.4 Sequestro non indeterminato, di lunga durata

Riprendendo il tema della limitazione del sequestro, la prassi del "**sequestro non indeterminato**" è bene affermata, ed è funzionale al rispetto dei principi di proporzionalità ed adeguatezza.[71] Consiste nel delimitare il più possibile le cose da sottoporre a sequestro.

Dal punto di vista **materiale**, devono essere sequestrati solo elementi che possono contenere **dati utili**. Sono pertanto escluse le periferiche del computer (schermo, tastiera, mouse...), a meno che non siano strettamente **necessarie** (ad esempio, adattatori per supporti particolari). In secondo luogo, se è possibile identificare compiutamente dei **singoli file**, vanno sequestrati solo quelli, e **non l'intero disco**.[72]

In ogni caso, non è sempre pratico compiere una selezione dei file a priori. Una difficoltà ulteriore è data dal P.M. che potrebbe insistere per l'acquisizione completa del supporto, per **esigenze di completezza** delle indagini, secondo la filosofia del *melius abundare quam deficere*. In fondo, i dati digitali sono fragili, quindi è meglio operare per preservarli nel migliore modo possibile. Tuttavia, un limite naturale a questa linea di pensiero è data dalla lunga **durata** delle operazioni di acquisizione, e dal **costo** e **difficoltà organizzative** di archiviare grandi quantità di dati.

[71] Cfr. la pronuncia del Tribunale di Brescia in Antonio Sapio (2016). ≪L'utilità della computer forensics nel processo penale: nuove tecnologie e prassi≫. Tesi di laurea mag. URL: http://tesi.luiss.it/17040/, pp. 20–21.

[72] Pittiruti, ≪Le indagini informatiche nel processo penale≫, pp. 48–50.

Una prassi diffusa è rappresentata dal **sequestro di lunga durata**. Il P.M. affida il dispositivo sequestrato al C.T., senza disporre la copia, o dare termini per l'esecuzione della procedura. La lunga durata del sequestro non è stata considerata una violazione della legge per la Cassazione, che preferisce un approccio il più prudente possibile, ai limiti della paranoia, arrivando anche a negare il dissequestro di dispositivi di cui è stata già acquisita copia forense, nell'eventualità in cui fosse necessario svolgere altre operazioni di analisi.[73]

5.9.5 Sequestro di corrispondenza

È possibile sequestrare oggetti di corrispondenza telematica, sempre che sussista un fondato motivo:

> Presso coloro che forniscono servizi postali, telegrafici, telematici o di telecomunicazioni è consentito procedere al sequestro di lettere, pieghi, pacchi, valori, telegrammi e altri oggetti di corrispondenza, anche se inoltrati per via telematica, che l'autorità giudiziaria abbia fondato motivo di ritenere spediti dall'imputato o a lui diretti, anche sotto nome diverso o per mezzo di persona diversa, o che comunque possono avere relazione con il reato.[74]

La specificazione "sotto nome diverso" è particolarmente valida per le comunicazioni via internet, perché è estremamente frequente l'uso di pseudonimi (*nicknames*), invece del proprio nome.

Si applicano delle precauzioni particolari se al sequestro procede un ufficiale di P.G.:

> Quando al sequestro procede un ufficiale di polizia giudiziaria, questi deve consegnare all'autorità giudiziaria gli oggetti di corri-

[73] Pittiruti, ≪Le indagini informatiche nel processo penale≫, pp. 54–56.
[74] Cfr. art. 254 co. 1 c.p.p.

> spondenza sequestrati, senza aprirli o alterarli e senza prendere altrimenti conoscenza del loro contenuto.[75]

Mentre queste precauzioni sono relativamente semplici da seguire per le lettere tradizionali, per i dati digitali è più difficile visionare solo i "metadati" (mittente, destinatario, ecc. . .) senza anche prendere cognizione dei dati. Si pensi al caso in cui si sequestrano delle e-mail, spesso i programmi di posta mostrano le prime righe del messaggio. In generale, si può affermare che se il sequestro tradizionale incide sul diritto di proprietà, il sequestro di dati informatici incide profondamente sul diritto alla privacy.

In ogni caso, tutti gli elementi che sono determinati essere non sequestrabili vanno restituiti:

> Le carte e gli altri documenti sequestrati che non rientrano fra la corrispondenza sequestrabile sono immediatamente restituiti all'avente diritto e non possono comunque essere utilizzati.[76]

Nel caso di e-mail, più che di **restituzione** si può parlare di **distruzione della copia sequestrata**. Per cercare di coniugare le esigenze di accertamento e privacy dell'utente è possibile operare in questo modo:

- Si compie un'**acquisizione in blocco** dei contenuti della casella di posta, per ridurre al minimo gli accessi al server, prendendo **più di quanto necessario**;
- Lavorando sulla copia acquisita, si **selezionano le mail** che sono effettivamente utili e quindi **sequestrabili**;
- Si **informa il fonitore** di servizi quali e-mail l'A.G. ha deciso di **sequestrare**, in modo che il fonitore usi **misure tecniche** – ad es., firma digitale e marca temporale – per **garantire la loro integrità ed esistenza** sul server al momento del sequestro;

[75] Cfr. art. 254 co. 2 c.p.p.
[76] Cfr. art. 254 co. 3 c.p.p.

- Le e-mail così "**certificate**" sono inviate all'A.G., e la **copia completa** acquisita all'inizio viene **distrutta**, così che rimanga traccia solo delle **mail utili**, con misure che ne garantiscano l'autenticità e l'affidabilità.

5.9.6 Sequestro di dati presso fornitori di servizi

La L. n. 48/2008 ha introdotto un articolo che regola il **sequestro di dati** (in generale) presso fornitori di servizi:

> L'autorità giudiziaria, quando dispone il sequestro, presso i fornitori di servizi informatici, telematici o di telecomunicazioni, dei dati da questi detenuti, compresi quelli di traffico o di ubicazione, può stabilire, per esigenze legate alla regolare fornitura dei medesimi servizi, che la loro acquisizione avvenga mediante copia di essi su adeguato supporto, con una procedura che assicuri la conformità dei dati acquisiti a quelli originali e la loro immodificabilità. In questo caso è, comunque, ordinato al fornitore dei servizi di conservare e proteggere adeguatamente i dati originali.[77]

L'espressione "per esigenze legate alla regolare fornitura dei medesimi servizi" è interessante. Le normali procedure di acquisizione richiedono che il dispositivo da acquisire non sia operativo,[78] e quindi è necessario sospendere il servizio, procedere all'acquisizione dei dati, e poi riavviarlo. Questo comporta un indesiderabile *downtime* (periodo in cui il servizio non è disponibile). Nel caso di gradi fornitori, che possono ricevere molte richieste di sequestri, questo modello sarebbe insostenibile.

Presentarsi con un decreto di sequestro indispone chiunque, ed è molto probabile che i fornitori cerchino di opporsi il più possibile, invece di fornire

[77] Cfr. art. 254-bis c.p.p.

[78] Diversamente, i dati contenuti cambierebbero in continuazione, e sarebbe come cercare di fotografare qualcosa che si muove.

i dati. Specie negli ultimi anni, la protezione della privacy ha assunto una dimensione notevole,[79] e pertanto i fornitori che riducono al minimo la raccolta di dati personali,[80] oppure che usano misure tecniche per garantire la confidenzialità delle informazioni,[81] sono visti favorevolmente dagli utenti, che vedono la loro privacy maggiormente tutelata.

Inoltre, le modalità di sequestro tradizionali possono essere lunghe e difficili se i dati si trovano in una giurisdizione estera. Pertanto, la modalità migliore non è tanto il tradizionale esercizio dell'autorità statale, bensi cercare la **collaborazione** con i fornitori.[82]

Le operazioni di acquisizione sono **delegate** ai fornitori. Il **vantaggio principale** è la possibilità di ottenere l'"**essenza**" dei dati, i bit **così come sono conservati sul server**. Come già visto, l'ispezione di dati conservati su server esterni restituisce al più l'"apparenza" di quei dati.[83]

Viceversa, è possibile considerare il fatto che i fornitori abbiano i dati nella loro disponibilità come un indizio di inaffidabilità intrinseca. Una cosa è il P.M. o la P.G. che delegano le operazioni di acquisizione ad un C.T.,[84]

[79] Ad es., i vari scandali riguardanti l'uso non autorizzato di dati personali da parte di Facebook, ed il regolamento europeo sulla protezione dei dati (GPDR), ecc...

[80] Ad es., molti servizi di VPN hanno una *no-log policy* – ossia, non tengono traccia di quali richieste sono state effettuate dai loro utenti, e quindi anche nel caso in cui ci sia una richiesta di informazioni, non avrebbero nulla da fornire.

[81] Ad es., l'uso di tecniche crittografiche per proteggere i dati, in modo che solo l'utente possa accedervi. Il *mud puddle test* (test della pozzanghera) permette di stabilire chi può accedere ai dati. Se un utente dimentica la password per accedere ai propri dati (dopo essere scivolato su una pozzanghera ed aver battuto la testa), ma può comunque accedervi con l'aiuto del fornitore, senza ricordare la password, allora significa che anche il fornitore ha pieno accesso ai dati. Cfr. Matthew Green (2012). *iCloud: Who holds the key?* URL: `https://blog.cryptographyengineering.com/2012/04/05/icloud-who-holds-key/`.

[82] Così come il proverbio, "S'acchiappano più mosche con una goccia di miele che con un barile d'aceto".

[83] Cfr. la sezione 5.6.3. Ad es., nel caso di pagine dinamiche, è possibile visualizzare il codice sorgente ed i dati originari che sono utilizzati dal server per produrre quelle pagine – la vera "essenza" del sito – mentre l'ispezione permette solo di vedere una "apparenza" del sito, quanto viene recapitato all'utente a seguito di manipolazioni.

[84] Cfr. rispettivamente l'art. 359 co. 1 c.p.p. e l'art. 348 co. 4 c.p.p..

ma altro è affidarsi completamente a terzi, e doversi fidare di quanto viene inviato.

È difficile prospettare la presenza fisica del P.M. e del suo C.T., pertanto assume enorme importanza la **documentazione delle operazioni tecniche** che sono state utilizzate da parte del gestore per ottenere i dati. In particolare, si deve dimostrare:

- La **conformità della copia** ai dati originali – pertanto, almeno l'hash dei due dati;
- L'uso di **misure tecniche** per **evitare le modifiche** ai dati originali – la documentazione di misure di *access control*;
- L'**integrità** dei dati – con l'apposizione di firma digitale per garantire l'integrità, e marca temporale per avere una data certa;

Tuttavia, sospettare eccessivamente dei gestori dei dati sembra inopportuno. Basta confrontare l'obbligo previsto per il **fornitore** di "conservare e proteggere adeguatamente i dati originali"[85] con l'obbligo del **custode** di "impedirne [per i dati, informazioni o programmi informatici] l'alterazione o l'accesso da parte di terzi"[86] per verificare che dal punto di vista degli obblighi, il fornitore **può essere considerato** a tutti gli effetti un **custode** dei dati sottoposti a sequestro.

A riprova, se le cose sequestrate non possono essere conservate – o non è opportuno conservarle – presso la cancelleria o segreteria del tribunale, si nomina un custode, e si determina la modalità di custodia.[87] I dati digitali possono essere **duplicati** ed essere **custoditi in più luoghi**, e quindi **anche fuori** dalla cancelleria o segreteria.

[85] Cfr. art. 254-bis c.p.p.
[86] Cfr. art. 259 co. 2 c.p.p.
[87] Cfr. art. 259 co. 1 c.p.p.

Dato che le **indicazioni sulla conservazione** dei dati digitali sono già **previste dalla legge**, e si **rivolgono direttamente al fornitore**, il fornitore **è un custode** di quei dati ai sensi dellart. 259 c.p.p.. Di conseguenza, sono applicabili le "**pene previste dalla legge penale** per chi trasgredisce ai doveri della custodia",[88] con la conseguenza che il fornitore ha tutto l'interesse, e nessun incentivo, a fornire prove falsate.

Un problema più sottile riguarda il rispetto delle *best practices* da parte dei fornitori. È ragionevole supporre che i fornitori siano al corrente delle migliori pratiche di archiviazione e riproduzione dei dati, proprio perché è il fondamento della loro attività. In ogni caso, a scanso di equivoci, l'A.G. può imporre al fornitore-custode l'uso di determinate misure di custodia e presentazione.

[88] Cfr. art. 259 co. 2 c.p.p.

Capitolo 6

Digital evidence e prova scientifica

6.1 *Digital evidence* come prova scientifica

La **prova scientifica** è il tipo di prova che è ottenuta mediante:[1]

- **Leggi scientifiche** – prova scientifica in senso stretto;
- **Metodi tecnologici** – prova tecnologica.

Le **scienze di riferimento** per la *digital evidence* sono:

- **Informatica** – la "scienza che studia l'elaborazione delle informazioni e le sue applicazioni [...] [la] rappresentazione [...] organizzazione e [...] trattamento automatico della informazione";[2]

[1] Di Pinto, ≪La prova scientifica nel processo penale≫, p. 909.

[2] Enciclopedia Treccani (n.d.). *Informatica*. URL: http://www.treccani.it/enciclopedia/informatica/.

- **Scienza forense** – lo studio del valore processuale dei fatti,[3] della ricerca delle prove, e del collegamento dei fatti ai soggetti per l'accertamento della loro responsabilità.[4]

La **natura tecnologica** della prova scientifica è *ipso facto*, perché la sua acquisizione e gestione richiede l'uso di strumenti tecnologici.

6.2 Ammissibilità della prova scientifica

6.2.1 Test di Frye

Tradizionalmente, il giudice era vincolato dalla *communis opinio* della scienza. Le **prove scientifiche ammissibili** al processo sono solo quelle fondate su **metodi scientifici generalmente attestati**.

Questo principio venne espresso nella sentenza del 1923 *Frye v. United States*.[5] Il caso concreto riguardava la possibilità di usare la misurazione della pressione sanguigna – una specie di "macchina della verità" – per verificare se un soggetto stesse mentendo:

> In other words, the theory seems to be that truth is spontaneous, and comes without conscious effort, while the utterance of a falsehood requires a conscious effort, which is reflected in the blood pressure. [...] If the subject is telling the truth, the pressure registers highest at the beginning of the examination, and gradually diminishes as the examination proceeds.

[3] Giovanni Ziccardi (2006). «Scienze forensi e tecnologie informatiche: la computer and network forensics». In: *Informatica e diritto* 25.2, pp. 103–125. URL: http://www.ittig.cnr.it/risorse/attivita-editoriale/rivista/indici/2-2006/, p. 105.

[4] Ibid., p. 109.

[5] United States Court of Appeals, District of Columbia Circuit. (1923). *Frye v. United States*. URL: https://en.wikisource.org/wiki/Frye_v._United_States/Opinion_of_the_Court.

La corte argomentò che la teoria della misurazione della pressione sanguigna non aveva ottenuto un riconoscimento generale (*general acceptance*) da parte del mondo scientifico:

> While courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.
>
> We think the systolic blood pressure deception test has not yet gained such standing and scientific recognition among physiological and psychological authorities [...]

6.2.2 Test di Daubert

La sentenza del 1993 *Daubert v. Merrell Dow Pharmaceuticals*[6] capovolge completamente il principio enunciato in *Frye v. United States*. Invece di vincolare il giudice al parere dominante della comunità scientifica, il giudice viene lasciato libero nella sua valutazione dell'ammissibilità della prova scientifica. Sono indicati alcuni punti da considerare:[7]

- **Falsificabilità** – le teorie o metodi devono essere falsificabili (secondo la definizione di Popper) ed essere comprovati empiricamente;
- ***Peer review*** – le teorie devono essere pubblicate, e la pubblicazione deve essere preceduta da procedure di *peer review*;
- **Tasso di errore** – il tasso di errore (noto o potenziale) deve essere il più ridotto possibile;

[6] United States Supreme Court (1993). *Daubert v. Merrell Dow Pharmaceuticals*. URL: `https://caselaw.findlaw.com/us-supreme-court/509/579.html`.

[7] Brighi e Maioli, ≪Un cambio di paradigma nelle scienze forensi. Dall'armonizzazione tecnico-giuridica a una nuova cornice epistemologica≫, p. 228.

- ***Communis opinio*** – le teorie devono essere generalmente accettate dalla comunità scientifica;

Prese insieme, queste caratteristiche permettono sia di evitare l'entrata della pseudoscienza (*junk science*) nei processi, sia – valorizzando in particolare i primi tre elementi – di utilizzare teorie nuove, che anche se non sono ancora generalmente accettate dalla comunità scientifica, hanno comunque dimostrato un certo grado di scientificità.

Per quanto riguarda la *digital evidence*, esiste già una *communis opinio* sulle modalità di **acquisizione** dei dati digitali. La possibilità di metodi di **analisi innovativi** può essere riconosciuta, purché la loro validità sia dimostrata in giudizio.

6.3 Margini di incertezza della prova scientifica

6.3.1 Rapporto fra scientificità ed attendibilità

In primo luogo, è importante chiarire la differenza fra "scientificità" ed "attendibilità" (della prova). Lo **scopo della scienza non è la ricerca della verità in sé**, ma la costruzione di **modelli teorici** che imitano il funzionamento della **realtà empirica**. Il metodo scientifico si ispira al modello argomentativo socratico: raggiungere la verità con **approssimazioni progressive**, avanzando una teoria (*pars construens*) e cercando un caso che dimostra che quella teoria è insufficiente (*pars destruens*).

Le teorie scientifiche vengono messe alla prova con esperimenti empirici. È erroneo affermare che una teoria sia "assolutamente valida". Piuttosto, è "valida fino a prova contraria", o più precisamente "capace di spiegare solo un numero limitato di casi". Al di fuori del numero degli esperimenti positivi, non è certo che la teoria sia affidabile. Pertanto, lo scienziato sa già che la **teoria** con cui sta lavorando **è inattendibile**, e quello che sta

cercando non è la verità, ma un **esperimento** che **dimostri** che la teoria è **incompleta**.

Ovviamente, nel tempo questo comporta che le teorie diventino **sempre più attendibili**, ma esisterà sempre un **margine di incertezza**, che può essere misurato, ma non eliminato.

6.3.2 Esempi di margine di incertezza

Ad esempio, l'**identificazione di persone dalla loro voce** è un'attività astrattamente possibile, e fondata su basi tecniche e scientifiche, ma c'è sempre un margine di errore.[8]

Ancora, esistono metodi per **identificare univocamente un singolo dispositivo**. Ad esempio *AmIUnique*[9] esegue una serie estensiva di test, per determinare quanto un dispositivo sia univocamente identificabile. Le **singole voci** presentano un certo **margine di errore** (più elevata è la percentuale, meno univoco è il valore), ma **se considerate insieme**, permettono di **identificare e tracciare lo stesso dispositivo**, anche se l'indirizzo IP cambia. Tuttavia, l'identificazione del soggetto che in concreto usa il dispositivo rimane comunque molto più difficile da determinare.

Un altro esempio interessante di incertezza è dato nel campo della ***audio/video forensics***,[10] campo che riguarda la trattazione ed analisi di registrazioni di suoni e videoriprese in formato digitale. Il caso tipico è l'uso di immagini o video catturati da dispositivi di videosorveglianza, "che pur registrando l'evento criminoso risultano inutilizzabili per la scarsa qualità del sistema (scarsa risoluzione, rumore, ecc.)". In questi casi, il C.T. può utilizzare tecniche di analisi dei dati che permettono di migliorare la qualità della prova.

[8] Di Pinto, ≪La prova scientifica nel processo penale≫, p. 917.
[9] V. `https://amiunique.org/`
[10] Battiato, Farinella e Puglisi, *Image/video forensics: casi di studio*.

Ovviamente, il software **non può creare dettagli** che non esistono nell'originale, ma si può cercare di migliorare la qualità da un punto di vista percettivo, mettendo in **evidenza** i **dettagli utili**, ed **isolandoli** dal **rumore** di sottofondo.

6.3.3 *Deepfakes*

Un altro campo che attualmente è solo emergente, ma nel futuro molto probabilmente assumerà grande importanza, è dato dai ***deepfakes***. Esistono tre tipi principali:

- ***Image deepfakes*** – dove un computer può generare l'immagine di un certo tipo di soggetto.[11]
- ***Video deepfakes***[12] – dove il volto di una persona viene sostituito con quello di un altra, in maniera fotorealistica;[13]
- ***Audio deepfakes***[14] – dove è possibile eseguire la sintesi vocale di un testo utilizzando il timbro di voce di una certa persona.[15]

La caratteristica più importante dei *deepfakes* è il fatto che la loro generazione è largamente automatizzata, e gestita dal software. Inoltre, mentre la tecnologia alla base del loro funzionamento è estremamente sofisticata, c'è una "democratizzazione" del software necessario per generarli. I programmi stessi tendono ad essere semplici da utilizzare, ed esistono anche servizi

[11] Ad es., vedi il sito `https://thispersondoesnotexist.com/`

[12] *Everybody Can Make Deepfakes Now!* (2020). URL: `https://youtu.be/mUfJOQKdtAk`.

[13] Per alcuni esempi, v. il canale Youtube *DrFakenstein*: `https://www.youtube.com/channel/UC9PLdpbloc1pfg3ds3My98w`

[14] *This AI Clones Your Voice After Listening for 5 Seconds* (2019). URL: `https://youtu.be/0sR1rU3gLzQ`.

[15] Per alcuni esempi, v. il canale Youtube *Vocal Synthesis*: `https://www.youtube.com/channel/UCRt-fquxnij9wDnFJnpPS2Q`

su internet, che non richiedono nemmeno l'installazione di programmi sul proprio computer.[16]

I *deepfakes* stanno diventando sempre più difficili da identificare,[17] con la conseguenza che per un osservatore umano, ad un certo punto sarà impossibile determinare se un determinato dato è "reale", oppure è un "falso" creato ad arte.

Così come è attualmente necessario ricorrere ad un esperto per valutare se un quadro è autentico, sarà necessario usare programmi di analisi specifici[18] per determinare se un determinato contenuto è un deepfake o meno. Pertanto, c'è la necessità di **affidarsi quasi ciecamente** al programma, perché l'occhio nudo è **meno capace** del software di analisi stesso.

Nel caso specifico dell'algoritmo di generazione chiamato GAN, l'identificazione di potenziali *deepfake* è già parte integrante del processo di generazione. Più precisamente, un sistema GAN ha due componenti:[19]

- Il ***generator*** è specializzato nel **generare** potenziali *deepfakes*, ed il suo scopo è di riuscire ad ingannare il *discriminator*;
- Il ***discriminator*** è specializzato nel **valutare** se un elemento è un *deepfake* o è reale, ed il suo scopo è di aiutare il *generator* a creare falsi che siano il più convincenti possibile.

Data la competizione fra *generator* e *discriminator* è ragionevole pensare che la capacità di identificazione dei *deepfakes* procederà di pari passo con la capacità di generazione, e che i *discriminator* più recenti siano in grado di identificare i *deepfake* creati dai *generator* venuti prima.

[16] Ad es., `https://deepfakesweb.com/` permette la creazione di *video deepfakes*.

[17] *It's Getting Harder to Spot a Deep Fake Video* (2018). URL: `https://youtu.be/gLoI9hAX9dw`.

[18] *DeepFake Detector AIs Are Good Too!* (2019). URL: `https://youtu.be/RoGHVI-w9bE`.

[19] *What's a Generative Adversarial Network? Inventor Explains* (2017). URL: `https://blogs.nvidia.com/blog/2017/05/17/generative-adversarial-networks/`.

In ogni caso, la valutazione del programma di analisi presenterà sempre un **margine di errore**, riducibile, ma non eliminabile. In particolare, si possono avere falsi positivi (una immagine reale è considerata un *deepfake*), o falsi negativi (un *deepfake* è riconosciuto come un'immagine reale).

Il problema più grave saranno i casi in cui il *discriminator* restituisce un valore che non va "oltre ogni ragionevole dubbio". La valutazione del *discriminator* è detta *confidence* – quanto è "convinto" della propria decisione – e consiste in un numero decimale nell'intervallo da 0.0 a 1.0, che può essere anche espresso come una percentuale.[20]

Più la *confidence* è vicina ad uno dei due estremi, più è certo che l'immagine sia o meno un *deepfake*. Ma nel caso ipotetico – e si spera che rimanga tale! – in cui la valutazione sia un valore intermedio, come 30%, e per un essere umano è impossibile distinguere se l'immagine è un *deepfake*, come si procederà? Ai posteri l'ardua sentenza.

6.4 Deriva tecnicistica del processo

Il problema più importante riguardo la prova scientifica è rappresentato non tanto dalla fallibilità della scienza, quanto dalla fallibilità dell'uomo. La figura del perito e dei CC.TT. esiste perché il giudice non è esperto in materie tecniche.

Il risultato è che il giudice tende a ritenere la prova scientifica **quasi una prova legale**, una "prova regina", così come era la confessione dell'imputato. Il **perito si sostituisce al giudice**, che prende semplicemente atto delle sue conclusioni e **non muove critiche**, ed il **potere decisionale passa di fatto al perito**. Ciò può avvenire per due ragioni: il giudice **sottovaluta il suo**

[20] Per una dimostrazione di un *discriminator* e della *confidence*, v.il sito `http://isitporn.com/`. Il servizio valuta se una immagine ha contenuti pornografici o meno, e restituisce una percentuale: più è elevata, più è probabile che l'immagine sia oscena.

ruolo (ritenendo di non poter contestare le conclusioni del perito), oppure **sopravvaluta l'affidabilità della prova scientifica** (perché la scienza è vista come più affidabile di mezzi di prova tradizionali e più fallibili, come la testimonianza).[21]

Per evitare la degenerazione del processo in un dibattito fra esperti, è necessario definire i ruoli e le responsabilità delle varie parti.

6.5 Ruolo dei periti e consulenti tecnici

Il ruolo dei consulenti tecnici è di **dimostrare l'affidabilità** dei metodi di analisi che usano, e delle conclusioni a cui sono giunti.

6.5.1 Conoscenza delle *best practices*

L'informatica forense è caratterizzata dall'essere una disciplina in continuo aggiornamento. Per esigenze di certezza, riproducibilità ed interoperabilità è necessario definire protocolli operativi per gli accertamenti informatici e la documentazione delle prove.[22]

I C.T. devono essere a conoscenza e dimostrare di applicare le *best practices*, specialmente nel caso di **acquisizione** di dati digitali, data la delicatezza della procedura. Nel caso dell'**analisi** di dati sussiste un maggiore margine discrezionale, ed esistono numerosi strumenti di analisi.

In questo caso si può affermare che la *best practice* è l'uso degli strumenti standard per il settore (*industry standards*). Tuttavia, grazie alla sentenza Daubert è possibile ipotizzare anche l'uso di **strumenti nuovi e sviluppati ad-hoc** per risolvere uno specifico problema durante quell'investigazione, purché se ne dimostri la validità.

[21] Di Pinto, ≪La prova scientifica nel processo penale≫, p. 916.

[22] Brighi e Maioli, ≪Un cambio di paradigma nelle scienze forensi. Dall'armonizzazione tecnico-giuridica a una nuova cornice epistemologica≫, pp. 233–234.

6.5.2 Documentazione delle operazioni

I C.T. devono produrre una documentazione il più completa possibile, che deve includere:

- **Strumenti di analisi** – è utile indicare la versione esatta. Nel caso di programmi commerciali, dimostrare che si è titolari di una licenza valida.[23] Nel caso di programmi *open source*, è necessario indicare come sono stati scaricati;[24]
- **Operazioni svolte** ed **parametri utilizzati** – in una maniera sufficientemente **dettagliata** da permettere la loro **riproducibilità** solo leggendo la relazione;[25]
- **Risultati conseguiti** – è utile **indicare esattamente i risultati** che devono essere conseguiti al termine delle operazioni, in modo che non sussistano incertezze in merito alla corretta riproduzione dell'analisi;[26]
- **Riferimenti tecnici** e **spiegazioni** – le varie operazioni devono essere motivate, con riferimenti a pubblicazioni scientifiche, standard o linee guida. In caso di operazioni semplici e di **routine**, è sufficiente una motivazione sintetica. Nel caso di operazioni complesse e l'uso di strumenti **ad-hoc**, è necessario spiegare nei dettagli le modalità di funzionamento;

[23] Come già visto nella sezione 5.3, è possibile usare programmi senza licenza, ma i risultati ottenuti sono inaffidabili.

[24] È preferibile allegare sia il codice sorgente, il programma eseguibile, ed indicare l'autore.

[25] In altre parole, se è necessario chiedere chiarimenti al C.T. che l'ha scritta, non è completa. Pertanto, nel dubbio, è sempre meglio essere più specifici. È utile inserire *screenshot* nella relazione, ed accompagnarla con una registrazione dello schermo con commento del C.T., in cui si dimostrano visivamente le operazioni compiute.

[26] Nella relazione si inserisce una rappresentazione dei risultati in forma di *screenshot*, ed hash. È opportuno allegare anche una copia dei risultati così come ottenuti dal programma nella versione digitale originale, ed aggiungere brevi note esplicative riguardo al metodo di analisi utilizzato.

- **Conclusioni** – mentre le parti precedenti sono più puramente tecniche, il C.T. deve esporre le conclusioni in modo che siano **comprensibili anche ai non-tecnici** (avvocati e giudici). È opportuno iniziare con un breve riassunto della parte tecnica, spiegata in maniera **semplificata**. L'obiettivo è di dimostrare al giudice che c'è un **fondamento scientifico** alla base delle conclusioni.[27] Infine, devono procedere all'illustrazione dei risultati raggiunti, e della loro rilevanza.

6.5.3 Sensibilizzazione verso le problematiche della *digital evidence*

L'informatica forense presenta una peculiarità subdola. Il presupposto che ha portato alla informatizzazione della società è che la tecnologia è diventata accessibile all'utente medio. Non serve essere esperti di informatica per usare un computer, e c'è un notevole interesse a progettare l'interazione fra utente e macchina in modo che sia il più semplice ed intuitivo possibile.[28]

Se una tecnologia è accessibile, il numero di utenti a cui può essere venduta aumenta notevolmente. Di fatto, la **persona media ha familiarità con i dispositivi informatici**, ma **questa conoscenza è solo a livello superficiale**: l'**utente medio** sa "**come usare**" un dispositivo, ma **pochi utenti** (tecnici o appassionati) sanno effettivamente "**come funziona**" esattamente il dispositivo, a livello di software e hardware.

[27] Quello che è rilevante non sono tanto i dettagli tecnici – ad es., le basi matematiche della funzione di hash – quanto menzionare la funzione dello strumento, e dimostrare il nesso tra strumento e conclusioni – ad es., "la funzione di hash serve a valutare l'integrità dei file, si usano più funzioni di hash per evitare il rischio di collisioni, dato che i valori corrispondono il file non è cambiato".

[28] Spesso l'interesse è più economico che filantropico. I programmi *open source* sono utilizzabili gratuitamente, e spesso sono tanto potenti quanto gli strumenti commerciali, ma le interfacce utente tendono ad essere più spartane. D'altro canto, più un programma è semplice da imparare ad utilizzare, più è facile venderlo.

Le problematiche dei dati digitali[29] rimangono fuori dal dominio delle conoscenze che è ragionevole aspettarsi dalla persona media. La sempre crescente importanza dei dati digitali impone – almeno ai giuristi ed alle forze dell'ordine – di **prendere coscienza** di queste caratteristiche.

Il C.T. di informatica forense ha un "dovere etico" di sensibilizzare i giuristi, ed in particolare, gli deve essere concessa l'"**autorità**" di poter **criticare** le **impostazioni tecniche erronee** date della Cassazione.

Ad esempio, nella **sent. Cass. n. 40963/2017**, si perviene ad una **affermazione criticabil**e: per le acquisizioni di **minore complessit**à – l'esempio citato è il contenuto di un singolo file – **non è necessario usare le tecniche di copia forense** per raggiungere le finalità previste dal codice.[30]

La critica della Corte da un lato è giustificabile: se si deve acquisire **un singolo file**, la **copia forense sembra eccessiva**, perché richiede la copia dell'intero supporto, e quindi comporta un dispendio di tempo e denaro, e sacrifica la riservatezza del patrimonio informativo della persona. In altre parole, c'è una irragionevole sproporzione tra mezzi e fine.

Dall'altro lato, rimane il fatto che la copia forense è la modalità di acquisizione che fornisce le **migliori garanzie**. Mentre è possibile ipotizzare altri metodi di acquisizione di un file,[31] sono tutte "meno affidabili" della copia forense: meno dati si catturano, e meno possibilità si hanno di trovare

[29] Fatta salva forse solo la degradazione dei supporti materiali, e la necessità di fare copie di backup – tutti hanno avuto un disco rigato, o perso dati...

[30] Bartoli, ≪Sequestro di dati a fini probatori: soluzioni provvisorie a incomprensioni durature≫, pp. 9–11.

[31] Se si è unicamente interessati al contenuto di un unico file, è possibile collegare il supporto ad un write-blocker, individuare il file, ed eseguire l'acquisizione dei contenuti della cartella in cui si trova. Mentre questa operazione è più affidabile di un semplice copia ed incolla (*drag-and-drop*), rimane comunque il fatto che si cattura il semplice contenuto del file – che è più vicina ad un'apparenza – e non come il file è effettivamente conservato sul disco – la sua vera essenza – e quindi eventuali ulteriori elementi che riguardano quel file, ma non sono contenuti al suo interno (ad es., file di log che lo menzionano) sono persi.

correlazioni fra i dati.[32]

Il fatto che il codice non elenchi le modalità di acquisizione non può essere abusato per usare metodi più veloci e semplici, ma che compromettono gravemente le garanzie previste dal codice. Altrimenti, il giudice "si vedrebbe a quel punto costretto, più che alla valutazione d'un elemento, a un atto di fede".[33] Di conseguenza:

> Le SU "avrebbero fatto meglio a sottolineare che il contraddittorio sul piano tecnico resta un valore da garantire, qualunque sia lo strumento impiegato" [...] Se si volesse fare a meno dell'immagine forense, si dovrebbe chiedere agli investigatori di documentare le operazioni in maniera meticolosa, tramite videoripresa o sistemi di auditing [...] [e] di limitare i rischi rendendo tracciabile ogni passaggio; si potrebbe così valutare il preciso impatto delle manovre compiute sul sistema: la ripetibilità non sarà più assicurata, ma si potrebbe svogere un controllo successivo sulla sostenibilità scientifica del procedimento seguito.[34]

6.6 Ruolo dei giuristi

6.6.1 "Ignoranza legittima" del giudice

I giuristi devono prendere atto della fallibilità della scienza, né lasciarsi tentare dal delegare di fatto il potere giurisdizionale al perito e C.T. La sentenza Daubert ha introdotto e rafforzato il principio per cui il **giudice** ha il **potere e dovere** di **valutare liberamente** la **prova scientifica**.[35]

[32] Cfr. la sezione 3.3.4.

[33] Bartoli, ≪Sequestro di dati a fini probatori: soluzioni provvisorie a incomprensioni durature≫, p. 10.

[34] Ibid., p. 11.

[35] Brighi e Maioli, ≪Un cambio di paradigma nelle scienze forensi. Dall'armonizzazione tecnico-giuridica a una nuova cornice epistemologica≫, p. 232.

Lo stato di "**legittima ignoranza**" del giudice nelle materie tecniche non può giustificare un "acritico affidamento" al contributo del perito, o dei CC.TT.,[36] o lasciare che il processo si riduca il più possibile ad un dibattito fra esperti. Gli **ausiliari** del giudice e delle parti **non sono a conoscenza** di **tutti gli altri elementi** del processo.[37] Non hanno cognizione delle **altre questioni di fatto**, o di **problematiche giuridiche**, hanno solo cognizione della parte tecnica.

La sent. Cass. n. 43789/2010 delinea chiaramente il limite del ruolo degli esperti: devono **aiutare il giudice** nell'**esame** e nella **comprensione** degli **studi scientifici**, e degli **elementi tecnici**. Successivamente, il giudice è tenuto a dare **motivazione razionale** della sua decisione, in ogni caso, **anche se è d'accordo** con le conclusioni del perito.[38]

Idealmente, i giudici dovrebbero avere una conoscenza di base delle caratteristiche dei dati digitali, e delle principali problematiche, per evitare gli errori di valutazione più gravi. Non si chiede una conoscenza specialistica delle *best practices*, che sono in continuo aggiornamento, ma solo avere una cognizione solida delle basi fondamentali della disciplina, le fondamenta su cui si basa l'intero edificio, e del loro rilievo che hanno per il processo e per l'affidabilità della prova.

6.6.2 Valorizzazione del contraddittorio tecnico

In secondo luogo, se si ricorre allo strumento della prova scientifica, allora è necessario garantire la **difesa tecnica** ed il **contraddittorio tecnico**.

Per coerenza con il **diritto di difesa**, la nomina di un C.T., la raccolta di elementi tecnico-scientifici, la loro presentazione al giudice, l'interrogazione

[36] Di Pinto, «La prova scientifica nel processo penale», p. 920.
[37] Ibid., p. 918.
[38] Brighi e Maioli, «Un cambio di paradigma nelle scienze forensi. Dall'armonizzazione tecnico-giuridica a una nuova cornice epistemologica», pp. 232–233.

dei propri consulenti e la contro-interrogazione di quelli delle controparti deve essere concessa a tutte le parti.[39]

Nel caso in cui il giudice nomini un perito è utile ascoltare le osservazioni dei CC.TT. di parte riguardo l'**operato del perito**, per avere piena cognizione delle **prospettive di parte**. In particolare, il C.T. del difensore è tenuto solo a ricercare e valorizzare elementi che vadano a favore della parte.

Se c'è disaccordo fra le **conclusioni** dei CC.TT. il giudice può sopperire con il suo **libero apprezzamento**, basandosi anche su altri elementi non-tecnici. Ma se c'è disaccordo fra i **metodi di analisi** utilizzati, allora è necessario sopperire la **perizia**, in modo che il giudice possa **valutare la loro scientificità ed affidabilità**.[40]

Infine, il contraddittorio deve essere assicurato sia dal punto di vista della:

- **Quantità** – il contraddittorio deve sussistere sia nella fase di **acquisizione** (perché tendenzialmente irripetibile), che nella fase di **analisi** (data la possibilità di usare più modalità, anche innovative) della *digital evidence*;
- **Qualità** – l'oggetto del contraddittorio deve riguardare sia la *digital evidence* in sé, ma anche tutte le **vicende che la riguardano**.[41] Più un'operazione è **irripetibile**, **idonea ad alterare i dati** o di **difficile esecuzione**, e più minuziosamente deve essere **documentata**, in modo da garantire il più completo controllo sull'affidabilità della prova.

Ovviamente, questi sono obiettivi ideali. Nella pratica, la concreta implementazione del contraddittorio tecnico è vincolata a limiti materiali: tempo,

[39] Di Pinto, ≪La prova scientifica nel processo penale≫, pp. 922–925.

[40] Ibid., pp. 925–926.

[41] Ad es., si devono registrare accuratamente le operazioni di acquisizione, si deve redigere una catena di custodia che elenchi chi è il custode della *digital evidence*, e quando e quali operazioni sono state compiute con il supporto materialeecc. . .

denaro, effettiva difficoltà del caso concreto, presenza di altre prove non tecniche, ecc. . .

In conclusione, è necessario trovare un equilibrio. Se si decide di usare la *digital evidence*, si devono fornire quanti più elementi di valutazione possibile, nei limiti di quanto è ragionevole fare nel caso concreto. Al tempo stesso, non si può scendere al di sotto di una certa soglia di **garanzie minime**, perché a quel punto, è **meglio non acquisire proprio la prova**, che acquisire un elemento che di "scientifico" ha solo il nome.

Capitolo 7

Conclusioni

7.1 Natura *sui generis* dei dati digitali

Come esaminato nella prima parte,[1] i dati digitali rappresentano un elemento ***sui generis***. La caratteristica più rilevante è la loro **natura dematerializzata**, ed in alcune occasioni anche **(quasi) immateriale**[2].

Le **cose materiali** hanno una **corporalità**, e possono essere percepite dai **sensi**. L'unico elemento corporale dei **dati digitali** è la loro **base materiale** – che per di più è **irrilevante**, dato che gli stessi dati possono essere memorizzati su vari tipi di basi materiali, usando diversi metodi di incorporamento – e l'**osservazione** a occhio nudo della base materiale **non dà informazioni**.

Anche usando le tecniche di analisi che permettono di visualizzare i dati così come sono materialmente memorizzati, deve seguire un processo di decodifica dal sistema binario, che richiede l'applicazione del codice adatto al tipo di dati. Solo alla fine di questo percorso si arriva effettivamente alle informazioni contenute in quel supporto materiale.

[1] Cfr. sezioni 2 e 3.
[2] Cfr. sezione 4.7.

Quello che si vuole evidenziare è l'abisso che passa fra il dominio "analogico" e quello "digitale". Si deve operare sulla presunzione che le **considerazioni tradizionali** che valgono per le prove materiali **non possano essere trasferite al digitale**.

Qualsiasi tentativo di costruire argomentazioni che possano servire tanto per le prove materiali che per quelle digitali è come cercare di mischiare acqua ed olio – con abbastanza sforzo, sembra possibile ottenere una miscela omogenea, ma basta aspettare poco tempo, e le differenze riemergono naturalmente.

Pertanto, il diritto **deve prendere conoscenza** della natura peculiare, e assolutamente aliena ai canoni tradizionali, dei dati digitali. Ogni **ragionamento** deve **partire** dalla **caratteristiche del dato digitale**, e **poi valutarne le conseguenze giuridiche**. Fare altrimenti è:

- Nell'ipotesi **migliore**, **controproducente**, perché si va a **svilire** non solo la **natura**, ma anche le **potenzialità** del dato digitale;
- Nell'ipotesi **peggiore**, **paradossale**, perché il giurista cerca di **piegare la scienza e la realtà empirica** alle sue convinzioni.[3]

7.2 Fragilità e forza dei dati digitali

7.2.1 Doppia natura

I dati digitali presentano una doppia natura:

- Sono **fragili**, e facilmente **modificabili**;[4]

[3] Cfr. rispettivamente le sezioni 3.2.3 (concezione erronea degli accertamenti riguardanti supporti materiali come ripetibili) e 4.4.2 (problematiche relative alla differenza fra i dati digitali e la loro rappresentazione materiale).

[4] Cfr. le sezioni 3.1 (volatilità), 3.2 (deteriorabilità), 3.3 (modificabilità).

- Sono **riproducibili** e facilmente **analizzabili**.[5]

Deve essere dato il giusto peso ad entrambe le caratteristiche.

7.2.2 Fragilità

Per quanto riguarda la **fragilità**, nell'agire in concreto si devono scegliere istituti giuridici che permettano la **maggiore partecipazione delle parti**, e del loro C.T..

Si deve sempre partire dalla premessa che ogni operazione sui supporti materiali è **irripetibile**, ed agire di conseguenza. Anche la modifica di un solo bit può avere conseguenze disastrose.[6] Inoltre, si devono seguire le *best practices*: l'uso di scorciatoie e presunzioni male si adatta alla necessità di proteggere i dati digitali. In particolare, l'hash assume un ruolo fondamentale. Se l'hash non viene calcolato, è impossibile verificare l'integrità dei dati nel tempo.

Il **mancato calcolo dell'hash** è come la mancata apposizione della firma su una sentenza,[7] deve essere considerata un vizio insanabile. Data l'estrema facilità con cui è possibile modificare i dati digitali, quel file è assolutamente inaffidabile.

La **soluzione più garantista** consiste non solo nel calcolo dell'hash, ma nell'apposizione di una firma digitale e marca temporale al file, al fine di associare anche una **identità** ed una **data certa** all'hash, e diminuire il più possibile la possibilità di contestazioni.

Tuttavia, mentre l'esigenza di proteggere i dati è fondamentale, al tempo stesso bisogna **evitare di comprimere eccessivamente** i diritti della persona. Il sequestro deve essere il meno invasivo possibile, e quindi riguardare

[5] Cfr. le sezioni 3.4 (analisi), 3.5 (verifica dell'integrità), 3.6 (riproducibilità).
[6] Specie nel caso di date, cfr. la sezione 5.1.
[7] Cfr. l'art. 161 co. 2 c.p.c..

singoli file, e non deve avere una durata eccessiva. Una volta che la copia è stata eseguita, il supporto va restituito il prima possibile.[8]

7.2.3 Analizzabilità

Dal punto di vista dell'**analizzabilità**, la prova digitale presenta l'indubbio vantaggio di poter analizzare anche grandi quantità di dati (*in bulk*) ed in maniera completamente automatica, di indicizzare i file ed i loro contenuti, estrarre informazioni utili, e di **trovare correlazioni** non solo all'interno dei dati contenuti nel singolo supporto materiale, anche fra più supporti materiali, o anche più casi.

Questo aumenta l'**efficienza** delle investigazioni. Gli elementi più comuni[9] sono analizzati in automatico. Questo comporta che l'investigatore può trovare più facilmente gli elementi necessari, e può risparmiare tempo ed energie per la ricerca di elementi più difficili da individuare o analizzare, che sono sfuggiti all'analisi automatica, e devono essere gestiti manualmente.

Tuttavia, ciò non deve portare a pensare che la prova digitale sia quasi onnipotente. In ogni caso, rimangono sempre due limiti di fondo:

- **Dati sottoposti ad analisi** – i dati possono essere stati **modificati per cercare di sviare le indagini**.[10] L'analisi non sempre può dimostrare che un determinato dato è stato alterato: questa è una valutazione tecnica, e rimane sempre un **margine di errore**. Ancora, le tecniche di ***anti-forensics*** puntano proprio all'impedire di recuperare dati con l'analisi, che quindi potrebbe non fornire risultati utili.

[8] Cfr. la sezione 5.9.4.

[9] Ad es., la ricerca di mail e di indirizzi email, della cronologia del browser, di file che sono già conosciuti come sospetti (gli *hash databases* di file pedopornografici, cfr. 3.4), *file carving*...

[10] Ad es., cancellazione sicura, modifica dei metadati...

- **Risultati** – l'analisi ha un'**ottica limitata**, perché riguarda solo i dati digitali trovati su un supporto. Inoltre, in molti i casi il risultato dell'analisi non è netto, ma permane un **margine di incertezza**.[11] Pertanto, è necessario verificare in maniera critica le risultanze, e confrontarle con le altre prove.

Anche qui bisogna trovare un equilibrio. Il **rischio più elevato** è l'**eccessiva fiducia nei risultati delle analisi**, più che la sfiducia, perché c'è una concezione della prova digitale come maggiormente affidabile. Non ci si può fermare alle apparenze, ma i risultati delle analisi devono essere confrontati con le altre prove raccolte – altrimenti, si rischia la deriva tecnicista del processo.

7.3 Modalità di ricerca della prova informatica

Le innovazioni della L. n. 48/2008 introducono i principi della **conformità della copia** e della **conservazione dei dati**, ed estendono la definizione dei **mezzi di ricerca della prova** in modo da includere anche le **fonti di prova di natura digitale**. L'intenzione di imporre uno standard di protezione e trattamento minimo dei dati è lodevole, ma i limiti diventano evidenti quando si considera che i mezzi di ricerca della prova contenuti nel c.p.p. sono stati pensati per prove materiali, non per dati dematerializzati.

Ispezione, perquisizione e sequestro sono ben distinte con le cose: si guarda, si tocca e si cerca, si appone un sigillo e si mette da parte. Con i dati, le differenze sfumano. Ispezione e perquisizione tendono a confondersi, e se si cerca di distinguerle, spesso l'ispezione assume un ruolo talmente ridotto

[11] Ad es., si sospetta la manomissione di un file, ma non ci sono elementi sufficienti a dimostrarlo.

da sembrare vestigiale. Ed in ogni caso, il sequestro avviene prima degli altri istituti.[12]

La valutazione non è completamente negativa. Nel caso di sequestro di **corrispondenza** e **dati conservati** presso **fornitori di servizi**, la previsione di **meccanismi di collaborazione** e la nomina dei **fornitori come custodi dei dati** sono norme che dimostrano la **considerazione delle difficoltà tecniche**, e delle possibilità offerte dai dati digitali.

In ogni casi, sarebbe più opportuno sviluppare nuovi **strumenti di ricerca della prova**, **specifici per i dati digitali**, data la loro natura *sui generis*. Il rischio principale è di **limitare** il più possibile il ricorso a **modalità atipiche di ricerca della prova**, perché forniscono meno garanzie.

La **ricerca della *digital evidence*** non **comprime** il diritto di proprietà, ma il diritto alla **riservatezza**. Pertanto, la **tutela della privacy** deve essere il punto di riferimento principale, nei limiti di quanto possibile. La procedura penale comporta inevitabilmente una violazione della riservatezza, ma ciò non comporta che l'invasione della sfera privata può essere illimitata.

Si devono imporre **limiti ragionevoli** alla raccolta indiscriminata di dati, **indicare le ragioni** per cui si sta procedendo ad acqusizione di dati, garantire la possibilità del soggetto sottoposto alle indagini di **chiedere un riesame** delle operazioni, **sollecitare le analisi** se non vengono compiute in tempi ragionevoli, e **distruggere i dati inutilizzabili o non più necessari** il prima possibile.

Dal punto di vista delle infrastrutture, ispirandosi alla disciplina delle intercettazioni, sarebbe utile creare laboratori forensi e sistemi di *storage* per i dati raccolti, preferibilmente interni ai tribunali stessi, oppure di promuovere la collaborazione dell'A.G. con i privati che hanno già a disposizione

[12] Cfr. rispettivamente le sezioni 5.6 (distinzione fra ispezioni e perquisizioni) e 5.9.3 (distinzione fra perquisizioni e sequestri).

questo tipo di strutture, per la migliore gestione ed elaborazione della *digital evidence*.[13]

7.4 Documentazione e valutazione

La *digital evidence* è uno strumento potente, ma che va utilizzato con **responsabilità**.

Da un punto di vista tecnico, le sue caratteristiche di fragilità, ed esigenze di ripetibilità e completezza della valutazione, è assolutamente necessario fornire una documentazione chiara e completa riguardo a come è stata acquisita ed analizzata.[14]

Allo stesso tempo, i giuristi devono avere una conoscenza di base delle caratteristiche dei dati digitali, in modo da evitare sia gli errori di valutazione più grossolani, che la sopravvalutazione della prova scientifica. La conoscenza condivisa degli assiomi fondamentali della *digital forensics* aiuta tutte le parti processuali ad intendersi meglio a vicenda, rendendo le operazioni relative alla *digital evidence* più efficienti e certe.

Il valore più importante da garantire è quello del **contraddittorio fra le parti**, sia giuridico che tecnico. Di conseguenza, si deve **potenziare** l'istituto delle **indagini difensive** e valorizzare l'**uso** di **prove tecniche**. Più le parti hanno la possibilità di addurre elementi all'interno del processo, più elementi il giudice ha a disposizione per decidere.

7.5 Osservazioni finali

L'informatizzazione della società è un fenomeno relativamente recente. La giurisprudenza, presa alla sprovvista, ha cercato di adeguare gli istituti

[13] Cfr. la sezione 3.2.3.
[14] Cfr. la sezione 6.5.2.

tradizionali ad esigenze nuove, ma come si è visto, questo tipo di soluzione può essere solo temporanea.

Al momento attuale, è auspicabile cercare di sviluppare nuovi istituti tipici, modellati secondo le caratteristiche peculiari dei dati digitali. La loro adozione richiederà tempo, ma i futuri operatori giuridici saranno più familiari con le caratteristiche della tecnologia rispetto a quelli attuali.

Come già accennato, la tecnologia offre sia possibilità di creare prove false, che di analizzarle: si dovrà pertanto fare attenzione all'attendibilità delle prove, ma senza rendere il processo un dibattito fra esperti.

In ultimo, sarà necessario considerare come bilanciare la tutela della riservatezza e la protezione delle informazioni con le esigenze di accertamento penale: un bilanciamento che forse non troverà mai un equilibrio stabile, e che dovrà essere decisa caso per caso.

Bibliografia

Adler, David (2018). *Silk Road: The Dark Side of Cryptocurrency*. URL: https://news.law.fordham.edu/jcfl/2018/02/21/silk-road-the-dark-side-of-cryptocurrency/.

Adobe (2017). *Applicazioni Web*. URL: https://helpx.adobe.com/it/dreamweaver/using/web-applications.html.

aescripts + aeplugins (2018). *Datamosh Quick Start Tutorial*. URL: https://youtu.be/S1jIw4fufP8.

Alvarez, Paul (2004). «Using extended file information (EXIF) file headers in digital evidence analysis». In: *International Journal of Digital Evidence* 2.3. URL: https://www.utica.edu/academic/institutes/ecii/publications/articles/A0B1F944-FF4E-4788-E75541A7418DAE24.pdf.

Aruba.it (n.d.[a]). *Cos'è la Firma Digitale*. URL: https://www.pec.it/firma-digitale.aspx.

– (n.d.[b]). *Cos'è la Marca Temporale*. URL: https://www.pec.it/marche-temporali.aspx.

Aterno, Stefano (2014). «Digital forensics (investigazioni informatiche)». In: *Digesto delle Discipline Penalistiche, VIII Aggiornamento*. UTET Giuridica.

Bartoli, Laura (2016). «La catena di custodia del materiale informatico: soluzioni a confronto». In: *Anales de la faculdad de derecho – Universidad*

de La Laguna 33, pp. 145–162. URL: https://www.ull.es/revistas/index.php/derecho/article/view/86.

Bartoli, Laura (2018). «Sequestro di dati a fini probatori: soluzioni provvisorie a incomprensioni durature». In: *Archivio penale* 1. URL: http://www.archiviopenale.it/sequestro-di-dati-a-fini-probatori-soluzioni-provvisorie-a-incomprensioni-durature/articoli/15338.

Battiato, Sebastiano, Giovanni Maria Farinella e Giovanni Puglisi (2011). *Image/video forensics: casi di studio*. URL: https://www.academia.edu/2867895/Image_Video_Forensics_Casi_di_Studio.

Bitcoin Wiki contributors (n.d.). *Address*. URL: https://en.bitcoin.it/wiki/Address.

Brighi, Raffaella e Cesare Maioli (2015). «Un cambio di paradigma nelle scienze forensi. Dall'armonizzazione tecnico-giuridica a una nuova cornice epistemologica». In: *Informatica e diritto* 24.1-2, pp. 217–234. URL: http://www.ittig.cnr.it/risorse/attivita-editoriale/rivista/indici/1-2-2015/.

Busch, Jack (2011). *Friday Fun: Use Chrome to Create Fake Screen Captures*. URL: https://www.groovypost.com/howto/howto/friday-fun-chrome-create-fake-screen-captures/.

Callisch, David (2010). *Coping with Wi-Fi's biggest problem: interference*. URL: https://www.networkworld.com/article/2215287/coping-with-wi-fi-s-biggest-problem-interference.html.

Cavicchioli, Marco (2020). *Bitfinex: sequestro di 96 bitcoin. Una ulteriore prova del non anonimato di BTC*. URL: https://cryptonomist.ch/2020/01/28/bitfinex-sequestro-96-bitcoin/.

Consiglio d'Europa (n.d.). *Convenzione sulla criminalità informatica*. URL: https://www.coe.int/it/web/conventions/full-list/-/conventions/treaty/185.

Crompwell (2018). *Intro To Photogrammetry*. URL: https://youtu.be/3EENC9rFWhc.

Dan's Tools (n.d.). *Unix Permissions Calculator – Further Information*. URL: http://permissions-calculator.org/info/.

De Vivo, Maria Concetta e Giovanna Ricci (2012). «Diritto, crimini e tecnologie». In: *Informatica e diritto* 21.2, pp. 27–114. URL: http://www.ittig.cnr.it/risorse/attivita-editoriale/rivista/indici/2-2012/.

DeepFake Detector AIs Are Good Too! (2019). URL: https://youtu.be/RoGHVI-w9bE.

Di Pinto, Stefano (2018). «La prova scientifica nel processo penale». In: *Rivista di Polizia*, pp. 909–946. URL: https://www.associazionelaic.it/wp-content/uploads/2019/03/La-prova-scientifica-nel-processo-penale.pdf.

Disterer, Georg (2013). «ISO/IEC 27000, 27001 and 27002 for Information Security Management». In: *Journal of Information Security* 4.2, pp. 92–100. URL: http://dx.doi.org/10.4236/jis.2013.42011.

Ellis, Justin (n.d.). *Cable Interference*. URL: https://www.comms-express.com/infozone/article/cable-interference/.

Everybody Can Make Deepfakes Now! (2020). URL: https://youtu.be/mUfJOQKdtAk.

Ferrazzano, Michele (2014). «Indagini forensi in tema di scambio di file pedopornografici mediante software di file sharing a mezzo peer-to-peer». Tesi di dott. Università di Bologna. URL: http://amsdottorato.unibo.it/6697/.

ForensicsWiki contributors (n.d.[a]). *File Carving*. URL: https://forensicswiki.xyz/wiki/index.php?title=File_Carving.

– (n.d.[b]). *Raw Image Format*. URL: https://forensicswiki.xyz/wiki/index.php?title=Raw_Image_Format.

FourCC.org (2011). *What is a FOURCC?* URL: https://www.fourcc.org/fourcc.php.

Gammarota, Antonio (2016). «Informatica forense e processo penale: la prova digitale tra innovazione normativa e incertezze giurisprudenziali». Tesi di dott. Università di Bologna. URL: http://amsdottorato.unibo.it/7723/.

Green, Matthew (2012). *iCloud: Who holds the key?* URL: https://blog.cryptographyengineering.com/2012/04/05/icloud-who-holds-key/.

Gutmann, Peter (1996). «Secure Deletion of Data from Magnetic and Solid-State Memory». In: *Sixth USENIX Security Symposium Proceedings*. URL: http://softpres.org/cache/SecureDeletionOfDataFromMagneticAndSolidStateMemory.pdf.

Hollasch, Steve (2018). *IEEE Standard 754 Floating Point Numbers*. URL: https://steve.hollasch.net/cgindex/coding/ieeefloat.html.

Hopson (2017). *Secret Binary FIle [sic] Music - MSPaint.exe and aclui.dll*. URL: https://youtu.be/OfVERK_SreU.

International Organization for Standardization (ISO) (2015). *Information technology — Security techniques — Guidelines for the analysis and interpretation of digital evidence*. URL: https://www.iso.org/standard/44406.html.

– (2018). *Information technology — Security techniques — Information security management systems — Overview and vocabular*. URL: https://www.iso.org/standard/73906.html.

It's Getting Harder to Spot a Deep Fake Video (2018). URL: https://youtu.be/gLoI9hAX9dw.

Jacobi, Jon L. (2015). *Death and the unplugged SSD: How much you really need to worry about data retention*. URL: https://www.pcworld.com/article/2921590/.

Kernel.org Wiki contributors (n.d.). *ATA Secure Erase*. URL: https://ata.wiki.kernel.org/index.php/ATA_Secure_Erase.

Kissel, Richard et al. (2014). *Guidelines for media sanitization*. US Department of Commerce, National Institute of Standards e Technology. URL:

https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-88r1.pdf.

Krawetz, Neal (2007). *A picture's worth...* URL: http://hackerfactor.com/papers/bh-usa-07-krawetz-wp.pdf.

Lloyd-Davies, Frazer (2018). *The backup rule of three is a simple way to remember backup best practice.* URL: https://www.acronyms.co.uk/blog/backup-rule-of-three/.

Massone, Alessandro (2020). *La Procura di Roma ha bloccato l'accesso a Project Gutenberg, la più grande biblioteca di internet*. URL: https://thesubmarine.it/2020/05/25/procura-roma-bloccato-accesso-project-gutenberg/.

Mauchly, John William (1980). «The Eniac». In: *A History of Computing in the Twentieth Century*. Elsevier, pp. 541–550. URL: https://books.google.it/books?id=AsvSBQAAQBAJ&pg=PA546.

McHugh, Nat (2015). *Create your own MD5 collisions*. URL: https://natmchugh.blogspot.com/2015/02/create-your-own-md5-collisions.html.

Meghanathan, Natarajan, Sumanth Reddy Allam e Loretta A. Moore (2009). «Tools and techniques for network forensics». In: *International Journal of Network Security & Its Applications (IJNSA)* 1.1, pp. 14–25. URL: https://arxiv.org/abs/1004.0570.

Microsoft Docs Contributors (2017). *What is a driver?* URL: https://docs.microsoft.com/en-us/windows-hardware/drivers/gettingstarted/what-is-a-driver-.

Montgomery, Chris (2012). *24/192 Music Downloads... and why they make no sense*. URL: https://web.archive.org/web/20200426050432/https://people.xiph.org/~xiphmont/demo/neil-young.html.

Murgia, Severino (n.d.). «Prova informatica e processo penale». Tesi di dott. Università degli Studi di Pavia. URL: https://iris.unipv.it/

retrieve/handle/11571/1203330/184892/Severino%20Murgia%20-%20Prova%20informatica%20e%20processo%20penale.pdf.

Nostalgia Nerd (2018). *Magnets vs. Floppy Disks — Nostalgia Nerd*. URL: https://youtu.be/Nn2R7bIzDtc.

Ontrack.com (2015). *Recovering data from black boxes*. URL: https://www.ontrack.com/uk/blog/concepts-explained/recovering-data-black-boxes/.

PassMark Software (n.d.). *Booting a forensics image on a Virtual Machine*. URL: https://www.osforensics.com/faqs-and-tutorials/booting-image-virtual-machine.html.

Pennisi, Michele (2015). «Rilievi ed accertamenti di polizia giudiziaria. I problemi esegetici posti dalla normativa vigente e gli sviluppi dottrinali e giurisprudenziali in materia». Tesi di laurea mag. Università di Bologna. URL: http://www.giurisprudenzapenale.com/wp-content/uploads/2017/09/Tesi-Pennisi.pdf.

Pittiruti, Marco (2015). «Le indagini informatiche nel processo penale». Tesi di dott. Università degli Studi Roma Tre. URL: http://hdl.handle.net/2307/5026.

Preshing, Jeff (2011). *Hash Collision Probabilities*. URL: https://preshing.com/20110504/hash-collision-probabilities/.

ProStorage (2017). *LTO and LTFS: The Pros and Cons of Linear Tape-Open and Linear Tape File System*. URL: https://getprostorage.com/blog/lto-ltfs-archiving/.

Russinovich, Mark (2018). *SDelete v2.02*. URL: https://docs.microsoft.com/en-us/sysinternals/downloads/sdelete.

Sapio, Antonio (2016). «L'utilità della computer forensics nel processo penale: nuove tecnologie e prassi». Tesi di laurea mag. URL: http://tesi.luiss.it/17040/.

Schneier, Bruce (2018). *MD5 and SHA-1 Still Used in 2018*. URL: https://www.schneier.com/blog/archives/2018/12/md5_and_sha-1_s.html.

Shaaban, Ayman e Konstantin Sapronov (2016). *Practical Windows Forensics*. Packt Publishing.

Sobti, Rajeev e G. Geetha (2012). «Cryptographic hash functions: a review». In: *International Journal of Computer Science Issues (IJCSI)* 9.2, pp. 461–479.

Stack, Liam (2019). *Update Complete: U.S. Nuclear Weapons No Longer Need Floppy Disks*. URL: https://www.nytimes.com/2019/10/24/us/nuclear-weapons-floppy-disks.html.

Tanenbaum, Andrew S. e Todd Austin (2013). *Structured computer organization*. Pearson.

Technology Connections (2018). *Nyquist–Shannon; The Backbone of Digital Sound*. URL: https://youtu.be/pWjdWCePgvA.

The iPhone Wiki contributors (n.d.). *Brick*. URL: https://www.theiphonewiki.com/wiki/Brick.

ThioJoe (2018). *What if You SHAKE a Hard Drive WHILE It's Running?* URL: https://youtu.be/Z3LQX9V90Vo.

This AI Clones Your Voice After Listening for 5 Seconds (2019). URL: https://youtu.be/0sR1rU3gLzQ.

Tonnellotto, Maurizio (2014). «Evidenza informatica, computer forensics e best practices». In: *Rivista di Criminologia, Vittimologia e Sicurezza* 8.2, pp. 68–103. URL: http://www.vittimologia.it/rivista/articolo_tonellotto_2014-02.pdf.

Torre, Marco (2015). «Aspetti giuridici e tecnici relativi al trattamento della prova digitale nel processo penale. La prova informatica nella legge 18 marzo 2008, n. 48». In: *Informatica e diritto* 24.1-2, pp. 65–104. URL: http://www.ittig.cnr.it/risorse/attivita-editoriale/rivista/indici/1-2-2015/.

Treccani, Enciclopedia (n.d.). *Informatica*. URL: http://www.treccani.it/enciclopedia/informatica/.

United States Court of Appeals, District of Columbia Circuit. (1923). *Frye v. United States*. URL: https://en.wikisource.org/wiki/Frye_v._United_States/Opinion_of_the_Court.

United States Court of Appeals, Ninth Circuit (2006). *United States v. Romm*. URL: https://caselaw.findlaw.com/us-9th-circuit/1231820.html.

United States Supreme Court (1993). *Daubert v. Merrell Dow Pharmaceuticals*. URL: https://caselaw.findlaw.com/us-supreme-court/509/579.html.

Vaciago, Giuseppe (2011). «Digital evidence: profili tecnico-giuridiche e garanzie dell'imputato». Tesi di dott. Università degli Studi di Milano-Bicocca. URL: https://boa.unimib.it/handle/10281/20472.

Vandeven, Sally (2014). *Forensic Images: For Your Viewing Pleasure*. URL: https://www.sans.org/reading-room/whitepapers/forensics/paper/35447.

Vocabolario on line Treccani (n.d.). *Digitale*. URL: http://www.treccani.it/vocabolario/digitale2/.

What's a Generative Adversarial Network? Inventor Explains (2017). URL: https://blogs.nvidia.com/blog/2017/05/17/generative-adversarial-networks/.

Wikipedia contributors (n.d.[a]). *Click of death*. URL: https://en.wikipedia.org/wiki/Click_of_death.

– (n.d.[b]). *Commodore Datasette*. URL: https://en.wikipedia.org/wiki/Commodore_Datasette.

– (n.d.[c]). *Decimal computer*. URL: https://en.wikipedia.org/wiki/Decimal_computer.

– (n.d.[d]). *Firmware*. URL: https://en.wikipedia.org/wiki/Firmware.

– (n.d.[e]). *List of file signatures*. URL: https://en.wikipedia.org/wiki/List_of_file_signatures.

– (n.d.[f]). *Mojibake*. URL: https://en.wikipedia.org/wiki/Mojibake.

Williams, Matt (2017). *Safeguarding Data: How Ransomware Can Affect the Master Boot Record (MBR)*. URL: https://www.faronics.com/news/blog/safeguarding-data-ransomware-can-affect-master-boot-record-mbr.

Ziccardi, Giovanni (2006). ≪Scienze forensi e tecnologie informatiche: la computer and network forensics≫. In: *Informatica e diritto* 25.2, pp. 103–125. URL: http://www.ittig.cnr.it/risorse/attivita-editoriale/rivista/indici/2-2006/.