

I SOFTWARE PER LA RICERCA ED ELABORAZIONE DELLE FONTI INFORMATICHE*

Donatello Cimadomo



Esigenze di ordine sistematico impongono di distinguere, per quanto ora d'interesse, i *software* per la **ricerca** delle fonti informatiche dai *software* per l'**elaborazione** delle fonti informatiche; e di precisare che per **fonte informatica** s'intende la fonte da cui trarre un elemento elettronico, che può essere un documento elettronico o un sistema informatico da cui estrapolare delle informazioni elettroniche.

Esula in parte da questo discorso un particolare tipo di software come il *trojan*, perché oltre ad essere utilizzato come software per la ricerca di fonti elettroniche (i.e. ispezione, perquisizione online e sequestro informatico) può essere impiegato – come comunemente accade – per le intercettazioni di comunicazioni.

Rappresenta, quindi, un software “multifunzione” o “multiuso”, avendo una natura prismatica.

Le legislazioni nazionali spesso non includono riferimenti diretti ed espliciti alle fonti di prova digitale e prova informatica, né tantomeno propongono una loro definizione specifica ed esclusiva. Ciò perché è complesso dare un inquadramento concettuale.

Ora, la nostra attenzione deve concentrarsi sulla terminologia utilizzata, perché una cosa è la ricerca ed elaborazione **delle** fonti informatiche, altro è la ricerca ed elaborazione di informazioni **nelle** fonti informatiche.

E infatti, una cosa è lo strumento fisico, materiale che esiste quale contenitore di dati elettronici come può essere un *server* o un sistema informatico qualsiasi (dunque, ricerca ed elaborazione *nella* fonte informatica); altro è invece ricercare una fonte digitale (quale sottoinsieme della fonte informatica) che incorpora la rappresentazione di un fatto o un atto (ricerca ed elaborazione *della* fonte informatica, o più specificamente, della fonte digitale) – banalmente, un file che incorpora in sé una serie di bit che restituiscono, letti insieme, una informazione, un elemento, un dato, un fatto, un atto.

Ritorna così la funzione prismatica del *trojan*, che oltre ad essere strumento di intercettazione, diventa anche strumento per la ricerca **nella** fonte di prova (inoculato nel sistema informatico) e **della** fonte di prova, perché utilizzato per cercare un file, quale fonte immateriale, ritenuto utile ai fini investigativi.

Possiamo dire che quando parliamo di software per la ricerca e l’elaborazione di fonti elettroniche ci riferiamo a strumenti utili per la ricerca ed elaborazione delle prove digitali e che ben potrebbero prescindere dal concetto di strumento impiegato *nella* fonte di prova. Si pensi, ad esempio, all’impiego di software per la ricerca ed elaborazione delle prove digitali (anche conservate nei cloud) ricevute da un *internet service provider* mediante lo strumento dell’ordine europeo di produzione o con gli strumenti previsti dal Secondo protocollo alla Convenzione di Budapest o con il sistema, direi poco garantista, della *voluntary disclosure* di dati elettronici.

Questa sistemazione risente anche delle definizioni di fonte di prova informatica e digitale che si sono avute nel tempo:

per prove digitali debbano intendersi «quelle fonti di prova memorizzate in qualsiasi tipo di strumento

informatico come ad esempio gli *smartphone*, palmari, telefoni cellulari, server aziendali, postazioni di lavoro dei dipendenti, o altri sistemi cosiddetti informatici o telematici complessi (es. Cloud Computing)» (ATERNO).

per fonti di prova digitale si intendono le «informazioni generate, memorizzate o trasmesse mediante dispositivi elettronici che possono essere utilizzate in giudizio» (Consiglio d'Europa, *Guida alla prova digitale*).

È bene chiarire la netta distinzione tra documenti informatici-*fonte* (il file, inteso come bene o cosa immateriale, la successione di bit che, nell'insieme restituiscono una informazione elettronica) e sistema informatico-*fonte*, laddove il sistema informatico consente l'archiviazione o l'elaborazione di dati elettronici.

La giurisprudenza definisce il sistema informatico come un complesso di apparecchiature destinate a compiere una qualsiasi funzione utile all'uomo attraverso l'utilizzazione anche parziale di tecnologie informatiche che sono caratterizzate, per mezzo di una attività di "codificazione" e "decodificazione", dalla "registrazione" o "memorizzazione" tramite impulsi elettronici, su supporti adeguati, di "dati", cioè, di rappresentazioni elementari di un fatto, effettuata attraverso simboli (*bit*) in combinazioni diverse, e dalla elaborazione automatica di tali dati, in modo da generare informazioni costituite da un insieme più o meno vasto di informazioni organizzate secondo una logica che consente loro di esprimere un particolare significato per l'utente. Più in generale, un dispositivo elettronico assurge al rango di sistema informatico o telematico se si caratterizza per l'installazione di un *software* che ne sovrintende il funzionamento, per la capacità di utilizzare periferiche o dispositivi esterni, per l'interconnessione con altri apparecchi e per la molteplicità dei dati oggetto di trattamento (Cfr., Cass., SS.UU., 24 aprile 2015, n. 17325).

Ricorrendo ad un'immagine, il sistema informatico è l'armadietto da cui estraiamo il faldone di carte da studiare e analizzare.

La nostra attenzione quindi si concentra sui *software* che ricercano e elaborano i c.d. *file*.

I primi, i software di ricerca, sono programmi destinati ad essere impiegati come *mezzi, strumenti, per la ricerca della prova informatica o digitale*; sinteticamente, il *thema* è quello della ricerca investigativa che implica la necessità dell'analisi del dato rilevante, comunemente tratto da fonti aperte (non il *deep* ed il *dark web*), per la verifica, ad esempio, delle relazioni e dei collegamenti esistenti nel mondo reale fra persone, gruppi di persone aziende, organizzazioni e siti web; ma, anche, analizzare dei domini Internet, e degli indirizzi IP.

I software di elaborazione, invece, sono software che elaborano le informazioni elettroniche restituendo all'utente una informazione "composta", frutto dell'analisi che lo strumento (il programma informatico) ha compiuto sulla massa di dati/informazioni inseriti. Può giovare di algoritmi e basarsi su varie forme di intelligenza artificiale.

Tali software si suddividono in software in funzione predittiva (es. GIOVE della Polizia di Stato) e software di analisi in funzione investigativo-probatoria (es. SARI, acronimo di Sistema Automatico di Riconoscimento Immagini).

Altri, ancillari profili sono quelli connessi alle c.d. *machine evidence* ossia alle prove elaborate da software installati su auto a guida automatizzata e che vengono elaborate in caso di incidenti, ed all'analisi in funzione di supporto alla decisione del giudice (i.e. software utilizzati in fase decisoria).

In particolare, gli algoritmi sono utilizzati in funzione di supporto alla valutazione della prova, e si fondano sulle reti bayesiane e reti neurali. Servono per ricostruire il passato attraverso una valutazione razionale.

Sono stati sviluppati anche algoritmi in funzione predittiva, da un lato in grado di formulare *giudizi prognostici sulla pericolosità (risk assessment tools)*; dall'altro di predire il contenuto di una decisione.

Nel caso di software di elaborazione delle fonti a supporto della decisione di un giudice, è indispensabile chiarire che, se un impiego deve esserne fatto, non può che essere solo ai fini di ausilio e non in un'ottica di sostituzione dell'umano. Umanità della decisione penale, ragionevole dubbio e funzione reale della motivazione sono pilastri fondamentali.

Si tratta, quindi, di software che dovrebbero aiutare il giudice ad interrogarsi, a ragionare, a migliorare la sua attività decisoria; punto di partenza è lo sviluppo di uno schema logico condiviso da utilizzare, di regole da impiegare, di sbarramenti inferenziali da porre al fine di evitare le fallacie logico-probabilistiche.

Va, infine, segnalata la Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (Legge sull'intelligenza artificiale).

Articolo 54

Ulteriore trattamento dei dati personali per lo sviluppo nello spazio di sperimentazione normativa per l'IA di determinati sistemi di IA nell'interesse pubblico

1. Nello spazio di sperimentazione normativa per l'IA i dati personali legalmente raccolti per altre finalità sono trattati ai fini dello sviluppo e delle prove nello spazio di sperimentazione di determinati sistemi di IA innovativi alle seguenti condizioni:

a) i sistemi di IA innovativi sono sviluppati per salvaguardare un interesse pubblico rilevante in uno o più dei seguenti settori:

i) la prevenzione, l'indagine, l'accertamento e il perseguimento di reati o l'esecuzione di sanzioni penali, incluse la salvaguardia contro le minacce alla sicurezza pubblica e la prevenzione delle stesse, sotto il controllo e la responsabilità delle autorità competenti. Il trattamento si basa sul diritto degli Stati membri o dell'Unione.

** Contributo tratto dalla relazione svolta nell'ambito del convegno "Per uno statuto dei nuovi mezzi di ricerca della prova di fronte alla società digitale", tenutosi in Roma in data 22.9.2023.*