Nuove questioni di informatica forense
ISBN 979-12-5994-643-0
DOI 10.53136/979125994643018
pp. 343-361 (gennaio 2022)

L'IA applicata all'analisi dei metadati: un'alternativa alla rottura della crittografia per le autorità di contrasto alla criminalità*

Isadora Neroni Rezende[1], Pier Giorgio Chiara[2]

1. Introduzione: la natura prismatica delle esigenze di sicurezza nell'era dell'Internet delle cose

Questo elaborato esamina l'impatto, in termini di diritti fondamentali, di alcune tecnologie crittografiche di sicurezza digitale e analisi dei metadati nel contesto delle attività di informatica forense attuate dalle autorità di contrasto. Lo studio della disciplina normativa – unitamente ad un'analisi più tecnica – mira a chiarire se l'impiego di modelli di intelligenza artificiale (IA) nell'analisi del traffico telematico criptato possa essere d'aiuto per ridurre l'invasività delle operazioni di sorveglianza delle forze dell'ordine.

L'analisi proposta assume particolare rilevanza alla luce della progressiva diffusione dei dispositivi dell'Internet delle cose – meglio noto nell'acronimo inglese "IoT", ovvero *Internet of Things* – sempre più incorporati in ambienti definiti "intelligenti" (ad esempio, le cd. *smart cities*). L'IoT è, in via di prima approssimazione, una complessa

* Il presente lavoro è stato discusso dagli autori alla conferenza internazionale *TILTing Perspectives 2021*, organizzata dall'Università di Tilburg. Questo lavoro è il risultato di una ricerca comune e condivisa condotta da entrambi gli Autori. Tuttavia, nel dettaglio, Pier Giorgio Chiara è l'autore dei §§ 1, 2 e 3, mentre Isadora Neroni Rezende è l'autrice dei §§ 4, 5, 6.

Questo studio è finanziato dal programma di ricerca e innovazione Horizon 2020 dell'Unione Europea attraverso l'accordo di sovvenzione Marie Sklodowska-Curie ITN EJD "Law, Science and Technology Rights of Internet of Everything" No 814177.

[1] Dottoranda di ricerca in Diritto delle Nuove Tecnologie presso l'Università Autonoma di Barcellona | Last-JD-RIoE.

[2] Dottorando di ricerca in Diritto delle Nuove Tecnologie presso l'Università del Lussemburgo | Last-JD-RIoE.

rete di cose, che incorpora e – grazie ad Internet – connette tra loro sensori, attuatori, software di intelligenza artificiale (IA) e componenti hardware[3]. Questa enorme classe di componenti dei sistemi IoT è molto meno standardizzata della componentistica per i PC[4]. Infatti, l'IoT riunisce sotto di sé dispositivi con CPU, memoria e potenza di elaborazione limitata, come i sensori di pressione che, per essere economicamente competitivi, richiedono un hardware piccolo e a basso costo[5], e dispositivi con processori potenti e grande memoria. Ne consegue che la prima classe di dispositivi non ha la necessaria memoria e potenza di calcolo per eseguire i più aggiornati e robusti protocolli crittografici; così, la ricerca si concentra sui cd. algoritmi crittografici *lightweight*, più adatti a tali contesti "limitati".

Questi fanno sempre più affidamento su metodi crittografici per garantire i diritti fondamentali degli individui, in particolare il diritto alla *privacy* e alla protezione dei dati personali. La dottrina dedica una grande attenzione alle sfide normative (giuridiche, etiche e sociali) portate dall'IoT nei campi della *cybersicurezza*, della *privacy* e della protezione dei dati[6]. Recenti studi hanno infatti dimostrato che il paradigma dell'IoT pone seri e diversi rischi al godimento da parte degli individui dei diritti fondamentali alla *privacy* e alla protezione dei dati, indipendentemente dal fatto che siano in vigore protocolli di crittografia, siano questi "leggeri" o più tradizionalmente robusti. Queste preoccupazioni sono vieppiù amplificate nel contesto delle attività di contrasto da parte delle forze dell'ordine, poiché gli strumenti digitali forniscono alle autorità pubbliche mezzi sempre più intrusivi di monitoraggio, aventi tanto più efficacia quanto più gli ambienti in cui operano sono informatizzati. Diversamente dalle relazioni orizzontali intrattenute tra attori privati, in ambito civile, infatti, i cittadini soffrono più intensamente le asimmetrie di potere presenti nelle loro relazioni verticali con lo Stato.

[3] A. RAYES E S. SALAM, *Internet of Things: from Hype to Reality*, Springer, 2019, p. 2.

[4] P. MARWEDEL, *Embedded System Design: Embedded Systems Foundations of Cyber-Physical Systems, and the Internet of Things,* Springer, 2018, p. 126.

[5] T. STAPKO, *Practical Embedded security: Building Secure Resource-Constrained Systems*, Newnes, 2008, p. 85.

[6] Si veda *ex multis* U. PAGALLO, M. DURANTE E S. MONTELEONE, *What is new with the Internet of Things in Privacy and Data protection? Four legal challenges on sharing and control in IoT* in R. Leenes, R. Van Brakel, S. Gutwirth e P. De Hert (a cura di) *Data Protection and Privacy: (In)visibilities and Infrastructures*, Springer, 2017.

L'importanza del tema è altresì evidenziata dalla recente pubblicazione, da parte del Consiglio dell'UE, di una risoluzione intitolata "Crittografia - Sicurezza attraverso la crittografia e sicurezza nonostante la crittografia". La risoluzione propone la creazione di un quadro di certezze giuridiche volto a garantire la capacità delle autorità di *intelligence* – in ambito di sorveglianza preventiva – e della giustizia penale – con particolare riguardo al campo dell'informatica forense – di accedere legalmente ai dati digitali criptati e in modo mirato. Tuttavia, nel contesto delle attività di sorveglianza preventiva e di indagine, la crittografia *end-to-end (E2E)* rende l'analisi del contenuto delle comunicazioni estremamente difficile o praticamente impossibile, senza il ricorso alla cd. "rottura" della tecnologia crittografica.

Nel contesto delineato, due diversi livelli di complessità sembrano emergere. Essi riguardano: (i) la possibilità di raggiungere un equilibrio tra i valori protetti dalla crittografia, soprattutto quella E2E, e gli obiettivi delle autorità di contrasto, con particolare riguardo al cd. *legal hacking*[7]; e, (ii) la capacità di alcuni modelli di IA di preservare i vantaggi della crittografia, permettendo di dedurre informazioni preziose dal traffico delle comunicazioni, anziché dal contenuto vero e proprio delle stesse.

Dopo aver fatto luce sui rischi paradossalmente connessi all'impiego di tecnologie crittografiche, la risposta agli interrogativi posti prenderà le mosse da una disamina tecnico-giuridica dei mezzi di IA applicati all'analisi del traffico delle comunicazioni. Nell'esporre le potenzialità offerte da tali tecniche, verranno evidenziate nuove prospettive nel concepire la distinzione tra contenuto e dati esterni alle comunicazioni (o metadati). In particolare, si sosterrà come l'analisi dei metadati offra maggiori possibilità di bilanciamento tra i diritti alla *privacy* e alla protezione dei dati, da un lato, e le esigenze di sicurezza, dall'altro. Le argomentazioni proposte saranno infine sostenute dalla più recente giurisprudenza europea in materia di sorveglianza.

[7] Si veda G. ZICCARDI, *Parlamento Europeo, captatore e attività di hacking delle Forze dell'Ordine: alcune riflessioni informatico-giuridiche*, in *Arch. pen.*, 2017 (1).

2. I rischi della crittografia

La crittografia ricomprende sotto di sé diverse tecnologie cruciali per proteggere e promuovere diritti fondamentali come il diritto alla *privacy*, alla protezione dei dati personali e alla libertà di espressione[8]. D'altra parte, una tecnologia crittografica particolarmente "robusta" può altresì ostacolare le attività di contrasto da parte delle forze dell'ordine[9].

Diversi elementi sul piano normativo (giuridico, etico, sociale), mettono in luce il conflitto di valori sotteso alle tecniche di rottura della crittografia. Nel 2020, la cd. "crypto-guerra"[10] è stata rivitalizzata da una relazione congiunta dei cosiddetti "cinque occhi", cioè Stati Uniti, Regno Unito, Australia, Canada e Nuova Zelanda, attraverso la quale si è esplicitato che la sicurezza pubblica non possa essere protetta senza compromettere la *privacy* o la sicurezza informatica[11]. Dal canto suo, il Consiglio dell'UE ha pubblicato nel novembre 2020 una risoluzione volta a garantire la capacità delle autorità competenti nel settore della sicurezza e della giustizia penale, come le forze dell'ordine e le autorità giudiziarie, di accedere ai dati in modo legale e mirato[12].

In questo contesto, è interessante notare che il considerando 54 della proposta di revisione della direttiva sulla sicurezza delle reti e dell'informazione (di seguito, NIS 2) afferma che «l'uso della crittografia E2E dovrebbe essere "conciliato" con i poteri degli Stati membri di garantire la tutela della sicurezza pubblica e dei loro interessi essenziali in materia di sicurezza, nonché di consentire l'indagine, l'accertamento e il perseguimento di reati nel rispetto del diritto dell'Unione»[13]. La

[8] W. WIEWIÓROWSKI, *Keynote: Data protection needs encryption*, EDPS Online IPEN Workshop, 3 giugno 2020.

[9] Si veda R. BRIGHI, *Requisiti tecnici, potenzialità e limiti del captatore informatico. Analisi sul piano informatico-forense,* in Revisioni normative in tema di intercettazioni. Riservatezza, garanzie difensive e nuove tecnologie informatiche, Torino, Giappichelli, 2021, pp. 231 – 256.

[10] B. J. KOOPS, E. KOSTA, *Looking for some light through the lens of "cryptowars" history: policy options for law enforcement authorities against "going dark"*, in *CLSR*, 2018, 34.

[11] Si veda: https://www.gov.uk/government/publications/international-statement-end-to-end-encryption-and-public-safety/international-statement-end-to-end-encryption-and-public-safety-accessible-version.

[12] CONSIGLIO DELL'UNIONE EUROPEA, *La sicurezza attraverso la crittografia e nonostante la crittografia,* 13084/1/20, 2020.

[13] COMMISSIONE EUROPEA, *Proposta di direttiva del Parlamento Europeo e del Consiglio relativa a misure per un livello comune elevato di cibersicurezza nell'Unione, che abroga la direttiva (UE) 2016/1148,* COM(2020) 823 final, considerando 54.

formulazione del considerando è piuttosto problematica perché queste "soluzioni", volte a facilitare l'accesso da parte delle forze dell'ordine, si tradurrebbero molto probabilmente in un abbassamento della sicurezza complessiva degli utenti di tali sistemi a causa di una superficie di attacco[14] più largamente estesa. A tal proposito, il Garante Europeo della Protezione dei Dati (EDPS), nel suo parere sulla proposta per la NIS 2, suggerisce di modificare il considerando 54 affinché nulla nella proposta possa essere interpretato come un'approvazione dell'indebolimento della cifratura *end-to-end* attraverso *backdoors*[15].

La dichiarazione del Garante europeo deve essere letta insieme alla suddetta risoluzione del Consiglio. Il Consiglio sottolinea la necessità di soluzioni tecniche che consentano alle autorità competenti di ottenere un accesso legittimo e mirato ai dati cifrati, nel rispetto dei principi di legalità, trasparenza, necessità e proporzionalità. Eppure, il Consiglio, nel tentativo di contemperare tutte le istanze in gioco, esclude esplicitamente qualsiasi misura che possa indebolire la crittografia, come le *backdoors*. Secondo diversi esperti di sicurezza informatica, firmatari di una lettera aperta in risposta alla risoluzione, il Consiglio andrebbe cercando una "via di mezzo" che semplicemente non esiste[16]. Allo stato dell'arte, è chimerico pensare a tecnologie informatiche ragionevolmente sicure che, da una parte, facilitino i poteri investigativi per decriptare contenuti e, dall'altra, non vedano decrescere le loro aspettative di sicurezza, garantendo pertanto i diritti fondamentali degli individui – in particolare, il diritto alla *privacy* e alla protezione dei dati personali. Una volta rotto il meccanismo crittografico, anche se solo relativamente alle forze dell'ordine o agenzie di intelligence, l'aspettativa di sicurezza per ogni utente inevitabilmente diminuisce.

Poiché l'indebolimento della crittografia E2E – tramite creazione di *backdoors* – si è rivelato non essere un'opzione, conseguenza inevitabile di quella che in dottrina è stata chiamata come "seconda cripto-

[14] Il Glossario del NIST definisce la superficie d'attacco di un sistema, un elemento del sistema o di un ambiente come l'insieme dei punti sul perimetro in cui un aggressore può cercare di entrare, causare un danno o estrarre dati da quel sistema, elemento del sistema o ambiente.

[15] EDPS, *Opinion 5/2021 on the Cybersecurity Strategy and the NIS 2.0 Directive*, 2021, p. 16. Il Glossario del NIST definisce la *backdoor* come un modo non documentato di accedere al sistema informatico.

[16] Si veda: https://sites.google.com/view/scientists4crypto/start.

guerra"– è stata la creazione di cornici legislative relative ai poteri di *hacking* degli Stati[17]. In questo contesto, Europol ha annunciato alla fine del 2020 una piattaforma di decrittazione innovativa, sviluppata con il Centro di Ricerca della Commissione europea, che mira ad aumentare significativamente le capacità dell'agenzia europea e delle forze dell'ordine nazionali di decrittare le informazioni legalmente ottenute nelle indagini penali[18].

Per rendere le cose più complesse, alla fine di marzo 2021, Google ha unilateralmente fermato un cyberattacco – che sfruttava vulnerabilità cd. "*zero-days*"[19]– condotto da un governo occidentale "alleato degli USA" a fini antiterroristici, sollevando inevitabilmente questioni politiche, etiche e legali intorno alle pratiche di *hacking* di Stato[20] – non per questo necessariamente lecite[21].

La crittografia non deve essere considerata una tecnologia esente da "costi": paradossalmente, tali tecnologie di sicurezza digitale generano dei rischi per i beni che si prefiggono di tutelare, ossia la sicurezza e i diritti fondamentali degli individui, in quanto i governi spesso rivendicano, in un'ottica di *trade-offs*, la supremazia della sicurezza sulle libertà.

3. L'utilizzo dell'AI nell'analisi del traffico criptato: nuove frontiere per l'analisi dei metadati

Una valida alternativa alla decriptazione dei contenuti comunicativi è l'analisi dei cd. "dati esterni al traffico" criptato, cioè l'analisi dei metadati, alimentata da algoritmi di IA. Tali tecniche preservano

[17] B. J. KOOPS, E. KOSTA, *op. cit.*, pp. 898-899.

[18] EUROPOL, *Europol and the European Commission inaugurate a new decryption platform to tackle the challenge of encrypted material for law enforcement investigations*, 18 dicembre 2020.

[19] Trattasi di una vulnerabilità nel sistema non ancora di dominio pubblico e di conseguenza utilizzabile senza alcun tipo di blocco da parte di chi ne è a conoscenza almeno sino a quando il produttore non introdurrà una correzione (*patch*).

[20] P. O'NEILL, *Google's top security teams unilaterally shut down a counterterrorism operation,* in *MIT Technology Review*, 2021.

[21] Si veda lo scandalo "Pegasus" emerso, dal luglio 2021, grazie a diverse inchieste del *The Guardian*: https://www.theguardian.com/news/series/pegasus-project.

l'integrità della crittografia, anche della tecnologia E2E, in quanto non ricorrono alla "rottura" del meccanismo al fine di vedere il contenuto della comunicazione. In primo luogo, da un punto di vista tecnico, questa sezione mostrerà fino a che punto tali pratiche permettano di trarre inferenze di valore e quindi ottenere profili accurati degli individui. Poi, si analizzerà il quadro giuridico europeo, esistente e futuro, applicabile al trattamento dei metadati.

Mentre i dati di contenuto presentano elementi di ambiguità e interpretazione soggettiva, i metadati di contesto sono – sotto certi aspetti – più significativi e oggettivi, dal momento che riguardano l'ora, la data, la fonte, la destinazione e i campi di lunghezza della comunicazione consegnata dal dispositivo; in altre parole, «definiscono le informazioni di contesto nel quale l'oggetto informazionale si è formato, è stato condiviso, utilizzato (e.g., geolocalizzazione, *micro-tagging*)»[22]. Alcuni autori hanno però sostenuto come la tradizionale distinzione tra dati e metadati, metaforicamente rappresentata dal binomio *contenuto – busta*, non sarebbe più efficace per un duplice ordine di motivi. In primo luogo, potrebbe suggerire surrettiziamente l'ipotesi che il contenuto della comunicazione sia più sensibile dei metadati; e, in secondo luogo, i confini tra contenuto e metadati, nella comunicazione in rete, sono sempre più sfumati[23]. Un approccio relativistico per distinguere contenuto e metadati considera il contesto specifico in cui l'informazione si è formata o viene utilizzata: ad esempio, un URL è una "istruzione di consegna", cioè *metadato*, e allo stesso tempo *dato*, in quanto significa essenzialmente inviare un messaggio che dice "per favore rimandami la pagina trovata a questo URL"[24].

Dal punto di vista giuridico, con specifico riferimento alla materia della *privacy* e protezione dei dati, è pacifico che i metadati siano essenzialmente dati personali e, pertanto, che il loro trattamento non sia meno delicato dei dati di contenuto. Il cd. "test di identificabilità" stabilito dal considerando 26 del GDPR, da leggere in combinato dispo-

[22] M. Palmirani, M. Martoni, *Big data, governance dei dati e nuove vulnerabilità*, in XXXV POLITEIA, 2019, 136, p. 10.

[23] L. Mitrou, *Communications Data Retention: A Pandora's Box for Rights and Liberties?* in A. Acquisti et al. (a cura di) *Digital privacy: Theory, technologies, and practices*, Auerbach Publications, 2008, p. 422. Cfr. con Opinione dell'AG Bobek, *Rīgas satiksme*, C-13/16, §95.

[24] C. Conley, *Metadata: Piecing Together a Privacy Solution,* ACLU report, 2014, pp. 4-5.

sto con l'Art. 4 del Regolamento, stabilisce infatti che «si debba tener conto di tutti i mezzi di cui il titolare del trattamento o un terzo può ragionevolmente avvalersi per identificare una persona fisica, direttamente o indirettamente». Inoltre, il medesimo considerando precisa che «per accertare la ragionevole probabilità di utilizzo dei mezzi per identificare la persona fisica, si dovrebbe prendere in considerazione l'insieme dei fattori obiettivi, tra cui i costi e il tempo necessario per l'identificazione, tenendo conto sia delle tecnologie disponibili al momento del trattamento, sia degli sviluppi tecnologici»[25]. Il test sviluppato dal considerando 26 del GDPR abbraccia essenzialmente un approccio basato sul rischio per determinare se i dati sono personali o meno[26].

Applicando tali criteri al caso dei metadati, il concetto di *fingerprinting* aiuta a fare luce, da un punto di vista tecnico, sul carattere personale nonché estremamente sensibile di tali particolari trattamenti. Osservando, attraverso i metadati, alcune proprietà dei dati criptati, è possibile creare delle sequenze di dati che attribuiscano queste proprietà ai *file* o ai siti *web* e financo dispositivi corrispondenti[27]. Mentre il *website fingerprinting* mira a identificare su quale sito *web*, o parte di esso, un utente stia navigando, analizzando il solo traffico criptato dell'utente[28], il *file fingerprinting*, attraverso l'adozione di algoritmi di *machine learning* (ad esempio, "alberi decisionali" e "random forest"), permette di rilevare dati noti, anche in tempo reale, in canali di comunicazione criptati con alta precisione[29]. Infatti, recenti studi dimostrano che l'IA, come i classificatori di apprendimento automatico e le reti neurali, migliora la precisione delle inferenze rispetto

[25] GDPR, Considerando 26.

[26] M. FINCK, F. PALLAS, *They who must not be identified—distinguishing personal from non-personal data under the GDPR,* in *Int. Data Priv. Law*, 2020, 10, 1, p. 15: «se ci fosse un ragionevole rischio di identificazione, i dati dovrebbero essere trattati come dati personali. D'altra parte, se il rischio fosse trascurabile, i dati potrebbero essere trattati come dati non personali, e questo anche se l'identificazione non possa essere esclusa con assoluta certezza». Tuttavia, tale lettura del Regolamento orientata al rischio ha incontrato qualche resistenza, soprattutto nel cosiddetto "approccio assolutista" del Gruppo di lavoro "Articolo 29", seguito da alcune autorità di controllo, come quella francese e irlandese.

[27] ENISA, *Encrypted traffic analysis,* 2019, p. 27.

[28] P. SIRINAM ET AL., *Deep fingerprinting: Undermining website fingerprinting defenses with deep learning,* in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2018.

[29] K. BOTTINGER ET AL., *Detecting Fingerprinted Data in TLS Traffic,* in *Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security*, 2015, pp. 633-638.

all'attività e al comportamento degli utenti. Certamente, questi modelli computazionali – perfezionati da ricercatori in contesti di contrasto a nuove minacce alla *privacy* degli utenti – possono essere impiegati ai fini della prevenzione, di indagine e del perseguimento dei reati.

Nel contesto della "*IoT-surveillance*", inoltre, gli algoritmi di IA si dimostrano particolarmente efficaci nell'analisi di contenuti video criptati. Sulla base delle caratteristiche statistiche dei pacchetti di dati protetti da crittografia, un modello di *machine learning* può classificare il comportamento dell'utente alla base del traffico dati del video di sorveglianza, a condizione che gli schemi del traffico dati delle telecamere di monitoraggio differiscano significativamente quando gli utenti svolgono attività diverse nelle schermate di monitoraggio[30].

L'identificazione del dispositivo, per di più, potrebbe essere argomento ancor più pregnante circa la "sensibilità" dei metadati IoT. Chiunque, attori malintenzionati o forze dell'ordine, potrebbe osservare il volume e tasso di traffico, unitamente ai metadati di intestazione dei pacchetti del traffico di rete per dedurre dettagli sensibili sugli utenti. Il soggetto monitorante potrebbe dedurre informazioni dettagliate sugli stati, le azioni e le tipologie di dispositivi intelligenti e sensori utilizzate, così come le attività correlate degli utenti[31]. La sorveglianza passiva del traffico di rete, modalità con cui si sono sempre svolte le indagini, diventa funzionale ad alimentare algoritmi di apprendimento automatico, i quali, estrapolando da tali dati diverse caratteristiche del dispositivo (ad esempio, orari, caratteristiche dei sensori o dei dispositivi, ecc.) dal traffico criptato, consentono l'identificazione accurata delle interazioni che hanno causato il traffico di rete[32]. In altre parole, colui che monitora può venire a conoscenza delle interazioni tra un utente e un dispositivo attraverso una vasta gamma di categorie, permettendo potenzialmente il *profiling* e altre

[30] J. WANG ET AL., *User Behavior Classification in Encrypted Cloud Camera Traffic,* in *IEEE Global Communications Conference*, 2019.

[31] N. TAKBIRI ET AL., *Matching anonymized and obfuscated time series to users' profiles,* in *IEEE Trans. Inf. Theory*, 2018, 58, 2.

[32] A. ACAR ET AL., *Peek-a-boo: i see your smart home activities, even encrypted!,* in *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, 2020.

tecniche invasive della privacy[33]. Infatti, dopo l'identificazione del dispositivo, è possibile dedurre le attività dell'utente dai cambiamenti nei tassi di traffico, correlati ai cambiamenti di stato del dispositivo.

Nel contesto specifico dell'IoT, la protezione dei metadati con mezzi crittografici, che pur è stata sostenuta altrove come una possibile mitigazione contro le attività invasive nella privacy degli utenti[34], è improbabile se non impossibile. I dispositivi IoT sono per lo più limitati in termini di memoria e potenza di calcolo. Ne consegue che non sono adatti per implementare sistemi di comunicazione anonima come The Onion Router (TOR), molto esigenti in termini di capacità di elaborazione e memorizzazione.

Alla luce di ciò, è "ragionevole" attendersi che un soggetto interessato possa individuare una persona fisica mediante l'identificazione delle interazioni di un utente con un dispositivo. Pertanto, queste analisi sui metadati soddisfano la soglia stabilita dal considerando 26 del GDPR.

Il carattere personale dei metadati è ulteriormente confermato dalla normativa speciale applicabile al settore delle comunicazioni elettroniche. In forza del principio *lex generalis – lex specialis*[35], l'attuale quadro giuridico applicabile ai dati personali è in tale ambito la direttiva 2002/58/CE sulla *privacy* e le comunicazioni elettroniche (direttiva ePrivacy). La direttiva ePrivacy è stata adottata nel 2002 – poi modificata nel 2009 – e ha previsto una frammentazione terminologica dell'ampio concetto di metadati: l'articolo 2(b) stabilisce la definizione di dati sul traffico, mentre l'articolo 2(c) definisce i dati di localizzazione. Il considerando 15 specifica, inoltre, che questi dati possono riferirsi all'instradamento, alla durata, al tempo o al volume di una comunicazione, al protocollo utilizzato, all'ubicazione dell'apparecchiatura terminale del mittente o del destinatario, alla rete su cui la comunicazione ha origine o termina, all'inizio, alla fine o alla durata di una connessione.

Al fine di garantire un più elevato *standard* di protezione alla *privacy* nelle comunicazioni elettroniche, la Commissione ha proposto

[33] J. REN ET AL., *Information Exposure From Consumer IoT Devices: A Multidimensional, Network-Informed Measurement Approach,* in *IMC '19: Proceedings of the Internet Measurement Conference*, 2019, pp. 276-277.

[34] W. SCHULZ, J. VAN HOBOKEN, *Human Rights and Encryption*, UNESCO, 2016, pp. 22-23.

[35] All'interno di una stessa materia, la normativa speciale prevale su quella di carattere generale.

nel 2017 un Regolamento ePrivacy. Senza soffermarsi troppo sul merito della Proposta nei confronti del trattamento dei metadati, è importante sottolineare che l'articolo 4(3)(a) fa perno su una più ampia definizione di "dati delle comunicazioni elettroniche" per coprire sia il contenuto delle comunicazioni elettroniche che i metadati, specificando ulteriormente questi ultimi all'articolo 4(3)(c). In tal modo, la Proposta elimina essenzialmente i termini "*dati relativi al traffico* e *dati relativi all'ubicazione"*, al fine di ottenere chiarezza giuridica e coerenza con una terminologia più "standardizzata". Tuttavia, come giustamente rilevato dall'EDPS, questa definizione, seppur ampia, non è esaustiva in quanto esclude qualsiasi metadato «che non sia richiesto ai fini della trasmissione, distribuzione o scambio di contenuti di comunicazione elettronica né trattato per la fornitura del servizio», come i dati di localizzazione in un'applicazione di messaggistica istantanea[36]. In conclusione, il futuro Regolamento ePrivacy dovrebbe fornire un quadro giuridico più chiaro e comprensivo per il trattamento di tutti i tipi di metadati.

Avendo inquadrato i termini della questione, cioè la natura altamente sensibile del metadato nel contesto di una ancor timida normativa, siamo pronti a rivolgere ora la nostra attenzione alle specifiche questioni giuridiche sollevate dal dispiegamento di modelli di IA per le attività di contrasto e la sorveglianza in ambienti intelligenti: attraverso l'impiego dei suddetti algoritmi, specialmente nell'analisi dei metadati, le forze di polizia e autorità giudiziarie eviterebbero di ricorrere a pratiche di *hacking*, non violando la crittografia.

4. Metadati e contenuto delle comunicazioni: prospettive alternative nell'ambito della giustizia penale

La panoramica fin qui offerta evidenzia in modo chiaro le potenzialità dell'IA nell'analisi dei metadati, vista come potenziale alternativa ai meccanismi di rottura della crittografia nel traffico delle comunicazioni. Tuttavia, tale osservazione di partenza non ha come necessario precipitato una visione del trattamento incrociato dei dati esterni come

[36] EDPS, *EDPS recommendations on specific aspects of the proposed ePrivacy Regulation*, 2017, p. 7.

operazione "meno invasiva" nei diritti alla *privacy* e alla protezione dei dati. È difatti un dato ormai largamente consolidato sia in giurisprudenza[37], sia in dottrina[38], che l'ingerenza nei diritti fondamentali generata da tali tecniche non sia meno significativa rispetto a quella associata alla captazione del contenuto stesso delle comunicazioni.

Cionondimeno, una prima differenza tra le due tecniche di indagine potrebbe esser apprezzata non tanto dal punto di vista del *vulnus* da esse arrecato al diritto alla *privacy*, bensì da uno "epistemologico". Pur essendo entrambi in grado di fornire informazioni dettagliate sulla sfera privata degli individui, metadati e contenuto delle comunicazioni possono non essere sempre muniti della stessa efficacia persuasiva in ordine ai fatti oggetto di interesse per le autorità di contrasto. Da un lato, a dispetto del ruolo sempre giocato dall'interpretazione, la captazione del contenuto comunicativo offre una rappresentazione diretta delle parole pronunciate, e può dunque costituire prova dichiarativa di un determinato fatto[39]. Dall'altro, lo stesso non può dirsi dell'analisi dei metadati. Nonostante la loro importanza quale fonte di informazioni per gli organi di indagine, i dati esterni non possono essere qualificati come atti comunicativi e possono dunque costituire solo fonti di *inferenza*, cioè punto di partenza per una ricostruzione a contenuto probabilistico di una determinata dinamica. In altre parole, i dati esterni non possono essere considerati prova diretta del fatto che determinate parole siano state pronunciate, o che un determinato soggetto si trovasse in un dato luogo ad un determinato momento; all'opposto, essi possono solo dimostrare – seppur con una certa pro-

[37] Cfr. Corte EDU, 13 settembre 2018, *Big Brother and Others v UK*, nn. 58170/13, 62322/14 and 24960/15), §356; CGUE, 8 aprile 2014, Digital Rights Ireland, cause riunite C-293/12 e C-594/12, §27; CGUE, 21 dicembre 2016, *Tele2 Sverige and Watson and Others*, cause riunite C-203/15 e C-698/15, §99; CGUE, 6 ottobre 2020, *La Quadrature du Net*, cause riunite C-511/18, C-512/18 e C-520/18, §117. In ambito statunitense, v. *Carpenter v. United States*, No. 16-402, 585 U.S. (2018), pp. 12-15. È doveroso notare, tuttavia, che nel panorama italiano la Suprema Corte continui a sostenere un approccio di segno opposto. Si veda, ad esempio, Cass., Sez. III, 19 aprile 2019 (dep. 23 agosto 2019), n. 36380, annotata da I. NERONI REZENDE, *Dati esterni alle comunicazioni e processo penale: questioni ancora aperte in tema di* data retention, in *Sistema Penale*, 2020, 5, pp. 183-198.

[38] V. *supra* sez. 3.

[39] Sulla distinzione tra prova dichiarativa e critico-indiziaria, v. P. FERRUA, *La prova nel processo penale*, in *Rev. Bras. de Direito Processual Penal*, 2018, 4, 1, pp. 81-128, 96-105. In particolare, si veda la descrizione del carattere complesso dell'intercettazione proposta dall'Autore.

babilità – che determinate parole siano state pronunciate, o che determinati luoghi siano stati visitati da soggetti identificati. In questo senso, i metadati possono qualificarsi come un'importante fonte di informazione per la ricostruzione delle dinamiche alla base di reti criminali, nonché servire quali fondamento di misure a carattere provvisorio o investigativo (come la disposizione di una misura cautelare o di un'intercettazione stessa). Quello che difficilmente sembrano poter fare i metadati, invece, è fondare una condanna oltre il ragionevole dubbio, se non in connessione con altri elementi probatori[40].

Un secondo, e cruciale, profilo di distinzione tra analisi dei metadati e captazione del contenuto delle comunicazioni risiede nelle diverse possibilità di bilanciamento tra diritti fondamentali ed esigenze di sicurezza prospettate da tali tecniche. Come sopra sottolineato, la rottura della crittografia implica per definizione un gioco a somma zero, dove l'accesso al contenuto della comunicazione non può che risolversi in una piena valorizzazione delle esigenze di contrasto alla criminalità a detrimento di quelle di riservatezza (e viceversa). Usando un linguaggio metaforico, per accedere al contenuto è fondamentale avere la giusta chiave per la porta: o non la si ha, e il contenuto rimane protetto dentro le mura di casa; o la si trova, e il contenuto diviene accessibile agli osservatori esterni.

A fronte di questa alternativa binaria, il trattamento dei metadati si presenta invece come un meccanismo più flessibile, in cui l'ingerenza nel diritto alla *privacy* risulta "modellabile" sia per estensione, sia per profondità. Ad esempio, come chiarito dalla stessa Corte Edu, le misure di sorveglianza di massa non configurano sempre una grave restrizione del diritto alla *privacy* e alla protezione dei dati dal punto di vista *individuale*[41]. In tali casi, difatti, l'attenzione delle autorità di contrasto o degli organi delegati è spesso "diluita" nell'ingente mole di informazioni analizzata, con conseguenze certamente più lievi per

[40] Cfr. Art. 192(2) c.p.p.

[41] Tale affermazione non deve essere letta nel senso di sminuire le crescenti preoccupazioni rispetto al concetto di *group privacy*. Se queste assumono particolare rilevanza nel mondo commerciale, consentendo pratiche quali il *targeted advertising*, nell'ambito della giustizia penale le attività delle forze dell'ordine sono (quasi) sempre votate all'identificazione del singolo, e le conseguenze più sensibili per la sfera giuridica degli interessati si realizzano solo quando l'attenzione degli organi competenti si focalizzi su individui ben identificati.

la sfera privata dei singoli interessati[42]. La quantità di dati trattati, l'arco di tempo sul quale le operazioni di trattamento si estendono, nonché il livello di invasività delle inferenze ottenute costituiscono d'altro canto valide coordinate attraverso le quali valutare la serietà dell'ingerenza nei diritti fondamentali dell'individuo[43]. L'IA come tecnologia privilegiata per l'analisi di ingenti volumi di informazioni (i cd. *big data*) sembra dunque essere, in alternativa alle tecniche di *hacking*, un valido strumento per conseguire interferenze più bilanciate nei diritti fondamentali, soddisfacendo al contempo le istanze di sicurezza collettiva.

Nel paradigma di una sempre più crescente produzione di dati nel contesto IoT, vedremo ora come anche la giurisprudenza delle due Corti sovranazionali europee sembri manifestare una maggiore consapevolezza rispetto alle potenzialità di un tale approccio diversificato.

5. Proporzionalità e analisi dei metadati nella giurisprudenza delle Corti europee

Uno sguardo complessivo alla giurisprudenza più recente della Corte Edu e della CGUE rivela una tendenza comune. Nella valutazione dei – sempre discussi – programmi di sorveglianza di massa emerge infatti una maggiore accettazione della compatibilità di tali misure con il sistema europeo multilivello di protezione dei diritti fondamentali. Un primo segno in questo senso arriva dalla Corte di Strasburgo che, nel caso *Big Brother Watch*, stima che gli Stati Parte alla Convenzione abbiano ancora un ampio margine di apprezzamento quanto alla scelta della "fisionomia" dei sistemi di sorveglianza da attuare[44]. Anche alla luce del principio di proporzionalità, dunque, le misure di monitorag-

[42] B. VAN DER SLOOT, E. KOSTA, *Big Brother Watch and Others v UK: Lessons from the Latest Strasbourg Ruling on Bulk Surveillance*, in *Eur Data Prot L Rev*, 2020, 5, p. 258. Il caso di specie prende le mosse da una serie di ricorsi promossi da diverse associazioni davanti alla Corte Edu, all'indomani delle cd. rivelazioni di Edward Snowden nel 2013. Tali ricorsi, facendo leva sull'Art. 8 CEDU, censuravano i programmi di intercettazione di massa condotti dai servizi di intelligence inglesi sulla base dell'Art. 8(4) del RIPA (*Regulation of Investigatory Powers Act 2000*), nonché le iniziative di *intelligence sharing* portate avanti tra tali organi e le autorità di intelligence statunistensi.

[43] Cfr. CGUE, 2 marzo 2021, *H.K./Prokuratuur*, caso C-746/18, §37.

[44] *Big Brother Watch*, *op. cit.*, §356.

gio indiscriminate non devono essere considerate di per sé contrarie all'Art. 8 CEDU.

Tuttavia, lo scrutinio della Corte diviene necessariamente più severo sulle modalità di realizzazione dei sistemi di sorveglianza in questione. In particolare, la Grande Camera distingue in tali scenari quattro passaggi di crescente interferenza nei diritti fondamentali: la raccolta e iniziale conservazione del contenuto e dati esterni alle comunicazioni; l'applicazione di criteri-filtro prestabiliti (*selectors*) ai dati raccolti; l'esame da parte di analisti delle informazioni filtrate; la conservazione di tali dati e il loro uso per la produzione di report, anche ai fini della condivisione con altri Stati[45]. Man a mano che si procede in tali passaggi, rimarca crucialmente la Corte, maggiori garanzie devono essere apprestate nei confronti degli individui finiti sotto la lente di ingrandimento delle autorità di contrasto[46].

I giudici del Lussemburgo non hanno tardato ad allinearsi almeno in parte al dato di partenza di quest'impostazione, seppur non abbandonando *tout court* la linea costantemente adottata dal caso *Digital Rights Ireland*. Non si pretende, nei confini del presente contributo, di offrire una disanima dettagliata dell'ormai ricca giurisprudenza della CGUE sul tema. Tuttavia, due elementi devono essere posti in luce nel recente caso *La Quadrature du Net*[47]. Anzitutto, come la Corte Edu, la CGUE chiarisce che gli obiettivi di protezione della sicurezza nazionale possono legittimare programmi di sorveglianza *indiscriminati*, per il tempo necessario alla neutralizzazione di una minaccia prevedibile e concreta all'integrità dello Stato[48]. Tale scenario giusti-

[45] Corte Edu, 25 maggio 2021, *Big Brother Watch et al c. Regno Unito*, nn. 58170/13, 62322/14 e 24960/15, §325. Si deve notare che, in questa sistematizzazione delle operazioni di sorveglianza, la Corte non distingue tra programmi di trattamento dei dati esterni e vere e proprie misure di intercettazione.

[46] *Id.*, §330.

[47] Il caso in questione, sulla scia di *Tele2/Watson*, si inserisce in un contenzioso promosso davanti alle giurisdizioni francesi, belga e britanniche da alcune associazioni di attivisti per la *privacy* (tra cui, per l'appunto, la francese *Quadrature du net*). Nei ricorsi presentati, in particolare, si censurava la conformità con la Carta delle legislazioni nazionali adottate in forza dell'Art. 15(1) della direttiva e-Privacy, la quale consente ai *provider* dei servizi di telecomunicazione di procedere a una raccolta indiscriminata dei dati degli utenti ai fini della lotta alla criminalità grave e al terrorismo.

[48] *La Quadrature du Net*, *op. cit.*, §§136-137.

ficherebbe anche l'attivazione di un obbligo legislativo per i *service provider* che – dietro richiesta delle autorità di contrasto – sono tenuti ad effettuare screening indiscriminati dei metadati generati dalle attività degli abbonati, attraverso l'utilizzo di criteri prestabiliti ed in continua evoluzione[49]. Diversamente, in linea con l'approccio anteriore, la lotta alla "mera" criminalità grave giustificherebbe unicamente misure il cui ambito di applicazione sia delimitato da criteri oggettivi (e.g. geografici, soggettivi), i quali dimostrino un nesso di pertinenza tra i dati raccolti e gli obiettivi perseguiti[50].

La Corte si spinge anche oltre nell'esaminare ulteriori ipotesi di interferenza nei diritti fondamentali. Ai nostri fini, in particolare, è significativo il caso della raccolta in tempo reale di dati da parte delle autorità di contrasto. Per la gravità dell'ingerenza, che implica un accesso diretto alle informazioni da parte dell'organismo pubblico, una tale misura limitativa dei diritti dev'essere autorizzata *ex ante* da un'autorità indipendente; il principio di proporzionalità, inoltre, impone che essa sia limitata ad individui determinati, già identificati in quanto soggetti probabilmente connessi ad attività terroristiche[51]. In qualche modo, tale misura viene vista tra l'altro come un'operazione logicamente successiva rispetto alle attività di screening "allargato", volte all'individuazione di target promettenti per le autorità di contrasto.

È proprio in questa prospettiva, difatti, che si inserisce l'argomentazione al cuore di questo contributo. L'IA, quale strumento di contrasto alla criminalità, non dev'essere esclusivamente vista quale tecnologia alla base di invasive attività di monitoraggio; in diverse ipotesi, essa può anche essere parte della soluzione, ponendosi quale mezzo di delimitazione delle attività di sorveglianza più focalizzate sull'individuo, e dunque foriere delle interferenze più significative nei diritti fondamentali. Nella sfera preventiva, questa potenzialità si esprime nei suoi termini più efficaci. In un ambito in cui, a priori, è difficile determinare quali informazioni possano essere utili alla lotta alla criminalità, e a specifici atti illeciti non ancora materializzatisi, l'IA è dotata delle capacità di trattare grandi masse di dati, limitando

[49] *Id.*, §177.
[50] *Id.* §§143-144, 147-148; *Privacy International*, *op. cit.*, §75.
[51] *La Quadrature du Net*, *op. cit.*, 191.

le inferenze più significative ai *target* che sembrano rappresentare un rischio maggiore per la generalità dei consociati.

Certamente non si ignorano i pericoli connessi alle argomentazioni prospettate. La dottrina ha da tempo messo in luce le controindicazioni legate alla proliferazione di sistemi di sorveglianza di massa nelle società democratiche, anche alla luce delle recenti prese di posizione delle Corti europee appena esposte[52]. È tuttavia importante mantenere una prospettiva aperta rispetto alle potenzialità offerte da tecnologie come l'IA, soprattutto in connessione con il cambiamento di paradigma portato dall'avvento dell'IoT. È difatti la crescente complessità delle interazioni e flussi di dati negli "ambienti intelligenti" a richiedere un ripensamento delle categorie finora utilizzate per analizzare le attività di sorveglianza in campo penale, e altrove. L'aumento incontrollato della produzione di dati, nonché dei mezzi disponibili per ricavarne informazioni utili, sembra infatti portare la Corte Edu e la CGUE ad adottare un approccio meno restrittivo rispetto alle misure non chiaramente circoscritte nel loro ambito di applicazione. Per un verso, almeno in campo penale, si riconosce che il trattamento dei dati nei sistemi di sorveglianza di massa non sempre sfocia in gravi interferenze nella vita privata dei singoli; per l'altro, si limitano le misure più invasive – quali quelle che implicano un accesso diretto e in tempo reale da parte delle autorità di contrasto – a soggetti già identificati. Se questo orientamento può chiaramente sfociare in una "normalizzazione della sorveglianza di massa"[53], l'adozione di un rigoroso approccio procedurale – informato al principio di proporzionalità e implementato grazie alle tecniche di IA – sembra essere adatto ad impedire tali derive.

Per arrivare a questo punto, però, un controllo più severo si renderebbe necessario, almeno da parte della Corte Edu. Due infatti sono i fattori che indeboliscono allo stato l'orientamento sostenuto dalla Corte, con l'affioramento di seri pericoli per le garanzie individuali. Un primo problema riguarda l'adozione di un approccio "olistico"

[52] N. NI LODEAIN, *Not So Grand: The Big Brother Watch ECtHR Grand Chamber Judgement*, in *Information Law and Policy Center,* 28 maggio 2021, https://infolawcentre.blogs.sas.ac.uk/2021/05/28/not-so-grand-the-big-brother-watch-ecthr-grand-chamber-judgment/; M. MILANOVIC, *The Grand Normalization of Mass Surveillance: ECtHR Grand Chamber Judgments in Big Brother Watch and Centrum för rättvisa*, in *Ejil:Talk!*, 26 maggio 2021, https://www.ejiltalk.org/the-grand-normalization-of-mass-surveillance-ecthr-grand-chamber-judgments-in-big-brother-watch-and-centrum-for-rattvisa.

[53] M. MILANOVIC, *op. cit.*, *The Grand Normalization of Mass Surveillance*.

nell'esame dei sistemi di sorveglianza censurati. Secondo una linea simile a quella già applicata nei casi incentrati su violazioni dell'Art. 6 CEDU[54], la Corte tende ad esaminare i requisiti di legittimità dei regimi di sorveglianza "nel loro complesso". Ciò significa che le lacune concernenti alcune specifiche garanzie sono suscettibili di essere "compensate" da altre previsioni del sistema[55], con conseguente "diluizione" di determinate protezioni individuali.

Una seconda debolezza riguarda invece l'approccio erroneamente diversificato per la captazione di contenuto comunicativo e metadati. A fronte di una pari invasività dal punto di vista della *privacy*, la Corte non sembra richiedere per l'analisi dei metadati lo stesso standard di garanzie rispetto alle intercettazioni, in particolare rispetto alla supervisione dei criteri di selezione delle informazioni[56]. Ciò avverrebbe anche nel caso in cui la misura di monitoraggio si focalizzi su un individuo ben identificato, ma non si "agganci" un dispositivo specifico[57]. Come sottolineato dal giudice Pinto de Albuquerque, tale rilievo non solo sembra ignorare l'equipollenza di contenuto e dato esterno in termini di potenziali lesioni della sfera privata, ma si lega ancora ad una obsoleta e sofistica tipizzazione delle misure di sorveglianza, che separa quelle riferite al *dispositivo* intercettato da quelle riferite all'*individuo*[58].

6. Conclusioni

Il presente contributo ha evidenziato l'irriducibile *trade-off* di valori che varie tecniche di "intercettazione" – tra cui, l'*hacking* di Stato – pongono in essere, presentando rischi sempre maggiori per gli individui che si muovono in ambienti "intelligenti". Nelle moderne società dell'informazione, così caratterizzate da una proliferazione del flusso

[54] V. MANES, M. CAIANIELLO, *Introduzione al Diritto Penale Europeo*, Giappichelli, Torino, pp. 217-221.

[55] *Big Brother Watch, op. cit.*, §170. Nel caso di specie, tale tecnica è servita ad evitare la censura del sistema di intelligence britannico, che mancava di una precisa tipizzazione dei "reati-presupposto" che legittimavano l'attivazione della misura.

[56] *Big Brother Watch*, *op. cit.*, §421.

[57] *Big Brother Watch*, *op. cit.*, §421.

[58] *Big Brother Watch*, *op. cit.*, Opinione in parte concorrente e in parte dissidente del Giudice Pinto de Albuquerque, §12.

di dati ed un'evoluzione delle aspettative di *privacy*, le maglie della sorveglianza digitale vengono progressivamente allargate, specialmente in senso preventivo. Se tale fenomeno può tradursi in una lesione incontrollata delle libertà fondamentali, almeno in campo penale la tendenza sembra essere quella di rafforzare le garanzie mano a mano che l'attenzione delle autorità di contrasto si focalizza sul singolo.

Anche in questo caso, tuttavia, un bilanciamento più ragionevole tra i diritti fondamentali e le esigenze di sicurezza sembra essere offerto dalle capacità di trattamento dei dati dell'IA. Diversamente dalla rottura della crittografia, che si traduce sempre un gioco a somma zero, l'analisi dei metadati permette di introdurre una pluralità di parametri nel test di proporzionalità, così ottenendo *inferenze* – e dunque, *interferenze* – più calibrate agli obiettivi di sicurezza perseguiti.

In definitiva, a dispetto delle debolezze tuttora presenti negli orientamenti delle due Corti europee, l'IA sembrerebbe ad oggi aprire strade di maggiore equilibrio rispetto alla rottura della crittografia, offrendo così adeguate possibilità di bilanciamento tra i diritti fondamentali alla *privacy* da un lato, e le istanze di protezione della collettività dall'altro.