LIBER AMICORUM
PER
PASQUALE COSTANZO

ARIANNA VEDASCHI

INTELLIGENZA ARTIFICIALE E MISURE ANTITERRORISMO ALLA PROVA DEL DIRITTO COSTITUZIONALE

17 FEBBRAIO 2020

CONSULTA ONLINE
PERIODICO TELEMATICO ISSN 1971-9892

LIBER AMICORUM
PER
PASQUALE COSTANZO

Arianna Vedaschi

Intelligenza artificiale e misure antiterrorismo alla prova del diritto costituzionale*

SOMMARIO: 1. L'azione di contrasto al terrorismo jihadista: tra prevenzione della minaccia e percezione di (in)sicurezza. – 2. Il ruolo della tecnologia nelle misure antiterrorismo: dai *big data* all'*artificial intelligence*. – 3. Le applicazioni dell'intelligenza artificiale alla lotta al terrorismo internazionale: uno scenario ancora "*work in progress*". – 3.1. *La rimozione dei contenuti terroristici online: dalle* partnerships *su base volontaria ai primi approcci regolatori*. – 3.2. *Intelligenza artificiale,* counter-terrorism measures *e diritto costituzionale: una triangolazione "pericolosa"?* – 4. Osservazioni conclusive.

1. *L'azione di contrasto al terrorismo jihadista: tra prevenzione della minaccia e percezione di (in)sicurezza*

A far tempo dagli attentati dell'11 settembre 2001, che – come noto – hanno tragicamente colpito le Twin Towers e il World Trade Center di New York, le democrazie occidentali[1] si sono trovate ad affrontare un'inedita minaccia, non assimilabile ad altri fenomeni terroristici che, ben prima dell'inizio del nuovo millennio, avevano interessato diversi Paesi europei. Si pensi alla Spagna e al Regno Unito, messi alla prova dal terrorismo separatista-indipendentista rispettivamente basco e nordirlandese, ma anche all'Italia e alla Germania, che, negli anni Settanta, hanno tristemente conosciuto il terrorismo politico. Almeno apparentemente, il terrorismo internazionale di matrice jihadista non presenta – soprattutto nell'immediato *post*-9/11[2] – una dichiarata finalità politica, ad esempio, di carattere indipendentista o sovversiva(-anarchica) in senso lato, né sembra proporre una propria concezione di Stato. La "nuova" violenza terroristica si scaglia contro il "diverso", cioè il "mondo occidentale" (*rectius*, lo stile di vita occidentale), senza una precisa strategia e, almeno fino alla proclamazione del c.d. Stato islamico (IS)[3], persino a prescindere dal perseguimento di un obiettivo prestabilito. Financhè la continua mutazione delle tattiche impiegate per la realizzazione degli attentati rende, se non impossibile, assai difficile prevederli, nonostante l'attenzione continua delle autorità di *law enforcement* e il potenziamento dei servizi di *intelligence*[4], costantemente impegnati in un'azione di contrasto dalla forte curvatura preventiva. L'imprevedibilità è, in ultima analisi, il vero elemento caratterizzante del "terrore islamista".

Vieppiù, nella maggior parte dei casi, l'attentatore-terrorista muore suicida nel commettere l'azione criminale, il che rende impossibile assicurarlo alla giustizia e sottoporlo a misure sanzionatorie *ex post facto*.

---

* *L'Autrice desidera ringraziare la dott.ssa Chiara Graziani per il supporto nella ricerca bibliografica e per l'*editing *delle note.*

[1] Benché la prospettiva del presente lavoro sia focalizzata sull'Occidente, non si può non ricordare che il terrorismo internazionale ha colpito anche – e, in alcuni momenti storici, in misura addirittura maggiore – l'Africa e l'Asia; per avere un'idea della distribuzione geografica degli attacchi terroristici in tempi recenti, si veda Europol, *European Union Terrorism Situation and Trend Report*, 2019.

[2] Il terrorismo internazionale è invero un fenomeno dinamico e multiforme; infatti, ha nel tempo subito diverse trasformazioni, tanto nella struttura quanto nelle ambizioni, sfociate nel 2014 nell'autoproclamazione dello Stato Islamico, successivamente sconfitto sul campo, stanti le perdite territoriali che si sono susseguite dal gennaio 2015 in poi, culminate nel marzo 2019, quando le Forze democratiche siriane – coalizione comprendente milizie curde, siriane e arabe – hanno liberato anche l'ultima roccaforte del "Califfato", Baghuz. Per una ricostruzione dei fatti del marzo 2019 e alcune considerazioni sulle sorti dello Stato islamico nel "*post*-Califfato", F. MARRONE-M. OLIMPIO, *La minaccia jihadista dopo il Califfato,* nelle pubblicazioni dell'ISPI (28 marzo 2019).

[3] Per un approfondimento, sia consentito il rinvio ad A. VEDASCHI, *Da al-Qāʿida all'IS: il terrorismo internazionale* si è fatto *Stato?*, in *Riv. trim. dir. pubbl.*, 1/2016, 41 ss.

[4] Per un quadro aggiornato sul ruolo dei servizi segreti e sul loro potenziamento nel contrasto al terrorismo, si veda D. MARTIN JONES-P. SCHULTE-C. UNGERER-M.L.R. SMITH (eds.), *Handbook of Terrorism and Counter Terrorism Post 9/11*, Elgar, Cheltenham and Northampton, 2019.

Questo scenario spiega la reazione dei governi e dei legislatori, che all'approccio *repressivo*, ovverosia volto a punire chi abbia già commesso il fatto di reato, hanno preferito quello *preventivo*, mirato ad una sorta di "tutela anticipatoria", cioè finalizzata ad evitare che situazioni potenzialmente "pericolose" evolvano nell'effettiva commissione di atti criminali. Pertanto, non pare eccessivo affermare che l'azione a forte connotazione preventiva rappresenta la vera e propria "chiave di lettura" delle attuali strategie di contrasto al terrorismo internazionale[5].

Dall'imprevedibilità del terrorismo internazionale di matrice jihadista deriva poi la sua attitudine a generare *insicurezza* nella società. Di qui l'ambizione delle competenti autorità pubbliche di adottare misure tese a rispondere alla *percezione* di insicurezza della popolazione. Anzi, si può sostenere che, nelle agende politiche contemporanee, l'obiettivo di rassicurare gli individui risulta, almeno a tratti, prevalente, rispetto a quello di proteggere effettivamente la collettività[6].

In questo quadro emerge la duplice dimensione dell'azione di contrasto al terrorismo internazionale, giacché, a quella *oggettiva*, tesa alla concreta applicazione delle misure antiterrorismo per *garantire* l'*effettiva sicurezza*, si affianca quella *soggettiva*, tutta ruotata sulla capacità delle *counter-terrorism measures* di *far sentire* i singoli individui al sicuro da possibili attentati, così da rassicurare la collettività nel suo insieme. Quello soggettivo è un aspetto fortemente simbolico dell'antiterrorismo, ma non per questo meno centrale nelle valutazioni, di stampo prettamente politico, tenute presente dalle autorità pubbliche nell'adozione dei provvedimenti volti a combattere l'inedita minaccia del "terrore jihadista".

Nelle democrazie avanzate, le misure di contrasto al terrorismo trovano, però, il loro limite nel nucleo essenziale dei diritti della persona, elencati dai testi costituzionali, dalle carte sovranazionali e dai trattati internazionali. Negli ultimi decenni, invero, il delicato e complesso bilanciamento tra la tutela dei diritti e le esigenze della sicurezza non solo è diventato un tema ampiamente studiato dalla dottrina[7], italiana e straniera, ma è stato spesso portato all'attenzione delle corti[8], tanto nazionali quanto sovranazionali. Nel decidere i casi controversi, non di rado, le corti hanno dovuto rilevare la violazione dei limiti minimi di tutela dei diritti fissati dai cataloghi costituzionali o sovranazionali. Anzi, anche allorquando non hanno escluso la possibilità teorica della legittimità della misura *counter-terrorism* sottoposta al loro scrutinio, ad esempio affrontando la questione della legittimità della sorveglianza di massa[9], ne hanno comunque riscontrato la non conformità, sul piano pratico, alle garanzie dei diritti tutelati dalle fonti sovranazionali o costituzionali. Da questa prospettiva, si può dunque evidenziare il *gap* fra la dimensione teorica e quella pratica. In termini più chiari, benché l'azione antiterroristica, messa a punto dai *law-makers* per salvaguardare la sicurezza (o almeno la sua percezione sociale), venga, sul piano teorico-astratto, ritenuta

[5] Per uno studio sul carattere preventivo delle diverse misure antiterrorismo, si veda C.C. MURPHY, *EU Counter-Terrorism Law: Pre-Emption and the Rule of Law*, Hart Publishing, Oxford, 2012.

[6] L'incertezza, generata dalla minaccia terroristica, viene spesso sfruttata dalle derive populiste, che strumentalizzano la legittima richiesta di sicurezza dei cittadini, v. M. BARBERIS, *Il cavaliere oscuro. Euristiche della (in)sicurezza*, in *Rassegna di diritto pubblico europeo*, 2/2019 (in corso di pubblicazione).

[7] In particolare, vi è chi ritiene che il rapporto libertà-sicurezza debba continuare ad essere letto secondo il canone regola-eccezione, com'era prima degli attentati del 9/11; in questo senso, si veda per tutti A. PACE, *Libertà e sicurezza. Cinquant'anni dopo*, in *Diritto e società*, 2/2013, 177 ss. Altri, invece, elevano la sicurezza al rango di diritto fondamentale, v. G. DE VERGOTTINI, *La difficile convivenza fra libertà e sicurezza. La risposta delle democrazie al terrorismo*, in *Rass. parl*., 2/2004, 427 ss. Non manca poi chi configura la sicurezza come vero e proprio «valore superprimario»: G. CERRINA FERONI-G. MORBIDELLI, *La sicurezza: un valore superprimario*, in *Percorsi costituzionali*, 1/2008, 31 ss.

[8] Con specifico riguardo alla Corte costituzionale italiana, si rinvia a P. PASSAGLIA (a cura di), *Giurisprudenza costituzionale e terrorismo*, Servizio studi della Corte costituzionale, agosto 2018.

[9] Esempio evidente è il Parere 1/15 della Corte di giustizia dell'Unione europea, con cui i giudici di Lussemburgo, pur legittimando, a livello teorico, la sorveglianza di massa come misura potenzialmente adatta a prevenire attacchi terroristici e, quindi, a tutelare la pubblica sicurezza, censurano però l'accordo internazionale stretto tra Canada e Unione europea sullo scambio di dati riguardati il traffico aereo. Corte di giustizia dell'Unione europea, A-1/15, 26 luglio 2017. A. VEDASCHI, *L'Accordo internazionale sui dati dei passeggeri aviotrasportati (PNR) alla luce delle indicazioni della Corte di giustizia dell'Unione europea*, in *Giur. cost.*, 4/2017, 1913 ss.

compatibile con il quadro normativo di riferimento, sul piano pratico, non supera però il vaglio delle corti, che, in qualche caso, si spingono a formulare una serie di linee guida utili per la necessaria (successiva) opera di revisione delle misure da applicare in concreto[10].

2. *Il ruolo della tecnologia nelle misure antiterrorismo: dai* big data *all'*artificial intelligence

Nel panorama appena delineato, la tecnologia ha assunto un ruolo chiave per una molteplicità di ragioni, correlate tanto alla necessità, per le autorità pubbliche, di considerare la dimensione soggettiva della sicurezza, ovverosia della sua percezione collettiva, quanto al tentativo di bilanciare le esigenze della sicurezza con i principi cardine dello Stato di diritto (e, in particolare, garantire l'effettiva salvaguardia delle libertà personali).

In chiave diacronica, va, anzi tutto, rilevato che il diffondersi della minaccia terroristica internazionale si è combinato con la rapidissima evoluzione della tecnologia[11], che ha indubbiamente caratterizzato i primi venti anni di questo secolo. Infatti, da strumenti a disposizione di una ristretta *élite*, Internet e le altre tecnologie digitali si sono trasformati in vere e proprie "pertinenze" di ciascun individuo. Tale pervasività dell'utilizzo del mezzo tecnologico lo ha trasformato in una sorta di "arma a doppio taglio": da una parte, è ben noto come i terroristi islamisti la sfruttino quale efficace mezzo di radicalizzazione, indottrinamento e reclutamento[12]; dall'altra, sono altrettanto indubbi i vantaggi offerti dalla tecnologia alle autorità pubbliche impegnate nella lotta al terrorismo internazionale.

A tal proposito, sembra utile richiamare, benché in estrema sintesi, i principali *tools* tecnologici impiegati per contrastare il terrorismo di matrice jihadista.

Fin dalle prime reazioni dei legislatori alla minaccia terroristica, è risultata evidente la scelta di utilizzare le tecnologie a scopo di raccolta, conservazione e analisi di dati potenzialmente utili a monitorare attività sospette[13]. Con il passare del tempo, scandito dai progressi della tecnologia, le attività di sorveglianza hanno assunto un carattere massivo. La possibilità di avere a disposizione, incrociare e analizzare rapidamente, grazie alla c.d. *big data analytics*[14], un significativo numero di dati ha richiesto la predisposizione di strumenti giuridici che regolassero l'utilizzo di queste tecniche (cioè dei meccanismi usati per processare le informazioni)[15]. Su tale versante, va segnalato

[10] Emblematico, in questo senso, è ancora il Parere 1/15, adottato dalla Corte di giustizia dell'Unione europea il 26 luglio 2017, cit., nota 9, spec. §§ 154-231.

[11] Invero, la tecnologia ha un impatto rilevante su un amplissimo novero di attività umane; cfr. P. COSTANZO, *Il fattore tecnologico e le sue conseguenze*, in *Rass. parl.*, 4/2012, 811 ss.

[12] La rete è usata anche per lo scambio di informazioni tra i membri delle "cellule del terrore" e, addirittura, per l'organizzazione di attentati. Tuttavia, queste comunicazioni si svolgono solitamente sul c.d. *dark web*, ossia una parte della rete che non viene indicizzata dagli ordinari motori di ricerca e che necessita di chiavi di accesso particolari. Sul punto v. G. WEIMANN, *Going Dark: Terrorism on the Dark Web*, in 39 *Studies in Conflicts and Terrorism*, 2016, 195 ss.

[13] Si veda, per esempio, il *Patriot Act*, che costituisce l'immediata risposta degli Stati Uniti d'America agli attacchi dell'11 settembre 2001. *Patriot Act*, P.L. 107-56, Oct. 26, 2001. Su questa legislazione, tra i tanti, D. COLE, *Oversight of the USA Patriot Act: Hearing Before the S. Comm. on the Judiciary, 109th Cong., Apr. 5, May 10, 2005*, in *Georgetown University Law Center*, 2005 e K.L. SCHEPPELE, *Exceptions that Prove the Rule. Embedding Emergency Government in Everyday Constitutional Life*, in J.K. TULIS-S. MACEDO (eds.), *The Limits of Constitutional Democracy*, Princeton University Press, Princeton, 2010, 124 ss.

[14] Va notato, infatti, che per funzionare gli algoritmi hanno necessità di un'enorme quantità di dati. Sul rapporto fra algoritmi e *big data*, v. P. COSTANZO, *La democrazia digitale (precauzioni per l'uso)*, in *Diritto pubblico*, 1/2019, 71 ss.

[15] V., ad esempio, la possibilità di analizzare i c.d. metadati, cioè dati di comunicazione che non rivelano il contenuto della stessa, ma consentono di conoscere significative informazioni afferenti ad essa (la durata di una chiamata telefonica, il luogo da cui viene inviata una email, l'orario in cui si accede ad un determinato sito, ecc.). Sul punto, E. GUILD-S. CARRERA, *The Political and Judicial Life of Metadata: Digital Rights Ireland and the Trail of the Data Retention Directive*, CEPS Paper in Liberty and Security, CEPS, Brussels, 2014.

l'affinamento di metodi quali la criptoanalisi (*signals intelligence*)[16] e la costruzione di enormi *databases* in grado di conservare e confrontare i c.d. metadati, in modo da consentire la profilazione[17] degli individui, ora non più volta, come era prima dell'11 settembre 2001, a fini meramente commerciali, bensì orientata in un'ottica criminal-preventiva.

Non di rado, si sono posti dubbi sulla compatibilità di questi strumenti con i diritti della persona, quali quelli della *privacy*, della *data protection*, ma anche con principi fondamentali, come ad esempio la presunzione di innocenza. Chiamate a verificare questi dubbi di legittimità, le corti si sono spesso pronunciate censurando specifici aspetti delle normative adottate per autorizzare e regolare l'impiego delle menzionate tecniche di sorveglianza[18].

Le tecnologie in parola, sebbene portate a livelli assai avanzati grazie ai progressi della ricerca, si caratterizzano, almeno in un primo momento, per essere strumenti che non si autoprogrammano. In altre parole, essi, ricevuti certi dati come *input*, sono in grado di analizzarli, combinarli e confrontarli, in modo assai rapido ed efficiente, ma non riescono a fare operazioni autonome capaci di sfociare in un *output* originale.

La situazione è, tuttavia, cambiata. Da alcuni anni, la c.d. intelligenza artificiale sta conoscendo un rapidissimo sviluppo, testimoniato dall'elaborazione di algoritmi "intelligenti"[19], che vengono sfruttati per facilitare, velocizzare e migliorare un altissimo numero di attività umane[20].

L'"intelligenza artificiale" consente agli algoritmi di programmarsi e affinarsi in autonomia. In questo senso, sono intuibili le implicazioni etiche, tant'è che la Commissione europea, dopo avere istituito, nel giugno 2018, un gruppo di esperti indipendenti sull'intelligenza artificiale, ha reso pubblico, nell'aprile 2019, un documento intitolato "Orientamenti etici per un'IA affidabile"[21]. L'intelligenza artificiale è qui definita come quell'insieme di «[s]istemi software (ed eventualmente hardware) progettati dall'uomo che, dato un obiettivo complesso, agiscono nella dimensione fisica o digitale percependo il proprio ambiente attraverso l'acquisizione di dati, interpretando i dati strutturati o non strutturati raccolti, ragionando sulla conoscenza o elaborando le informazioni derivate da questi dati e decidendo le migliori azioni da intraprendere per raggiungere l'obiettivo dato». Lo stesso documento precisa, poi, che questi sistemi possono persino «adattare il loro comportamento analizzando gli effetti che le loro azioni precedenti hanno avuto sull'ambiente». Si tratta, quindi, di assetti che sono in grado di simulare i processi logici propri dell'intelligenza

[16] Si tratta di un'attività consistente nel raccogliere informazioni grazie all'intercettazione e all'analisi di comunicazioni cifrate che vengono decrittate. Sul punto, M.V. HAYDEN, *Balancing Security and Liberty: The Challenge of Sharing Foreign Signals Intelligence*, in *Notre Dame Journal of Law, Ethics and Public Policy*, 2005, 247 ss.

[17] In tema di profilazione, si veda H. DUFFY, *The 'War on Terror' and the Framework of International Law*, Cambridge University Press, Cambridge, 2015, 637 ss.

[18] Cfr. supra, § 1, spec. nn. 9 e 10. Già prima del Parere 1/15 sulla raccolta e analisi dei dati PNR, la Corte di giustizia dell'Unione europea si era distinta per un attento scrutinio sulle tecniche di sorveglianza di massa. Tra le decisioni più importanti, si vedano: Corte di giustizia dell'Unione europea, sent. 8 aprile 2014, C-293/12 e C-594/12, *Digital Rights Ireland*; Id., sent. 6 ottobre 2015, C-362/14, *Schrems*; Id., sent. 21 dicembre 2016, C-203/15 e C-698/15, *Tele2 Sverige AB*. Altri casi sono attualmente pendenti dinanzi la Corte di Lussemburgo; in particolare: causa C-523/17, *Privacy International*; cause riunite C-511/18, *La Quadrature du Net et al*. e C-512/18, *French Data Network et al.*; causa C-520/18 *Ordre des barreaux francophones et germanophone et al.*; causa C-311/18, *Schrems*.

[19] Basati su tecniche di *machine learning* e *deep learning*. Il *machine learning* è una branca dell'intelligenza artificiale volta a migliorare, grazie all'applicazione di metodi statistici, la *performance* di un algoritmo. Per questo motivo, si parla anche di "tecniche di apprendimento automatico". Il *deep learning* è, a sua volta, una branca del *machine learning*, i cui algoritmi funzionano "su più livelli" e ricalcano la struttura delle reti neurali del cervello umano (ci si riferisce a queste tecniche, infatti, anche con il nome di "reti neurali artificiali"). B. GANOR, *Artificial or Human: A New Era of Counterterrorism Intelligence?*, in *Studies in Conflict & Terrorism*, 1/2019, 1 ss.

[20] Si veda sul tema, con un approccio critico dal punto di vista del diritto costituzionale, C. CASONATO, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Dir. pubbl. comp. eur.*, Speciale/2019, 101 ss.

[21] Gruppo indipendente di esperti sull'intelligenza artificiale, istituito dalla Commissione europea nel giugno 2018, *Orientamenti etici per un'IA affidabile*, 8 aprile 2019.

umana, imparando e applicando un ragionamento, in alcuni casi addirittura autocorreggendosi (il che implica la possibilità che essi commettano errori)[22].

Alla luce del quadro delineato nel paragrafo introduttivo e con specifico riferimento al settore delle misure antiterrorismo, l'utilità di queste tecniche risulta palese. In un contesto in cui la prevenzione è il fattore alla base di tutte le strategie adottate, strumenti che applichino ragionamenti autonomi capaci di creare nuove conoscenze e, dunque, di prevedere (allo scopo di prevenire) possibili sviluppi di fatti e/o comportamenti pericolosi potrebbero, in effetti, costituire un punto di svolta nel *counter-terrorism*.

Sul versante parallelo, e di estremo interesse per il giurista costituzionalista, occorre però valutare se e fino a che punto l'applicazione di queste tecnologie resti compatibile con la tutela dei diritti della persona e non implichi restrizioni non proporzionate[23] degli stessi.

3. *Le applicazioni dell'intelligenza artificiale alla lotta al terrorismo internazionale: uno scenario ancora "*work in progress*"*

Ferma l'utilità dell'intelligenza artificiale, il suo concreto impiego nel *counter-terrorism* è, ad oggi, piuttosto limitato e, per larga parte, ancora in fase sperimentale. Pare tuttavia utile menzionare gli ambiti di effettiva o almeno potenziale applicazione.

In primo luogo, queste tecniche possono essere sfruttate per tentare di prevenire la radicalizzazione. A tal proposito, va subito detto che non esiste una definizione universalmente condivisa di radicalizzazione, per quanto non siano mancati sforzi definitori a livello internazionale[24]. Sul piano empirico, può enfatizzarsi il fine dell'azione radicalizzante, che, nel propagandare un'intonazione integralista dell'Islam, induce gli individui (non necessariamente geograficamente, culturalmente e/o religiosamente vicini all'estremismo islamico) ad abbracciare un'impostazione religiosa estremista e totalizzante, a prescindere dalla commissione di azioni criminali. L'idea è quella di far nascere un sentimento di inclusione e, quindi, un vincolo di appartenenza ad una comunità, appunto quella jihadista, in coloro che, non integrandosi nel tessuto sociale di provenienza, scoprono un'identità "nuova" nelle prospettive offerte dall'Islam radicale[25].

Orbene, l'intelligenza artificiale può essere utilizzata per identificare coloro che sono più vulnerabili e, perciò, maggiormente esposti agli sforzi radicalizzanti dei gruppi terroristici. Ad esempio, il c.d. metodo *Redirect* – sviluppato da Jigsaw, una *subsidiary* di Google – consiste in un sistema algoritmico in grado di riconoscere i profili degli individui potenziali *targets* della narrativa dell'odio. Nello specifico, si testa l'algoritmo sul discorso islamista, in modo da insegnargli i vocaboli e le frasi utilizzati dalla retorica jihadista. Una volta che si è così programmato l'algoritmo, lo si usa per intercettare l'utente di Internet sensibile al discorso terroristico, allo scopo di "re-indirizzarlo" a pagine *web* che contengano la c.d. contro-narrativa. Quest'ultima dovrebbe essere capace di deflazionare la possibilità di adesione ad ideologie violente o comunque estremiste.

---

[22] Su questi aspetti, v. K. MCKENDRICK, *Artificial Intelligence Prediction and Counterterrorism*, in Chatham House, (copiare il link sul browser) 9 August 2019. Sulla possibilità di errore, v. *infra* nel presente lavoro.

[23] Sul principio di proporzionalità e sul suo ruolo nel giudizio di bilanciamento tra valori, si veda, per tutti, A. STONE SWEET-J. MATHEWS, *Proportionality Balancing and Constitutional Governance. A Comparative and Global Approach*, Oxford University Press, Oxford, 2019.

[24] V., ad esempio, l'*Annual report of the United Nations High Commissioner for Human Rights and reports of the Office of the High Commissioner and the Secretary-General*, 21 July 2016, A/HRC/33/29, che definisce la radicalizzazione come «a process through which an individual adopts an increasingly extremist set of beliefs and aspirations. This may include, but is not defined by, the willingness to condone, support, facilitate or use violence to further political, ideological, religious or other goals».

[25] Sui legami tra radicalizzazione e marginalizzazione sociale, E. SANTORO, *La* trappola *dell'identità culturale: dal multiculturalismo alla radicalizzazione*, in G. CERRINA FERONI-V. FEDERICO (a cura di), *Strumenti, percorsi e strategie dell'integrazione nelle società multiculturali*, Edizioni Scientifiche Italiane, Napoli, 2019, 87 ss.

Una seconda forma di "tecnologia intelligente" utile nella lotta al terrorismo internazionale è offerta dalla capacità degli algoritmi di procedere al riconoscimento facciale[26], testato, in via ancora sperimentale[27], dalle autorità di *law enforcement* di vari Paesi, non solo europei, a fini antiterroristici e, più in generale, per attività di prevenzione del crimine. Come noto, a scopi identificativi, i sistemi di riconoscimento facciale estraggono e processano dati biometrici del viso di individui (raccolti grazie a telecamere poste in luoghi pubblici) per confrontarli con altri volti contenuti in *databases* di persone sospettate di terrorismo oppure effettivamente già condannate per questo tipo di reati[28].

Un terzo impiego dell'intelligenza artificiale si deve agli obblighi fissati dalle normative finalizzate a prevenire il riciclaggio del denaro e, in particolare, il finanziamento del terrorismo. Più nello specifico, il regolamento (UE) 2015/847[29] e la direttiva (UE) 2015/849[30] prescrivono alle banche e agli intermediari finanziari precisi doveri di monitoraggio sulle transazioni finanziarie poste in essere dai titolari di conto corrente, per evitare che vengano portati a compimento i reati appena menzionati. Le disposizioni euro-unitarie, invero, non richiedono che questi obblighi vengano adempiuti necessariamente impiegando strumenti di intelligenza artificiale; ciononostante, gli enti creditizi e gli intermediari di maggiori dimensioni hanno ritenuto vantaggioso investire in sistemi algoritmici in grado di tracciare le transazioni e rilevare spostamenti di denaro sospetti[31].

Un'ulteriore applicazione dell'intelligenza artificiale si basa sulla capacità degli algoritmi di analizzare i metadati delle comunicazioni che avvengono *online*[32], in modo da individuare comportamenti potenzialmente pericolosi e segnalarli alle competenti autorità pubbliche. A tal riguardo, merita di essere menzionata l'esperienza francese, dove la *loi* 2015-912[33] permette ai servizi di *intelligence* di utilizzare questo tipo di algoritmi, detti *boîtes noires*, su autorizzazione del

---

[26] Si veda lo studio della FUNDAMENTAL RIGHTS AGENCY, *Facial Recognition Technology: Fundamental Rights Consideration in the Context of Law Enforcement*, 21 November 2019.

[27] Nel campo dell'immigrazione, invece, questa tecnologia risulta già largamente in uso per i riconoscimenti alle frontiere. Regolamento (UE) 2018/1861 del Parlamento europeo e del Consiglio, del 28 novembre 2018, sull'istituzione, l'esercizio e l'uso del sistema d'informazione Schengen (SIS) nel settore delle verifiche di frontiera, che modifica la convenzione di applicazione dell'accordo di Schengen e abroga il regolamento (CE) n. 1987/2006, in *G.U.*, L 312, del 7.12.2018, 14–55.

[28] La prima pronuncia di una corte sull'utilizzo di un sistema di riconoscimento facciale si è avuta da parte della High Court of Justice of England and Wales, che, nel settembre 2019, ha esaminato l'utilizzo, nell'ambito di grandi eventi sportivi, di un sistema di riconoscimento facciale da parte della polizia locale del Galles del Sud. *R(Bridges) v. The Chief Constable of South Wales Police et al.*, [2019] EWHC 2341. V., per un primo commento della decisione, A. PIN, *Non esiste la "pallottola d'argento": l'*Artificial Face Recognition *al vaglio giudiziario per la prima volta*, in *DPCE Online*, 4/2019, 3075 ss.

[29] Regolamento (UE) 2015/847 del Parlamento europeo e del Consiglio, del 20 maggio 2015, riguardante i dati informativi che accompagnano i trasferimenti di fondi e che abroga il regolamento (CE) n. 1781/2006, in *G.U.*, L 141, del 5.6.2015, 1–18.

[30] Direttiva (UE) 2015/849 del Parlamento europeo e del Consiglio, del 20 maggio 2015, relativa alla prevenzione dell'uso del sistema finanziario a fini di riciclaggio o finanziamento del terrorismo, che modifica il regolamento (UE) n. 648/2012 del Parlamento europeo e del Consiglio e che abroga la direttiva 2005/60/CE del Parlamento europeo e del Consiglio e la direttiva 2006/70/CE della Commissione, in *G.U.*, L 141, del 5.6.2015, 73–117.

[31] In base alla normativa citata, i soggetti in questione devono altresì riferire tali movimenti alle *financial intelligence units* (FIUs), istituite presso ciascuno Stato membro. Le FIUs sono state introdotte negli Stati membri dell'UE ai sensi della precedente disciplina in tema di antiriciclaggio, ossia la Direttiva 2005/60/CE del Parlamento e del Consiglio, del 26 ottobre 2005, relativa alla prevenzione dell'uso del sistema finanziario a scopo di riciclaggio dei proventi di attività criminose e di finanziamento del terrorismo, in *G.U.*, L 309, 25.11.2005, 15–36. Essa è stata poi abrogata dalla successiva Direttiva (UE) 2015/849 (cit., nota 30), che perpetua l'obbligo, per tutti gli Stati membri, di mantenere una FIU.

[32] Ad esempio, gli indirizzi IP utilizzati, i siti *web* visitati. V. *supra*, nota 15. Sul punto, F. TRÉGUER, *From Deep State Illegality to Law of the Land: The Case of Internet Surveillance in France*, in *Archive ouverte en Sciences de l'Homme et de la Société*, 2016.

[33] *Loi* n° 2015-912 *du* 24 *juillet* 2015 *relative au renseignement*. V. M-A. GRANGER, *Oversight of the State Emergency in France*, in B.J. GOOLD-L. LAZARUS (eds.), *Security and Human Rights*, Hart Publishing, Oxford, 2019, 389 ss.

Primo Ministro, sentito il parere della *Commission Nationale de l'Informatique et des Libertés,* ancorché a prescindere da qualsiasi vaglio giurisdizionale di tipo preventivo.

L'ultimo settore di applicazione degli algoritmi, sempre in un'ottica *counter-terrorism*, è quello della rimozione dei messaggi a sfondo terroristico, che vengono quotidianamente diffusi attraverso la rete[34]. Data la difficoltà di esercitare un continuo ed efficace monitoraggio "umano" della narrativa dell'odio in generale, e di quella terroristica in particolare, *service providers*, *hosting providers*, *social networks* e motori di ricerca stanno investendo risorse e ricerca in apparati tecnologici in grado di riconoscere i contenuti terroristici, violenti e/o pericolosi[35], così da rimuoverli, in adempimento alle proprie *policies* interne e – a partire da tempi recenti – anche in ottemperanza di atti normativi[36]. Tale scenario merita di essere indagato da un punto di vista giuridico, soprattutto alla luce della possibile limitazione di diritti fondamentali – quale, ad esempio, quello della libertà di espressione – che può derivare da interventi legislativi in tal senso.

3.1. La rimozione dei contenuti terroristici *online*: dalle *partnerships* su base volontaria ai primi approcci regolatori

Per esaminare l'utilizzo delle tecniche "intelligenti" di intercettazione e rimozione dei contenuti terroristici *online*, al fine di indagarne l'impatto sui principi basici del diritto costituzionale, è opportuno, in via preliminare, focalizzarsi sulle fonti dalle quali deriva l'obbligo di eliminazione. Invero, la parola "obbligo", in questa sede, potrebbe rivelarsi parzialmente impropria. Infatti, perlomeno in un primo momento, non si rinvenivano atti giuridici vincolanti – né a livello nazionale né, tantomeno, a quello sovranazionale[37] – che richiedessero l'espunzione di tali contenuti "pericolosi" dai motori di ricerca e dalle principali piattaforme digitali. Esistevano (e sono ancora operative), piuttosto, una serie di *partnerships*, costituite su base volontaria, tra gli operatori della *Internet Communication Technology* (ICT) e le competenti autorità pubbliche, rispetto alle quali i primi si impegnano a monitorare i contenuti da loro ospitati e, se del caso, segnalarli per garantirne la rimozione[38]. Al giorno d'oggi, pur essendo innegabile la tendenza alla "normativizzazione" di queste collaborazioni, che sarà meglio illustrata nel prosieguo, si deve osservare che si tratta di un *trend* ancora in una fase iniziale.

Tra le forme di cooperazione volontaria, la prima che va menzionata, ai fini di questo studio, è il Global Internet Forum to Counter Terrorism (GIFCT). Il GIFCT è stato istituito nel 2017, ma rappresenta la formalizzazione di una collaborazione preesistente, benché meno strutturata, fra le più importanti società dell'ICT (Facebook, Microsoft, Twitter, YouTube), poi allargata a entità istituzionali, fra cui le Nazioni Unite[39]. Nell'aderire al GIFCT, i soggetti coinvolti si impegnano alla rimozione di contenuti potenzialmente pericolosi, senza pregiudicare l'essenza del fondamentale diritto alla libera manifestazione del pensiero[40]; e, in parallelo, si dedicano a finanziare la ricerca sullo sviluppo di mezzi tecnologici finalizzati a prevenire la diffusione *online* del terrorismo. Se si

---

[34] V. *infra*, § 3.1.

[35] S. STUART MACDONALD-S. GIRO CORREIA-A.-L. WATKINS, *Regulating Terrorist Content on Social Media: Automation and the Rule of Law*, in 15 *International Journal of Law in Context*, 2019, 183 ss.

[36] V. *infra*, § 3.1.

[37] Tuttavia, possono segnalarsi dei generici riferimenti al tema della collaborazione fra autorità pubbliche e operatori del *web*. Si vedano, ad esempio, le Risoluzioni del Consiglio di Sicurezza delle Nazioni Unite, che, a far tempo dal 2013, invitano gli Stati membri ad agire contro il terrorismo *online*, anche suggerendo forme di cooperazione con il settore privato e con la società civile: Consiglio di Sicurezza delle Nazioni Unite, Risoluzione 2129/2013, 17.12.2013, S/RES/2129(2013).

[38] C. KAVANAGH et al., *Terrorist use of the Internet and cyberspace: issues and responses*, in M. CONWAY et al. (ed.), *Terrorists' Use of the Internet: Assessment and Response*, IOS Press, Amsterdam, 2017, 1 ss.

[39] Più nel dettaglio, tale cooperazione vede come *partner* Tech Against Terrorism, un programma implementato dal Counter Terrorism Executive Directorate delle Nazioni Unite.

[40]. V. *Evolving an Institution*, in Global Internet Forum to Counter Terrorism, *GIFCT*.

escludono i riferimenti esplicativi rinvenibili sul sito *web* del GIFCT, poco è conoscibile, a livello concreto, sull'organizzazione del Forum e sulle sue modalità di azione. Anzi, diversi dubbi sono sollevati tanto sulla mancanza di trasparenza quanto sulla struttura della *governance* del GIFCT, ruotata intorno ad un *executive board* in cui siedono rappresentanti dei soli membri fondatori, i quali risultano peraltro essere società private (Facebook, Microsoft, Twitter, YouTube).

Altra esperienza simile al GIFCT, ma circoscritta all'ambito euro-unitario, è l'EU Internet Forum. Quest'ultimo è stato costituito nel 2015 – prima, quindi, del GIFCT – con il preciso compito di contrastare il terrorismo *online*. L'EU Internet Forum riunisce rappresentanti dei governi degli Stati membri UE – di regola, i Ministri dell'Interno –, di Europol e delle più importanti società ICT. Si tratta di un *forum* di dialogo fra i soggetti pubblici e i privati, finalizzato a discutere le misure più efficaci ad integrare una proficua collaborazione tra le competenti autorità e gli operatori del settore, al fine di prevenire e contrastare le attività radicalizzanti[41]. Questa cooperazione di natura volontaria[42] ha, tuttavia, un significativo impatto, se si considera che, proprio nel dare seguito all'attività del Forum, dal 2015 in poi, i "giganti del *web*" hanno apportato rilevanti modifiche alle proprie *policies*, che ora prevedono, nella quasi totalità dei casi, un'intensa attività di monitoraggio, segnalazione e, conseguente, eventuale rimozione dei massaggi a contenuto radicalizzante o terroristico in senso lato.

Queste *partnerships*, basate sulla collaborazione volontaria tra aziende private e autorità pubbliche, sono presenti anche all'interno dei singoli Stati. Si pensi, ad esempio, al Regno Unito – che solo di recente si è lasciato alle spalle l'esperienza euro-unitaria –, dove specifici corpi di polizia (Counter Terrorism Internet Referral Unit – CTIRU) lavorano in collaborazione con i gestori di motori di ricerca e di piattaforme digitali, nell'intento di individuare il "rischio radicalizzazione"[43].

Va poi rilevata la tendenza alla "normativizzazione" di queste esperienze, dovuta all'adozione di leggi e di altre fonti giuridicamente vincolanti che obbligano alla rimozione dei contenuti a sfondo terroristico. Ad esempio, già dal 2017, il legislatore federale tedesco ha imposto alle società ICT l'eliminazione della narrativa terroristica dalle proprie piattaforme, a pena di pesanti sanzioni pecuniarie, in caso di mancata *compliance*[44]. Sulla stessa linea, le Istituzioni europee stanno discutendo una proposta di regolamento, presentata dalla Commissione nel settembre 2018, «relativo alla prevenzione della diffusione di contenuti terroristici online»[45]. Il testo, attualmente sottoposto alla prima lettura del Consiglio, assicura ampi margini di autonomia ai "giganti della tecnologia", che sono chiamati non solo a monitorare i contenuti, ma pure, se del caso, a rimuoverli proattivamente[46].

È presumibile che da tale approccio normativo, nazionale e sovranazionale, finirà per rafforzare le menzionate *partnerships*, come il GIFCT e l'EU Internet Forum, la cui attività non rappresenterà più un'autonoma scelta dei vari operatori, bensì l'implementazione di un obbligo giuridico.

Tratto comune di tutte le forme di cooperazione – da quelle che dipendono dal mero consenso degli *stakeholders* a quelle "normativizzate" – è il non trascurabile margine discrezionale lasciato agli operatori tecnologici nelle modalità impiegate per la rilevazione dei contenuti da attenzionare e, eventualmente, eliminare. Di conseguenza, questi soggetti privati, anche sulla spinta del ruolo

[41] L'EU Internet Forum, in effetti, lavora a stretto contatto con il Radicalization Awareness Network, un *hub* che riunisce esperti di radicalizzazione impegnati ad implementare misure di prevenzione e di de-radicalizzazione.

[42] Si tratta di una misura adottata nell'ambito della strategia antiterrorismo dell'Unione europea approvata nel 2015. V. *Risoluzione del Parlamento europeo dell'11 febbraio 2015 sulle misure antiterrorismo* (2015/2530(RSP).

[43] Sul tema, si veda C. WALKER, *Blackstone's Guide to the Anti-Terrorism Legislation*, Oxford University Press, Oxford, 2014.

[44] *Netzwerkdurchsetzungsgesetz* (2017). T. VAN BENTHEM, *Social Media Actors in the Fight against Terrorism: Technology and its Impact on Human Rights*, in 7 *Cambridge International Law Journal*, 2018, 284 ss.

[45] COM (2018) 640 final.

[46] Per un'analisi critica della bozza di regolamento, M. SCHEININ, *The EU Regulation on Terrorist Content: An Emperor without Clothes*, in Verfassungsblog, 30 January 2019.

centrale che la prevenzione del terrorismo gioca nel discorso socio-politico contemporaneo[47], si sono impegnati in un uso efficiente della tecnologia a loro disposizione. In particolare, hanno introdotto forme di automazione (inclusa quella "intelligente", perché capaci di imparare dalla propria esperienza) volte ad adempiere agli obblighi – o perlomeno, a rispettare gli impegni – sottoscritti. Del resto, data l'enorme quantità di contenuti quotidianamente postati sulle diverse piattaforme, non ricorrere a forme di intelligenza artificiale significherebbe, di fatto, rinunciare alla possibilità di un controllo capillare su quanto condiviso dagli utenti. Sul piano empirico, i dati più aggiornati mostrano l'utilità di questi meccanismi: negli ultimi sei mesi del 2018, ad esempio, YouTube ha rimosso circa 150,000 video contenenti riferimenti al terrorismo e il 98% di essi era stato segnalato dagli algoritmi a disposizione della piattaforma[48].

È infine interessante notare che gli algoritmi impiegati dai principali motori di ricerca e *social media* sono di varia natura; e, benché lo studio che si propone non abbia un carattere tecnico, si ritiene comunque utile richiamarli sinteticamente, allo scopo di una migliore comprensione dei problemi di stampo giuridico che essi possono innescare. Una prima tecnica è quella dell'*image matching*: se un utente carica un'immagine, sia essa una foto o parte di un *link*, dotata di caratteristiche comuni con un'altra rappresentazione grafica che il sistema ha "imparato" essere di tipo terroristico, l'immagine in questione viene segnalata (*flag*) e, automaticamente, "candidata" alla rimozione. Vi è poi il c.d. *language understanding*: l'algoritmo viene programmato per essere in grado di identificare alcune parole e sintagmi ricorrenti in scritti di tipo terroristico ed è capace di riconoscerli nel caso in cui appaiano nei *posts* pubblicati dagli utenti. Una terza tecnica è il *terrorist cluster*: l'intelligenza artificiale segnala profili che hanno legami con potenziali terroristi: ad esempio, perché stringono collegamenti via Facebook con utenze che sono state rimosse per propaganda terroristica. Alcuni algoritmi sono, addirittura, in grado di individuare la recidiva, ossia riescono a comprendere quando un soggetto, il cui *account* è già stato disabilitato, perché ritenuto "pericoloso", crea altri profili "fittizi", così da riproporsi sulla piattaforma *online*, nonostante la precedente esclusione[49].

3.2. Intelligenza artificiale, *counter-terrorism measures* e diritto costituzionale: una triangolazione "pericolosa"?

Le tecniche appena descritte, adottate sia nell'ambito della cooperazione volontaria sia in adempimento di obblighi fissati da fonti giuridiche vincolanti – già in vigore o in corso di adozione nelle competenti sedi[50] – pongono diversi spunti di riflessione. In particolare, da una prospettiva pubblicistica, sembrano emergere potenziali "punti di attrito" tra l'innovazione tecnologica e il diritto costituzionale.

In prima battuta, è il caso di rilevare che tutti gli scenari descritti – dal GIFCT alla "normativizzazione" degli obblighi di *screening* del *web* e, conseguente, rimozione dei contenuti "pericolosi" – vedono al centro la necessaria cooperazione fra autorità pubbliche e operatori privati del settore (i c.d. giganti del *web*), che rispondono alla classica logica di mercato[51]. In quest'ottica,

[47] A tal proposito, è emblematico il discorso pronunciato da Theresa May durante il World Economic Forum 2018. L'ex Primo Ministro britannico rilevava la necessità «to move further and faster in reducing the time it takes to remove terrorist content online», anche «automatically».

[48] S. WOJCICKI, *Expanding our work against abuse of our platform*, YouTube Official Blog, 4 December 2017.

[49] Su queste tecniche, si rinvia a S. MACDONALD et al., *Regulating terrorist content on social media: automation and the rule of law*, in 15 *International Journal of Law in Context*, 2019, 183 ss.

[50] Si pensi alla proposta di regolamento dell'Unione europea, che, se adottata, armonizzerà le legislazioni degli Stati membri sul punto. V. *supra*, nn. 45 e 46.

[51] La massimizzazione del profitto, anziché dell'interesse pubblico. A tal proposito, non mancano però le teorie di coloro che che, nel rilevare il non coinvolgimento degli operatori di settore nelle "agende securitarie" di legislatori e governi, sostengono la possibilità di un approccio oggettivo e asettico dei privati nei confronti del bilanciamento

va segnalata una sorta di inesorabile "privatizzazione" di attività tipicamente di spettanza delle pubbliche autorità, peraltro particolarmente delicate, quali quelle della prevenzione e della repressione del crimine.

Come anticipato, questa "esternalizzazione" risulta, da un certo punto di vista, inevitabile: le autorità pubbliche non sempre dispongono della tecnologia necessaria né delle adeguate risorse finanziarie per un attento e costante monitoraggio del *web*; e, anche nel caso in cui gli strumenti tecnici a loro disposizione consentissero, in astratto, tale attività, non avrebbero la possibilità di accedere agli spazi virtuali messi a disposizione della rete, senza una previa autorizzazione giudiziaria. Nello sforzo di cooperazione finalizzato al contrasto del crimine, il raggio di azione accordato al settore privato dal potere pubblico è nondimeno piuttosto ampio. In effetti, persino i – pochi – atti normativi che affrontano il problema presentano tratti di vaghezza e genericità, da cui deriva il sistematico rinvio alle *policies* interne, adottate da attori come Google, Facebook e dagli altri operatori della rete.

Tra gli aspetti sui quali viene lasciata discrezionalità agli attori privati vi è, ad esempio, il profilo definitorio. Come è noto, l'individuazione dei caratteri propri del concetto di "terrorismo" occupa gli studiosi da ben prima dell'11 settembre 2001 e, ad oggi, non esiste una definizione universalmente accettata del fenomeno, né a livello di diritto positivo né da parte degli studi dottrinali[52]. Questa incertezza, se non vera e propria lacuna, lascia un non trascurabile margine alle piattaforme digitali: sono i loro algoritmi a decidere se e quando un contenuto presenta *standards* di pericolosità, ad esempio perché propaganda l'islamismo jihadista[53], e vada, di conseguenza, rimosso.

Abbandonare questa delicata operazione definitoria a soggetti privati significa però affidare loro un ruolo che può essere accostato al *law-making*: agli operatori di settore viene infatti consentito di individuare (*rectius*, delineare) le fattispecie alle quali si ricollega un effetto limitativo di diritti (anzi tutto, la libera manifestazione del pensiero). Si è davanti ad un'attività che si caratterizza, da un punto di vista sostanziale, come "para-legislativa".

Quella "para-legislativa" non è però l'unica funzione tipicamente pubblicistica che viene "delegata" a soggetti di natura privatistica. Nell'agire proattivamente – come peraltro espressamente richiesto, ad esempio, dalla proposta di regolamento dell'Unione europea, sopra menzionata – per la rimozione di *posts* e altro materiale a sfondo terroristico diffuso in rete, motori di ricerca, *social media* e altri attori del *web* assolvono a funzioni, *de facto*, di tipo esecutivo. Tuttavia, nella concezione classica del diritto costituzionale, qualsiasi azione dell'esecutivo e degli organi da esso dipendenti deve ispirarsi, almeno in linea generale[54], a canoni di trasparenza e di pubblicità, affinché sia garantito il vaglio tanto delle corti, viste come necessario "contropotere", chiamato a frenare possibili abusi e/o derive dei poteri politicamente sensibili, quanto dell'opinione pubblica,

---

sicurezza-libertà, con il minor rischio di sacrificare il secondo elemento del binomio a vantaggio del primo. In questo senso, J.D. MICHAELS, *All the President's Spies: Private-Public Intelligence Partnerships in the War on Terror*, in 96 *California Law Review*, 2008, 924 ss.

[52] Con riferimento al *pre*-9/11, si veda B. GANOR, *Defining Terrorism: Is One Man's Terrorist Another Man's Freedom Fighter?*, in 3 *Police Practice and Research*, 2002, 287 ss. In relazione ai tentativi definitori *post*-9/11, M. SCHEININ-M. VERMEULEN, *Unilateral Exceptions to International Law: Systematic Legal Analysis and Critique of Doctrines that Seek to Deny or Reduce the Applicability of Human Rights Norms in the Fight against Terrorism*, EUI Law Working Papers, 2010/08, 20 ss.

[53] Ad esempio, Twitter definisce terroristici le organizzazioni che «identify through their stated purpose, publications, or actions as an extremist group; have engaged in, or currently engage in, violence and/or the promotion of violence as a means to further their cause; and target civilians in their acts and/or promotion of violence» (Twitter, *Terrorism and Violent Extremism Policy*, updated March 2019). Nell'ambito della stessa *policy*, Twitter si impegna a identificare e rimuovere i relativi *posts*.

[54] Ferma l'eccezione del segreto di Stato, tradizionalmente considerato punto di tensione tra le esigenze di trasparenza dell'azione pubblica e la sicurezza nazionale. Sul punto, sia consentito rinviare ad A. VEDASCHI, *The Dark Side of Counter-Terrorism:* Arcana Imperii *and* Salus Rei Publicae, in 66 *The American Journal of Comparative Law*, 2018, 877 ss.

alla quale i poteri dello Stato, in ultima istanza, devono rispondere, almeno negli assetti democratici.

Orbene, nel caso in cui un motore di ricerca provveda alla rimozione di un contenuto digitale segnalato da un algoritmo "intelligente", tale trasparenza non viene garantita, giacché il procedimento seguito per decidere l'eliminazione del messaggio resta coperto dal segreto industriale. In altre parole, il funzionamento concreto dell'algoritmo non è mai reso interamente noto, poiché fa parte dell'assetto proprietario dell'azienda che lo ha elaborato. Dunque, se *l'iter* logico seguito dal decisore non è trasparente, è chiaro che l'eventuale contestazione dei suoi passaggi è oltremodo difficile, se non impossibile. Di conseguenza, altrettanto complesso risulta essere lo scrutinio, da parte di un giudice, circa la legittimità delle misure prese in esito al processo decisionale della macchina[55].

Infine, persino la funzione giurisdizionale pare essere toccata dal progressivo "assorbimento" nell'orbita dei "giganti del *web*". Di recente, Facebook ha annunciato l'imminente creazione di un *Oversight Board*[56], che dovrebbe entrare in funzione dai primi mesi del 2020. L'*Oversight Board* dovrà pronunciarsi sulle doglianze di utenti che lamentano l'ingiusta eliminazione di contenuti postati sulla rete. I membri di questo organo saranno, seppur indirettamente[57], nominati dal consiglio di amministrazione di Facebook. Pertanto, la composizione non assicura i caratteri di terzietà, indipendenza e imparzialità, propri invece della funzione giurisdizionale tradizionalmente intesa.

Sembra, dunque, che i poteri riconducibili alla classica triade montesquieuiana siano tutti, sebbene con diverse modalità e in differente misura, in qualche maniera "esternalizzati" e "delegati" a soggetti privati (gli operatori della rete), con le criticità sopra accennate.

Ne consegue, in linea generale, una chiara torsione dell'idea di sovranità, che cessa di configurarsi come concetto riconducibile *in toto* allo spazio pubblico, allo Stato, per diventare invece oggetto di una parziale "condivisione" tra pubblico e privato, se non di una vera e propria "cessione" da parte delle autorità statali a favore del settore ICT.

Last but not least, preme rilevare il ruolo fondamentale giocato dalle linee guida e dalle *policies* interne adottate dagli attori digitali. In effetti, la funzione para-normativa affidata, non *de jure* ma perlomeno *de facto*, agli operatori della rete – evidente in relazione sia al profilo definitorio del "terrorismo" sia a quello operativo riguardante le modalità di rimozione dei contenuti terroristici *online* – mette sotto *stress* il concetto classico di fonte del diritto. La non vincolatività delle linee guida nei confronti della generalità dei consociati sembra ricondurle alla categoria della *soft law*. Cionondimeno, la "paternità" della *soft law* dovrebbe, almeno secondo le teorie classiche, essere

[55] L'algoritmo è notoriamente soggetto al rischio di errore. L'errore algoritmico può dipendere da diversi fattori. In primo luogo, possono esistere *biases* interni al sistema, ossia l'algoritmo potrebbe essere stato costruito in modo discriminatorio nei confronti di particolari categorie di persone. In secondo luogo, l'errore potrebbe dipendere da malfunzionamenti tecnici; ad esempio, si è dimostrato che, nel caso del riconoscimento facciale, l'algoritmo distingue con maggiore difficoltà visi di persone di colore. Sui rischi di discriminazione rispetto al colore della pelle arrecati da algoritmi di riconoscimento facciale – utilizzati, però, nel settore commerciale – si rinvia a J. BUOLAMWINI-T. GEBRU, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, in 81 *Proceedings of Machine Learning Research*, 2018, 1 ss. Sui possibili *biases* dei sistemi di intelligenza artificiale, *Houston Federation of Teachers v. Houston Independent School District*, United States District Court, S.D. Texas, Houston Division, 251 F.Su3d 1168 (2017), che giudica discriminatorio un *software* algoritmico utilizzato per la valutazione delle *performance* degli insegnanti. V., inoltre, Supreme Court of Wisconsin, *State of Wisconsin v. Eric L. Loomis*, Case no. 2015AP157-CR, 5 April – 13 July 2016, in cui l'utilizzo di un algoritmo predittivo in ambito giudiziario penale viene invece ritenuto legittimo e non foriero di discriminazioni.

[56] Si veda il *post* di M. Zuckerberg, *A blueprint for content governance and enforcement*, 15 November 2018. L'*Oversight Board* è stato poi giornalisticamente chiamato "Facebook Supreme Court". Per una riflessione sui rischi derivanti dall'introduzione dell'*Oversight Board*, Q. WEINZIERL, *Difficult Times Ahead for the Facebook „Supreme Court"*, in *Verfassungsblog*, 21 September 2019.

[57] Grazie alla nomina di un *trust* che sceglie i membri dell'*Oversight Board*, ma i cui componenti sono, a loro volta, nominati dai vertici della società.

ricondotta ad un soggetto pubblicistico[58]. Diversamente, nel caso in esame, si è davanti ad una sorta di "*soft law* dei privati"[59], che, nello scenario di cui ci si occupa, impatta però sulle libertà personali e sui diritti ai quali si riconosce, nella generalità degli ordinamenti, rango costituzionale. Da questa prospettiva, si può forse intravedere un ripensamento del sistema delle fonti del diritto, non più letto seguendo una logica rigorosamente gerarchico-formale, ma ora caratterizzato da tendenze sostanzialistiche. In altre parole, la prevalenza della fonte di rango più elevato, rispetto a quelle gerarchicamente subordinate, sembra essere sempre meno scontata. Diversamente, a far premio sulle altre potrebbe essere la fonte che, da un punto di vista concreto, permette (effettivamente o almeno nella percezione collettiva) una più efficace tutela del bene "sicurezza" – nel caso di specie, eliminando il messaggio violento – sia pure essa un atto di "*soft law* dei privati".

4. *Osservazioni conclusive*

Dall'applicazione dell'intelligenza artificiale al *counter-terrorism* e, in particolare, dall'azione di rimozione dei contenuti terroristici dal *web* discende una non trascurabile serie di questioni sul piano del diritto costituzionale.

Su un piano generale, emerge subito che queste questioni sono collegate; anzi, a volte, addirittura consequenziali e dovute a dati di fatto difficilmente modificabili. Come si è visto, il ricorso a tecnologie altamente raffinate in questo settore rende imprescindibile la cooperazione tra il potere pubblico e il settore privato. Dalla conseguente "privatizzazione" di attività tipicamente pubblicistiche discende l'inedita torsione di tradizionali concetti del diritto costituzionale. In ultima analisi, è interessata da questo processo distorsivo la sovranità (v. *supra*). Del resto, anche la nozione di potere pubblico assume un'accezione prettamente oggettiva-funzionalistica, ossia non si qualifica come potere pubblico il soggetto che *presenta caratteristiche* pubblicistiche, bensì colui che *svolge funzioni* concretamente pubblicistiche[60]. In parallelo, questa situazione impatta sulla sistematica delle fonti del diritto, che viene ad essere riletta in una logica sostanzialistica, piuttosto che gerarchico-formalistica.

Orbene, lo scenario descritto comporta un alto margine di rischio per l'effettiva tutela dei diritti della persona; in effetti, non solo si lascia ampio raggio di azione ad operatori privati, che ragionano secondo logiche di mercato, anziché in un orizzonte di massimizzazione dell'interesse pubblico, ma si permette altresì una "de-umanizzazione" delle funzioni, grazie all'impiego, sempre più massiccio, dell'algoritmo, che non risponde ai caratteri di trasparenza, pubblicità e *accountability* propri del potere pubblico.

Alla luce di queste osservazioni, sembrano allora scontate le intuibili riserve del costituzionalista, che si interroga sull'opportunità (o meno) di ricorrere all'"algoritmo intelligente" per monitorare, identificare e, eventualmente, rimuovere i contenuti "pericolosi" dal *web*. Un atteggiamento di assoluta chiusura, tuttavia, non sembra realisticamente proponibile, in quanto il rinunciare all'impiego di questo tipo di strumentazione tecnologica rappresenterebbe un "passo indietro" che accorderebbe ai terroristi un'inaccettabile posizione di vantaggio. In via di fatto,

[58] Si pensi alle raccomandazioni della Commissione europea, fonti del diritto UE non giuridicamente vincolante, ma riconducibili ad un'Istituzione euro-unitaria.

[59] Sull'uso della *soft law* nell'ambito delle misure antiterrorismo, si veda ONU, Assemblea Generale, *Report of the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism*, 29.8.2019, A/74/335.

[60] Sembrano quindi essere superate teorie strettamente formalistiche sui poteri dello Stato, che identificavano come "pubblica" ogni funzione svolta da un soggetto formalmente "pubblicistico", a prescindere dal suo contenuto materiale. V., in questo senso, le teorizzazioni dei pubblicisti tedeschi di fine Ottocento e inizio Novecento, come Laband e Jellinek. P. LABAND, *Das Staatsrecht des deutschen Reiches*, Mohr, Tübingen, 1911; G. JELLINEK, *Allgemeine Staatslehre*, Springer, Berlin, 1929.

risulterebbe peraltro impossibile, per l'occhio umano, scandagliare decine di migliaia di contenuti in modo attento ed efficace, cioè minimizzando il rischio di errore.

Pertanto, in linea con quanto si legge tra le righe del Parere espresso dalla Corte di giustizia in tema di *Passenger Name Record* (PNR)[61], opporsi fermamente all'impiego delle più avanzate tecnologie sembra utopistico oltreché improduttivo nel mondo d'oggi, costantemente minacciato dagli attacchi terroristici, ma profondamente trasformato dall'avanzamento della tecnologia.

Cionondimeno, nel procedere ad un consapevole e legittimo uso della tecnologia, soprattutto nella sua forma più avanzata rappresentata dagli "algoritmi intelligenti", si dovrebbero osservare alcuni *caveat*. In un approccio di cauta apertura, si può osservare che la cooperazione fra pubblico e privato basata su atti normativi, di rango primario, sembra da preferirsi alle *partnerships* volontarie non strettamente regolate. Gli atti normativi sono infatti caratterizzati da una maggiore conoscibilità, il che implica una maggiore controllabilità da parte dell'opinione pubblica, con risvolti positivi sul piano della certezza del diritto e della prevedibilità delle conseguenze dei comportamenti di ciascun individuo. Inoltre, questi atti possono essere oggetto di scrutinio penetrante ad opera delle corti; si pensi ai giudizi di costituzionalità dinanzi agli organi di giustizia costituzionale interni, ma pure a rinvii pregiudiziali e ricorsi per annullamento, qualora si ragioni con riferimento all'Unione europea. Lo stesso non si può dire in relazione alle linee guida e *policies* adottate in autonomia dai "giganti del *web*".

Vieppiù, gli atti normativi in parola dovrebbero rispondere a requisiti stringenti in termini di contenuto. A cominciare dai profili definitori, sarebbe, ad esempio, opportuno l'inserimento di definizioni uniformi, chiare e precise, *in primis* circa il concetto di contenuto terroristico. Il perfezionamento del profilo definitorio, all'interno delle fonti normative, risolverebbe anche il problema dell'attuale frammentazione delle diverse definizioni al momento presenti nelle linee guida dei vari operatori del *web*. Adottare una definizione omogenea dei concetti menzionati eviterebbe disparità di trattamento dell'utente, che, allo stato dei fatti, si trova a poter essere censurato su una certa piattaforma, mentre potrebbe essere lasciato libero di esprimersi in caso di pubblicazione dello stesso contenuto in un altro "luogo virtuale". Inoltre, la maggiore precisione sarebbe altresì auspicabile con riguardo sia alle tecniche di rimozione da utilizzare sia ai requisiti minimi di conoscibilità dell'algoritmo operante in rete.

Infine, sulla considerazione che atti di carattere generale e astratto si prestano difficilmente, per loro stessa natura, a disciplinare dettagli tecnici, soprattutto in termini operativi, si potrebbe poi pensare ad un potenziamento delle fonti secondarie "di dettaglio" – sempre di matrice pubblicistica – nello sforzo di ridurre il più possibile lo spazio "pericolosamente" lasciato alle *policies* degli enti privati.

Questi accorgimenti, lungi dall'essere mere disquisizioni di tipo formalistico, potrebbero rappresentare un primo passo, affinché l'inevitabile ricorso all'intelligenza artificiale venga mantenuto all'interno del quadro costituzionale (o, se si preferisce ragionare in un'ottica di *common law*, nell'ambito del rispetto della *rule of law*).

[61] Cfr. *supra*, nn. 9 e 10 e riferimenti ivi citati.