Università degli Studi della Campania *Luigi Vanvitelli*

MATERIALI E STUDI DI DIRITTO PUBBLICO

COLLANA DIRETTA DA

GIULIANO BALBI - LORENZO CHIEFFI

GIULIANO BALBI - FEDERICA DE SIMONE
ANDREANA ESPOSITO - STEFANO MANACORDA
(*a cura di*)

DIRITTO PENALE E INTELLIGENZA ARTIFICIALE

"NUOVI SCENARI"

IUSTITIAM COLIMUS

G. Giappichelli Editore – Torino

Prefazione

MARIAVALERIA DEL TUFO

Questo volume, agile e interessante, è curato da Giuliano Balbi, Andreana Esposito, Stefano Manacorda e da Federica De Simone, che da qualche anno esplora con competenza e passione le interferenze tra nuove tecnologie e diritto penale. Il lavoro è il risultato di un progetto scientifico portato avanti da un gruppo di ricercatori dell'Università degli Studi della Campania "Luigi Vanvitelli" in rapporto al tema peculiare dei *corporate crimes*, arricchito da contributi di prestigiosi studiosi esterni che hanno partecipato al più ampio e vivace dibattito instauratosi all'interno dell'Ateneo su tale *focus* di ricerca, analizzato anche in una prospettiva più ampia.

La promozione di attività di studio e di confronto è infatti principalmente finalizzata a riflettere sugli scenari che le nuove tecnologie aprono alla scienza, alla politica e al diritto. Continuamente rafforzato da scoperte e nuove modalità applicative, l'attuale strumentario scientifico e tecnologico pone problemi di gestione e controllo, generando nel contempo accettazione e diffidenza.

Soprattutto nei mesi della pandemia, l'applicazione generalizzata, anche all'interno di istituzioni pubbliche, di mezzi informatici ha reso definitivamente evidenti e acquisiti i vantaggi che l'uso virtuoso della tecnologia può apportare alla funzionalità di un sistema; ha familiarizzato milioni di persone con i nuovi strumenti rendendoli utilizzabili su larga scala e ha aperto strade da cui non appare opportuno tornare indietro, apparendo opzione più saggia – e irrinunciabile – l'integrazione di modalità, procedimenti ed esperienze.

Tuttavia è difficile pensare che l'utilizzo delle nuove tecnologie, una volta rese accessibili, disponibili, meglio conoscibili e più fruibili da un elevato numero di consociati, possa essere limitato o messo a frutto soltanto in una prospettiva orientata a perseguire e osservare regole già consolidate. Lo strumento in sé è neutro ma suscettibile di impieghi insidiosi, e dovrebbe essere rimessa alla responsabilità, alla affidabilità e alla sensibilità democratica di coloro che ne fanno uso la capacità di valutare potenzialità e rischi della sua utilizzazione stabilendone di volta in volta limiti e condizioni di fruizione. Un uso sapiente dovrebbe comunque permettere di modificare le regole senza soverchi timori e senza scardinare i principi, in modo da usufruire delle opportunità limitando la portata del

pericolo di applicazioni improprie. O si tratta di un cammino senza ritorno, destinato a non essere illuminato dai principi, ma capace di travalicarli al servizio di altri assetti sociali basati sul controllo?

Va infatti riconosciuto che le prestazioni offerte sono di tale rilievo da aver già iniziato a produrre mutamenti profondi che investono non solo il mondo esterno ma anche il nostro approccio ad esso, determinando il nascere di nuove sensibilità e nuove visioni. Il che non è necessariamente un male, ma soltanto un passaggio con antecedenti storici similari, riscontrabili in occorrenza di mutate condizioni di contesto.

Venendo a quel che più direttamente ci compete, le innovazioni tecnologiche, già apportate con successo in alcuni settori e allo studio in altri, stanno dando vita anche a timori e inquietudini. Tuttavia, mi chiedo come un ragazzino di oggi, che può ottenere moltissime cose con un click, tendenzialmente inserito in un modello culturale improntato all'*hic et nunc*, possa un domani tollerare i tempi lunghi della giustizia o piuttosto preferire che vengano utilizzate diverse modalità di intervento legate alla tecnologia, in grado di fornire soluzioni più performanti e immediate, anche accettando margini di errore o di rischio. E se soluzioni di questo tipo siano effettivamente perturbanti.

In altri termini, dobbiamo prendere in conto che, nell'immediato, le nuove tecnologie possono sicuramente offrire vantaggi evidenti anche se rilette, utilizzate e gestite nell'ottica di una società già stabilizzata intorno a regole generali condivise – e penso ad esempio ai controlli da parte delle forze dell'ordine di luoghi pericolosi con mezzi elettronici, a vicende endo-processuali tecnologicamente risolvibili, al supporto offerto dall'intelligenza artificiale a giudici e avvocati. Tuttavia, a medio/lungo termine, laddove la scienza avrà raggiunto risultati più avanzati e il progressivo utilizzo dei nuovi mezzi avrà modificato nel profondo rapporti, mentalità, sensibilità e modi di percepire, il discorso si porrà in modo diverso, perché sarà proprio l'*acquis* culturale a essere cambiato e a determinare le esigenze, i risultati attesi e le soglie di accettazione. E allora vanno poste oggi le basi perché tutto ciò possa essere affrontato con consapevolezza e rigore, nel pieno rispetto dei principi irrinunciabili e della dignità umana.

I rischi non mancano, anche a livello macro. Possibilità prima precluse – e sicuramente molto inquietanti – si dischiudono agli Stati: penso ad esempio all'acquisizione di dati sulla popolazione che possono diventare strumenti efficienti in termini di controllo sociale o per l'implementazione di politiche attive anche di emarginazione, esclusione o negazione di diritti fondamentali. Uno scenario già in parte attuato, con cui è urgente confrontarsi.

Una *mise en alerte* sull'uso delle nuove tecnologie è dunque importante per

bilanciare con visione razionale e serena i vantaggi che esse indubbiamente apportano e per rendere avvertiti i consociati dei rischi derivanti proprio dai mutamenti ormai fortemente radicati nel nostro contesto sociale e nel nostro vissuto. Un osservatorio implementato sul nucleo di principi non negoziabili che connota un diritto penale democratico appare allora quanto mai necessario: e questo libro, problematico e acuto, costituisce un contributo importante alla riflessione.

Introduzione

FEDERICA DE SIMONE

Il progresso tecnologico che sta caratterizzando la storia recente dell'umanità si contraddistingue per la velocità con cui si impone e, stando alla celebre espressione dell'economista Rosenberg, costituisce ancora una *scatola nera* di cui non si conoscono compiutamente i legami con l'innovazione. Ciò determina, da un lato, un senso di disorientamento rispetto alle conseguenze che ne possono derivare, dall'altro, la difficoltà di immaginare gli scenari futuri e predisporre gli strumenti per affrontarli.

Il giurista è naturalmente portato a risolvere le innovazioni classificando e regolamentando; tuttavia, si avverte anche nel mondo del diritto l'inadeguatezza degli strumenti ordinari e l'inopportunità di una visione ristretta alle categorie tradizionali. Cionondimeno, diffusa è la consapevolezza delle potenzialità e dei benefici che l'implementazione delle nuove tecnologie può apportare alla collettività, ma altrettanto diffusa è la consapevolezza della necessità di garantire la tutela dei diritti fondamentali.

Il volume contiene le riflessioni di alcune autorevoli voci in dottrina e di giovani ricercatori che hanno preso parte al Progetto di Ricerca *AI.CO.CRI 5.0: The use of AI neural networks in the fight against corporate crimes* finanziato dall'Università degli Studi della Campania Luigi Vanvitelli nell'ambito del Programma *Valere 2020*. Con questa raccolta di scritti si intende indagare i rapporti tra il Diritto penale e le nuove tecnologie, anche fornendo possibili chiavi di lettura in riferimento ad alcuni temi specifici.

In particolare, dopo aver esaminato i possibili impieghi e i diversi ruoli che l'intelligenza artificiale può rivestire nell'ambito delle attività di polizia, nella decisione giudiziaria e nella valutazione della pericolosità criminale, si affronta il tema della responsabilità penale classica nell'ipotesi in cui il reato non sia commesso direttamente dall'uomo, bensì dalla macchina.

La prospettazione di una responsabilità diretta dell'intelligenza artificiale a fronte della sua capacità decisionale, piuttosto che di una responsabilità collettiva da ravvisare in capo a coloro che hanno provveduto a crearla, programmarla e gestirla, porta al tema dei rapporti tra i nuovi sistemi e la *compliance* aziendale. Il binomio responsabilità societaria e nuove tecnologie è una questione tra le

più interessanti da indagare, non soltanto in quanto entrambe fortemente caratterizzate dal paradigma della predizione, ma anche per i possibili impieghi nelle politiche di contrasto alla corruzione. Il sistema *Zero Trust* progettato in Cina promette di essere un efficace strumento nel contrasto nella maggior parte delle ipotesi di reati contro la pubblica amministrazione. Tuttavia, gli effetti distorsivi che ne possono derivare hanno determinato i suoi stessi ideatori a sospenderne l'utilizzo.

Entrambi gli argomenti costituiscono l'oggetto principale su cui si sono sviluppate le due direttrici di indagine del progetto di ricerca e che hanno portato ad alcune valutazioni di opportunità circa le possibilità che l'impiego delle nuove tecnologie contribuisca a migliorare l'efficienza della spesa pubblica, aumentare la trasparenza e combattere proprio la corruzione.

Rimanendo in tema di responsabilità delle persone giuridiche, si è approfondita la questione dei rischi connessi all'impiego delle criptovalute rispetto alla fattispecie di riciclaggio, nella misura in cui le monete virtuali possono rappresentare uno strumento di *money laundering* utilizzato nell'interesse e a vantaggio delle società. Nel volume è riportata l'opinione contraria di chi ritiene che tale tecnologia possa costituire, invece, un modello di comportamento idoneo proprio a prevenire i fenomeni di riciclaggio, non ostando a ciò l'anonimato garantito dal mezzo informatico.

Suggestiva la prospettazione di un sistema capace di smascherare le menzogne e garantire, così, l'affermazione della verità, tra le massime ambizioni della società contemporanea. È quanto propone di fare il sistema di *memory detection* denominato a-IAT e sperimentato nel processo penale e sulla cui validità e ammissibilità nel novero delle prove scientifiche si è operata una riflessione, a seguito delle pronunce della giurisprudenza di merito sul tema.

Altra questione trattata è quella relativa al rapporto tra i benefici per la sicurezza della collettività e i rischi che ne possono derivare per la tenuta del sistema dei diritti fondamentali in tema di utilizzo massivo degli strumenti di riconoscimento facciale. Il dibattito sull'opportunità di un simile uso è molto fervido – sia in Europa, sia in Italia – e solleva non poche perplessità sull'opportunità dell'affermazione di una società del controllo, dubbi atti a incidere anche sulle scelte del legislatore.

Le finalità di predizione, prevenzione e scoperta del crimine sono sottese agli strumenti di intelligenza artificiale sempre più diffusi tra le forze dell'ordine di tutto il mondo e in dotazione anche al corpo dei Carabinieri per le tradizionali attività di indagine; l'impiego dei nuovi strumenti si sta rivelando un fondamentale ausilio, sia per le attività relative al controllo del territorio, sia per quelle strettamente connesse al contesto investigativo.

Altrettanto interessante si è rivelato il tema dell'utilizzo delle nuove tecnologie, in particolare della *blockchain*, nel contrasto al fenomeno delle frodi ali-

mentari, ambito per il quale la normativa penalistica ha sempre mostrato tutti i suoi limiti in termini di efficacia. Proprio la sinergia con gli strumenti di recente acquisizione costituisce lo spunto per una riflessione sull'opportunità di una rivisitazione e sistematizzazione dell'intera materia, in modo da garantire la salute collettiva dai possibili rischi derivanti da condotte penalmente rilevanti in tema di sicurezza alimentare.

Prima ancora di fornire una chiave di lettura rispetto ai singoli temi imposti dalla rivoluzione tecnologica, i contributi racchiusi in questo volume costituiscono sono essi stessi lo specchio del dibattito in atto. Ad oggi, infatti, sembrano contrapporsi due schieramenti, chi predilige un approccio di *esaltazione progressista* della tecnologia e, diversamente, chi assume una posizione di sfiducia tecnologica. Eppure, una visuale laica sulla questione è possibile. È necessario contemperare una riflessione epistemologica del rapporto tra la tecnica e l'uomo con una visione antropocentrica, che ponga l'uomo al centro dei processi di innovazione, permettendo di superare la cd. reificazione della tecnologia, che riduce il pensiero computazionale a un processo di *datificazione*.

«*Dobbiamo lavorare per umanizzare la tecnologia ed evitare che la tecnologia ci disumanizzi*», e se lo dice il robot *Sophia*…

CAPITOLO I

Intelligenza artificiale e diritto penale: qualche aggiornamento e qualche nuova riflessione

FABIO BASILE

SOMMARIO: 1.1. Premessa. – 1.2. Che cosa intendiamo per intelligenza artificiale? – 1.3. Primo ambito – IA e attività di *law enforcement*. – 1.3.1. RoboCop: dalla fantascienza alla realtà? – 1.3.2. Sistemi di intelligenza artificiale e polizia predittiva. – 1.4. Secondo ambito – IA e decisione giudiziaria: la macchina-giudice? – 1.5. Terzo ambito – IA e valutazione della pericolosità criminale: gli algoritmi predittivi. – 1.6. Quarto ambito – IA e responsabilità penale: così intelligente da essere "responsabile"? – 1.6.1. IA strumento del reato. – 1.6.2. IA autore del reato? – 1.6.2.1. Irriducibilmente umano?

1.1. *Premessa*

Nel presente contributo ho scelto di riprendere sinteticamente alcune tematiche già trattate in un precedente scritto uscito nel 2019 [1], integrandole con alcuni aggiornamenti e alcune nuove riflessioni maturate grazie – oltre che a nuove letture – anche al confronto con colleghi e giovani studiosi che ho potuto avere in occasione di numerosi seminari e convegni sull'IA, tenutisi negli ultimi 24 mesi.

Nelle pagine seguenti cercherò, pertanto, di re-indagare i possibili ambiti all'interno dei quali la rivoluzione tecnologica messa in moto dall'IA già solleva, o è destinata a sollevare, problemi, dubbi e questioni, rilevanti per il diritto penale:

1. le attività di *law enforcement*, in particolare le attività di cd. *polizia predittiva*;
2. i cd. *automated decision systems*, che potrebbero in futuro conoscere un

[1] F. BASILE, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Dir. pen. uomo*, 2019, p. 1 ss. Ringrazio il dott. Alessandro Carlini per i preziosi suggerimenti e per la revisione del testo che qui si offre ai lettori.

impiego anche all'interno dei procedimenti penali, sostituendo, in tutto o in parte, la decisione del giudice-uomo;

3. i cd. *algoritmi predittivi*, impiegati per valutare la pericolosità criminale di un soggetto, vale a dire la probabilità che costui commetta in futuro un (nuovo) reato;

4. infine, le possibili ipotesi di coinvolgimento – come strumento, come autore, o come vittima – di un sistema di IA nella commissione di un reato.

Il punto di partenza è, ovviamente, il medesimo dal quale partivo due anni fa: l'intelligenza artificiale è ovunque [2]. Le sue applicazioni pratiche si trovano nelle abitazioni, nelle automobili, negli uffici, nelle banche, negli ospedali, nel cielo e in internet, incluso l'"internet delle cose". Le animazioni di Hollywood, i videogiochi, i navigatori satellitari, il motore di ricerca di Google, sono tutti basati su tecniche di intelligenza artificiale. E così a proseguire [3].

È quindi facile presagire che la rivoluzione tecnologica messa in moto dall'intelligenza artificiale potrà presto significativamente impattare anche con le pretese di tutela dei beni giuridici affidate al diritto penale [4].

E noi, come giuristi, come penalisti, non possiamo farci trovare impreparati, «giacché quello che è veramente inquietante» – scriveva Martin Heidegger – «non è che il mondo si trasformi in un completo dominio della tecnica. Di gran lunga più inquietante è che l'uomo non è affatto preparato a questo radicale mutamento del mondo» [5].

Pare, pertanto, opportuno continuare a condurre la riflessione, già avviata, sulle possibili implicazioni dell'IA sul sistema della giustizia penale, al fine di non aggravare il ritardo del diritto, in particolare del diritto penale italiano, di fronte all'evoluzione tecnologica.

In effetti, come è stato efficacemente rilevato, «il progresso irrompe, non chiede permesso. E nel contesto attuale disegnare questo nuovo rapporto tra esseri umani e macchine non è per niente facile. Anche perché le tecnologie digitali hanno una velocità impressionante. Le tecnologie di ieri, come ad esempio la

[2] M.A. BODEN, *L'intelligenza artificiale*, Il Mulino, Bologna, 2019, p. 3.

[3] Per una sistematica ricognizione sugli usi attuali dell'IA e nel futuro prossimo prevedibile, vedi P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, *Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel*, Stanford, 2016, p. 18 ss.

[4] In tema di rapporti tra IA e giustizia penale v. V. MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in *DisCrimen*, 15 maggio 2020.

[5] M. HEIDEGGER, *Gelassenheit*, 1959, trad. it. di A. Fabris, *L'abbandono*, Il Melangolo, Genova, 1995, p. 36.

TV, la radio, l'elettricità, l'automobile hanno impiegato più di 50 anni per raggiungere i 50 milioni di utenti. Ci hanno concesso tutto il tempo per abituarci alle loro innovazioni, per avere nuove regole sul loro utilizzo, e per organizzare le nostre vite e le nostre società di conseguenza. Oggi, le tecnologie digitali irrompono molto più velocemente, e non ci danno affatto il tempo per organizzarci e per abituarci alle loro dirompenti innovazioni. Un esempio evidente di questa velocità viene dalle reti sociali: Twitter ha impiegato meno di 3 anni per raggiungere i 50 milioni di utenti; Facebook e Instagram meno di 2 anni. Anche se il record della velocità è quello di Pokemon Go, che è riuscito a raggiungere i 50 milioni di download in soli 19 giorni!» [6].

1.2. *Che cosa intendiamo per intelligenza artificiale?*

Prima, però, di entrare nel merito delle tematiche "penalistiche", conviene richiamare l'attenzione su alcune caratteristiche dei sistemi di IA, rilevanti ai fini della nostra indagine.

1. Innanzitutto, quando parliamo di IA non dobbiamo necessariamente pensare ad un "umanoide" simile in tutto e per tutto all'essere umano: l'umanoide può essere, sì, un'applicazione di IA (forse la più eclatante), ma di certo non l'unica e non, almeno nella fase attuale, la più rilevante dal punto di vista pratico [7]. Per contro, possiamo affermare che oggi l'IA è, principalmente, un *software*, una componente algoritmica.

2. In secondo luogo, per quanto possa essere suggestivo parlare di intelligenza artificiale, occorre rimarcare che l'intelligenza (quella degli esseri umani, prima ancora che quella delle macchine), benché oggetto di numerosissimi studi di psicologi, biologi e neuroscienziati, costituisce ancora un concetto indeterminato [8]. Ad ogni modo, l'intelligenza artificiale è mimesi, è copiatura delle presta-

[6] G.F. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, in *Agendadigitale.eu*, 11 giugno 2019, p. 3.

[7] Così C. TREVISI, *La regolamentazione in materia di Intelligenza artificiale, robot, automazione: a che punto siamo*, in *Medialaws*, 21 maggio 2018, p. 1; L. FLORIDI, *What the Near Future of Artificial Intelligence Could Be*, in *Philosophy & Technology*, 32, 2019, p. 11 ss.

[8] Si noti, per altro verso, che proprio dagli studi sull'intelligenza artificiale stanno pervenendo importanti contributi per scoprire come funziona l'intelligenza umana e il cervello umano. Si veda, ad esempio, un recente progetto europeo di integrale simulazione del cervello umano, realizzato grazie all'impiego di tecniche di IA: Redazione (a cura di), *Il progetto europeo sul cervello umano*, in *Dir. pen. uomo*, 2 aprile 2019. Dall'altra parte dell'Atlantico, un progetto USA analogo è in svolgimento: Redazione (a cura di), *L'esortazione del Presidente*, ivi, 2 aprile 2019.

zioni umane: i sistemi di IA "apprendono" per correlazioni [9], e non seguono il ragionamento deduttivo-causale, tipico dell'intelligenza umana.

3. Oggi si riconosce unanimemente che i grandi e rapidi progressi, compiuti dall'IA in tempi recenti, sono stati consentiti dalla felice combinazione di due fattori [10]:

– da un lato, il recente, impressionante aumento delle capacità computazionali, grazie alle quali disponiamo di computer sempre più veloci, potenti, con capacità di memoria (e, quindi, tra l'altro, di archiviazione dati) straordinariamente grandi;

– dall'altro lato, il recente, impressionante aumento di dati digitali, raccolti anche grazie a sensori ad alta definizione e a basso costo: dati alla cui raccolta contribuiamo ogni giorno anche noi digitalizzando documenti, scattando foto, facendo video o inviando messaggi tramite le reti sociali o altri strumenti di messaggistica.

4. La combinazione di tali due fattori – unitamente ad altri progressi nella ricerca – ha, tra l'altro, consentito di elaborare e di diffondere su larga scala i sistemi di *machine learning* che possiamo, in estrema sintesi, descrivere così: il sistema di IA "impara" autonomamente dall'ambiente esterno (tramite i dati che immagazzina ed elabora), e modifica le proprie prestazioni adattandole agli esiti del procedimento di apprendimento [11]. In altri termini, il *software* di IA programma sé stesso nel tempo in modo funzionale all'obiettivo assegnato.

[9] Questo è particolarmente vero per i sistemi di IA che fanno uso del cd. *machine learning*. Tuttavia, esiste almeno un altro grande approccio all'IA, la cd. IA Simbolica, la quale tenta di riprodurre il ragionamento umano.

[10] Così, tra i tanti, J. KAPLAN, *Intelligenza artificiale. Guida al futuro prossimo*, Luiss University Press, Roma, II ed., 2018, p. 72; G.F. ITALIANO, *Intelligenza artificiale: passato, presente, futuro*, in F. Pizzetti (a cura di), *Intelligenza artificiale, protezione dei dati personali e regolazione*, Giappichelli, Torino, 2018, p. 220; R. CALO, *Artificial Intelligence Policy: a Primer and Roadmap*, in *University of Bologna Law Review*, 3, 2, 2018, p. 186.

[11] Sul *machine learning*, v., in una prospettiva tecnica, S. RUSSELL, P. NORVIG, *Artificial Intelligence: A Modern Approach,* Chennai, III ed., 2009, p. 634 ss.; L. FLORIDI, *What the Near Future of Artificial Intelligence Could Be*, cit., p. 4 ss.; P. DOMINGOS, *L'algoritmo definitivo: la macchina che impara da sola e il futuro del nostro mondo*, Bollati Boringhieri, Torino, 2016, p. 7 ss.; K. HAO, *What is machine learning*, in *MIT Technology Review*, 17 novembre 2018; C. COLAPIETRO, A. MORETTI, *L'intelligenza Artificiale nel dettato costituzionale: opportunità, incertezze e tutela dei dati personali*, in *BioLaw Journal*, 3, 2020, p. 365 ss.; F. SUMAN, *Dove sta andando oggi l'Intelligenza artificiale?*, in *Il Bo Live*, 11 marzo 2019; in una prospettiva giuridica, R. CALO, *Artificial Intelligence Policy*, cit., p. 185; H. SURDEN, *Machine Learning and Law*, in *Washington Law Review*, 89, 1, 2014, p. 87 ss.; S. QUINTARELLI, *Forum AI and Law*, in *BioLaw Journal*, 1, 2020, p. 493 ss.

1.3. *Primo ambito – IA e attività di* law enforcement

1.3.1. *RoboCop: dalla fantascienza alla realtà?*

Probabilmente molti di noi ricordano la figura di RoboCop, il poliziotto con un corpo di titanio e kevlar, un cervello informatico e sensori ultrapotenti: se nel 1987, anno di uscita del celebre film, tale immagine apparteneva decisamente alla fantascienza («il futuro della legge» era il sottotitolo del film), oggi la realtà ci propone alcune applicazioni delle tecnologie di IA – in uso, per lo più in via sperimentale, presso le forze di polizia di alcuni Stati – che si avvicinano molto a RoboCop [12]: si tratta, nella maggior parte dei casi, di macchine robotiche, non necessariamente umanoidi, utilizzate per una varietà di compiti, come ad esempio attività di pattugliamento, sorveglianza, disinnesco di bombe, individuazione di atteggiamenti sospetti, riconoscimento facciale, etc. [13].

Applicazioni di questo tipo, se da un lato hanno il gran merito di preservare da una serie di pericoli gli agenti (umani), e se in talune circostanze assicurano un ottimo livello di efficienza nelle prestazioni erogate, sollevano, dall'altro lato, una serie di problematiche:

– la questione della *privacy*, in considerazione della gran mole di dati che queste applicazioni (fornite, ad esempio, di sensori e telecamere avanzate) possono acquisire in relazione alla vita, anche privata, dei cittadini: dati che, peraltro, potrebbero essere manipolati abusivamente, sottratti, deformati, con grave pregiudizio per le persone cui essi si riferiscono;

– alcune di queste applicazioni sono equipaggiate con armi, non letali (ad esempio, il *taser* o lo *spray* al peperoncino) o letali (equiparabili alle classiche

[12] In argomento, v. N. SHARKEY, *2084: Big robot is watching you. Report on the future of robots for policing, surveillance and security*, 2008, reperibile al seguente indirizzo *web https://it.scribd.com/document/139971746/Noel-Sharkey-2084-Big-robot-is-watching-you-Future-Robot-Policing-Report-Final*; una versione più breve di tale saggio, intitolata *The robot arm of the law grows longer*, e originariamente pubblicata sulla rivista *Computer*, 2009, p. 113, può essere letta anche a questo indirizzo *web https://ieeexplore.ieee.org/document/5197441*; v. pure L. ROYAKKERS, R. VAN EST, *A Literature Review on New Robotics: Automation from Love to War*, in *International Journal of Social Robotics*, 7, 5, 2015, p. 549 ss.; E.E. JOH, *Policing Police Robots*, in *UCLA Law Review Discourse*, 2016, p. 516; L. PASCULLI, *Genetics, Robotics and Crime Prevention*, in D. PROVOLO, S. RIONDATO, F. YENISEY (a cura di), *Genetics, Robotics, Law, Punishment*, Padova University Press, Padova, 2014, p. 197 ss.; per una disamina più generale sull'uso della IA nelle attività di *law enoforcement*, v. P. STONE *et al.*, *Artificial Intelligence and Life in 2030*, cit., p. 36 ss.

[13] Un sofisticato programma di riconoscimento facciale – SARI, Sistema Automatico di Riconoscimento Immagini – è in dotazione anche alla Polizia scientifica italiana, stando a quanto si apprende dalle notizie giornalistiche, v. redazione ANSA, *Ladri individuati grazie al nuovo sistema di riconoscimento facciale*, 7 settembre 2018.

armi da fuoco), il che crea indubbie preoccupazioni in ordine al tasso di fallibilità di queste applicazioni e quindi in ordine all'individuazione del responsabile (uomo o macchina?) di eventuali uccisioni o lesioni commesse per errore, nonché in ordine alla presumibile assenza, in capo a questi dispositivi robotizzati armati, di doti tipicamente umane – la pietà, l'intuito, la capacità di improvvisazione, il cd. senso comune [14] – la cui presenza, in operatori della polizia, è sempre auspicabile [15];

– vi è, poi, il problema dell'ampiezza che il controllo umano deve assumere su tali applicazioni: il controllo dell'uomo si deve limitare alla scelta degli obiettivi, al monitoraggio, o deve essere un controllo più intenso, esercitato anche a costo di compromettere le prestazioni stesse del RoboCop?

Suona inquietante, se riguardato in questa prospettiva, il fatto che il *sequel* del film RoboCop, uscito nel 2014, avesse come sottotitolo: «chi avrà il controllo: l'uomo o il robot?» [16].

1.3.2. *Sistemi di intelligenza artificiale e polizia predittiva*

Oltre ai RoboCop, in relazione alle attività di *law enforcement* dobbiamo anche citare le possibili applicazioni dei sistemi di IA per finalità di polizia predittiva, laddove per "polizia predittiva" possiamo intendere l'insieme delle attività rivolte allo studio e all'applicazione di metodi statistici con l'obiettivo di "predire" chi potrà commettere un reato, o dove e quando potrà essere commesso un reato, al fine di prevenire la commissione dei reati stessi.

La predizione si basa fondamentalmente su una rielaborazione attuariale di diversi tipi di dati, tra cui quelli relativi a notizie di reati precedentemente commessi, agli spostamenti e alle attività di soggetti sospettati, ai luoghi, teatro di ricorrenti azioni criminali, e alle caratteristiche di questi luoghi, al periodo del-

[14] Come giustamente sottolinea M.B. MAGRO, *Biorobotica, robotica e diritto penale*, in D. PROVOLO, S. RIONDATO, F. YENISEY (a cura di), *Genetics, Robotics, Law, Punishment*, cit., p. 512, «ai robot dotati di intelligenza artificiale, dotati di conoscenze altamente specialistiche, manca, al di sotto di queste conoscenze, il livello di conoscenze comuni, il c.d. "senso comune", ciò che tutti gli umani posseggono senza aver fatto studi particolari. Il "senso comune" è quello che consente di collegare conoscenze specialistiche di campi diversi e di affrontare i problemi e di risolverli senza la rigidità tipica dell'approccio simbolico dell'intelligenza. Spesso una reazione intelligente ad una certa situazione è quella che, sì, tiene in considerazione il contesto, ma che non è capace di selezionare quale aspetto del contesto sia rilevante».

[15] Sulle cd. *autonomous weapons*, v. in particolare N. SHARKEY, *La robotica*, in J. AL-KHALILI (a cura di), *Il futuro che verrà*, Bollati Boringhieri, Torino, 2018, p. 195 ss.; R. CALO, *Artificial Intelligence Policy*, cit., p. 196, con ulteriori riferimenti.

[16] *https://www.youtube.com/watch?v=0BIWKVH8_GE*.

l'anno o alle condizioni atmosferiche maggiormente connesse alla commissione di determinati reati; tra i dati utilizzati a questi fini talora compaiono anche informazioni relative all'origine etnica, al livello di scolarizzazione, alle condizioni economiche, alle caratteristiche somatiche (… una rivincita di Lombroso?), riconducibili a soggetti appartenenti a determinate categorie criminologiche (ad es., potenziali terroristi), etc. [17].

In tempi recenti, l'impiego di *software* basati sull'IA ha consentito di fare un salto di qualità nelle attività di polizia predittiva, dal momento che è ora possibile l'acquisizione e la rielaborazione di una mole enorme di dati, che fa emergere connessioni prima difficilmente individuabili dall'operatore umano [18].

I *software* di polizia predittiva possono dividersi fondamentalmente in due categorie:

– quelli che, ispirandosi alle acquisizioni della criminologia ambientale, individuano le cd. "zone calde" (*hotspots*), vale a dire i luoghi che costituiscono il possibile scenario dell'eventuale futura commissione di determinati reati (ad es., il sistema informatico *X-LAW*, originariamente predisposto dalla Questura di Napoli, che parrebbe aver già ottenuto ottimi risultati sul territorio italiano nel campo della prevenzione di talune tipologie di reati [19]);

– quelli che, ispirandosi invece all'idea del *crime linking*, seguono le serialità criminali di determinati soggetti (individuati o ancora da individuare), per prevedere dove, come e quando costoro commetteranno il prossimo reato, identifi-

[17] Per un completo inquadramento della materia della *predictive policing*, v. W.L. PERRY, B. MCINNIS, C.C. PRICE, S.C. SMITH, J.S. HOLLYWOOD, *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*, Rand, Santa Monica, 2013.

[18] C. CATH, S. WACHTER, B. MITTELSTADT, M. TADDEO, L. FLORIDI, *Artificial Intelligence and the "Good Society": the US, EU, and UK approach*, in *Science and Engineering Ethics*, 2018, p. 505 ss.; L. BENNET MOSES, J. CHAN, *Algorithmic Prediction in Policing: Assumptions, Evaluation, and Accountability*, in *Policing and Society*, 2016, p. 1 ss.; G. MASTROBUONI, *Crime is Terribly Revealing: Information Technology and Police Productivity*, in *Review of Economic Studies*, 87, 6, novembre 2020, p. 2727 ss., consultabile *online* al seguente link: *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2989914*; per un sintetico quadro, in lingua italiana, dei sistemi di IA finalizzati ad attività di polizia predittiva, v. R. PELLICCIA, *Polizia predittiva: il futuro della prevenzione criminale?*, in *cyberlaws.it*, 9 maggio 2019; B. PEREGO, *Predictive policing: trasparenza degli algoritmi, impatto sulla privacy e risvolti discriminatori*, in *BioLaw Journal*, 2, 2020, p. 447 ss.; pur riferendosi a sistemi non solo di polizia predittiva v. C. MORELLI, *Furti e rapine: a sventarli ci pensa l'intelligenza artificiale!*, in *Altalex.com*, 6 maggio 2019.

[19] Notizie riferite da M. IASELLI, *X-LAW: la polizia predittiva è realtà*, in *Altalex.com*, 28 novembre 2018. Un sistema recentissimo di questo tipo, sempre di pregiata fattura italiana, è Pelta Suite, in sperimentazione nel Comune di Caorle. L. BARIELLA, *Polizia predittiva: al via la sperimentazione a Caorle*, in *Altalex.com*, 24 maggio 2021.

cando "la mano criminale", il *modus operandi* emergente dalla serie criminale (un esempio in tal senso è l'ormai noto *software* Delia della società KeyCrime).

Questi sistemi di polizia predittiva possono indubbiamente apportare grandi benefici, ma il loro utilizzo suscita più d'una perplessità [20]:

– essi possono fornire adeguate previsioni solo in relazione a limitate, determinate categorie di reati (ad esempio, reati attinenti alla criminalità da strada, come rapine e spaccio di stupefacenti), non necessariamente quelli più pericolosi per la democrazia e per la libertà democratica;

– il loro uso potrebbe implicare gravi attriti con la tutela della *privacy* (in considerazione della gran mole di dati personali raccolti), e con il divieto di discriminazione (nella misura in cui, ad esempio, identifichino fattori di pericolosità connessi a determinate caratteristiche etniche, o religiose o sociali) [21];

– si tratta, poi, di sistemi che in una certa misura si auto-alimentano coi dati prodotti dal loro stesso utilizzo, col rischio di innescare circoli viziosi, dando origine al fenomeno della "profezia che si auto-avvera": se, ad esempio, un *software* predittivo individua una determinata "zona calda", i controlli e i pattugliamenti della polizia in quella zona si intensificheranno, con inevitabile conseguente crescita del tasso dei reati rilevati dalla polizia in quella zona, che diventerà, quindi, ancora più "calda", mentre altre zone, originariamente non ricondotte nelle "zone calde", e quindi non presidiate dalla polizia, rischiano di rimanere, o di diventare, per anni zone "fredde", ove la commissione di reati non viene adeguatamente monitorata;

– inoltre, questi sistemi sollecitano una prevenzione dei reati attraverso l'intervento attivo della polizia, attraverso, quindi, una sorta di "militarizzazione" nella sorveglianza di determinate zone o di determinati soggetti, senza invece minima-

[20] Le considerazioni contenute nel prosieguo del testo rielaborano spunti e riflessioni formulati da L. PASCULLI, *Genetics, Robotics and Crime Prevention*, cit., p. 192, e da R. PELLICCIA, *Polizia predittiva*, cit., che rinvia, tra l'altro, alle ricerche compiute in materia, e alle relative perplessità espresse, dall'Human Rights Data Analysis Group (HRDAG), raccolte nel sito *https://hrdag.org/usa/*, alla voce *The Problem with Predictive Policing*. Sulle stesse problematiche, più di recente v. anche B. PEREGO, *Predictive policing*, cit., p. 452 ss.

[21] Su questi aspetti, v. A. BONFANTI, *Big data e polizia predittiva: riflessioni in tema di protezione del diritto alla privacy e dei dati personali*, in *MediaLaws*, 24 ottobre 2018; E. THOMAS, *Why Oakland Police Turned Down Predictive Policing*, in *vice.com*, 28 dicembre 2016; J. KREMER, *The end of freedom in public places? Privacy problems arising from surveillance of the European public space*, Helsinki, 2017, in particolare il capitolo 3.4.2, "*Prediction*", p. 269 ss.; in particolare, sul ruolo dei dati nella discriminazione algoritmica prodotta dai sistemi di polizia predittiva, v. R. RICHARDSON, J. SCHULTZ, K. CRAWFORD, *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice*, in *New York University Law Review*, 94, 2019, p. 192 ss.

mente mirare alla riduzione del crimine attraverso un'azione rivolta, a monte, ai fattori criminogeni (fattori sociali, ambientali, individuali, economici, etc.);

– infine, non si deve trascurare il fatto che la maggior parte di questi *software* sono coperti da brevetti depositati da aziende private, le quali, a buon diritto, sono gelose dei relativi segreti industriali e commerciali, sicché non si può disporre di una piena comprensione dei meccanismi del loro funzionamento, con evidente pregiudizio delle esigenze di trasparenza e di verificazione indipendente della qualità e affidabilità dei risultati da essi prodotti. D'altra parte, se anche i meccanismi di funzionamento fossero resi pubblici, la logica di molti di essi potrebbe risultare comunque intrinsecamente non intelligibile nemmeno per un esperto di IA, dal momento che questi meccanismi si "autoalimentano" in modo imperscrutabile tramite il *machine learning* (cd. *black box*).

1.4. *Secondo ambito – IA e decisione giudiziaria: la macchina-giudice?*

Algoritmi basati sull'IA vengono, già da qualche tempo, utilizzati anche a fini decisionali nei più svariati ambiti [22]: si tratta dei cd. *automated decision systems*, in via di crescente diffusione [23], sia in ambito privato, sia in ambito pubblico [24].

Tra le decisioni che siffatti algoritmi sono in grado di assumere vi sono, ovviamente, anche decisioni finalizzate a comporre, o prevenire, liti e risolvere controversie.

Anzi, in quest'ambito, le nuove tecnologie – grazie alla possibilità di attingere a quantità enormi di dati da fonti quali banche-dati giurisprudenziali, legislative, raccolte di precedenti, e simili – hanno già messo a punto dispositivi molto sofisticati, che utilizzano teoria dei giochi, analisi dei risultati positivi e strategie di negoziazione per risolvere le questioni [25].

Anche queste applicazioni presentano indubbiamente taluni vantaggi, tra i quali:

[22] J. KLEINBERG, H. LAKKARAJU, J. LESKOVEC, J. LUDWIG, S. MULLIANATHAN, *Human Decisions and Machine Predictions*, in *Quarterly Journal of Economics*, 2017, p. 237.

[23] D. REISMAN, J. SCHULTZ, K. CRAWFORD, M. WHITTAKER, *Algorithmic Impact Assessments: a Practical Framework for Public Agency Accountability*, 2018, reperibile al seguente link: *https://ainowinstitute.org/aiareport2018.pdf*.

[24] Sull'impiego, all'interno della pubblica amministrazione, di sistemi decisionali basati sull'IA in Italia e in Argentina, v. ad esempio D.U. GALETTA, J.G. CORVALÀN, *Intelligenza Artificiale per una Pubblica Amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto*, in *federalismi.it*, 6 febbraio 2019, p. 1 ss.

[25] Per un primo inquadramento del possibile impatto delle tecnologie di IA sul processo penale, mi permetto di rinviare al seguente link: *https://www.youtube.com/watch?v=Tl8an9paY8M*.

– impiegano una metodologia che i soggetti coinvolti percepiscono come oggettiva e priva di pregiudizi [26];

– comportano una riduzione dei tempi e significativi risparmi di spesa sia per i soggetti coinvolti, sia per i soggetti responsabili della decisione [27].

Esse, tuttavia, suscitano inevitabilmente talune preoccupazioni, soprattutto se si pensa ad un loro possibile impiego anche in sede penale [28]:

– potrebbero essere fonte di discriminazioni e automatismi;

– mettono in crisi la tradizionale idea di "giudice naturale precostituito per legge" (art. 25, comma 1, Cost.), idea che finora aveva anche una proiezione geografica: il giudice-macchina sarà un giudice unico per tutto il territorio nazionale?

– anche il principio espresso dall'art. 101, comma 1, Cost. ne risulta scosso: può una macchina agire "in nome del popolo"?

– e che dire della "soggezione soltanto alla legge", richiesta dall'art. 101, comma 2, Cost.? il giudice macchina sarà, probabilmente, molto più vincolato al precedente, di quanto lo sia oggi il giudice uomo (perlomeno nei sistemi di *civil law*), con il rischio, peraltro, di ostacolare interpretazioni evolutive;

– infine, sembra pressoché impossibile aspettarsi da un algoritmo la capacità di intendere e applicare la regola di giudizio, di cui all'art. 533, comma 1, c.p.p., basata sull'"oltre ogni ragionevole dubbio", dal momento che possiamo immaginare *software* capaci di dare risposte secondo una logica binaria (sì/no; bianco/nero; vero/falso), o anche secondo una logica probabilistica (sì al 70%; bianco all'80%; vero al 90%), ma difficilmente *software* capaci di esprimere valutazioni, nella cui assunzione giochino un ruolo irrinunciabile – per quanto non ponderabile in termini precisi – fattori irriducibilmente umani [29].

[26] J. KAPLAN, *Intelligenza Artificiale*, cit., p. 137 ss.

[27] E. LATIFAH, A.H. BAJREKTAREVIC, M.N. IMANULLAH, *Digital Justice in Online Dispute Resolution: The Shifting from Traditional to the New Generation of Dispute Resolution*, in *Brawijaya Law Journal – Journal of Legal Studies*, 6, 1, aprile 2019.

[28] Tra gli altri, v. G. CANZIO, *Il dubbio e la legge*, in *Dir. pen. cont.*, 2018, p. 1 ss.; M. GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, ivi, 29 maggio 2019, p. 1 ss.; A. NATALE, *Introduzione. Una giustizia (im)prevedibile?*, in *Quest. giust.*, 4, 2018, p. 1 ss.; nello stesso fascicolo, v. pure i contributi di C. COSTANZI, *La matematica del processo: oltre le colonne d'Ercole della giustizia penale*, e di C. CASTELLI, D. PIANA, *Giustizia predittiva. La qualità della giustizia in due tempi*; vedasi, infine, il fascicolo 7/2019 di *Giur. it.* che ospita un'ampia sezione monografica, a cura di U. Ruffolo ed E. Gabrielli, dedicata al tema *Intelligenza artificiale e diritto*.

[29] Sul punto, v. pure S. GABORIAU, *Libertà e umanità del giudice: due valori fondamentali della giustizia. La giustizia digitale può garantire nel tempo la fedeltà a questi valori?*, in *Quest. giust.*, 4, 2018, p. 11.

1.5. *Terzo ambito – IA e valutazione della pericolosità criminale: gli algoritmi predittivi*

Quali probabilità sussistono che un individuo, avente determinate caratteristiche, possa in futuro commettere un (nuovo) reato?

Si tratta di un quesito la cui risposta è necessaria, tra l'altro, quando si tratta di applicare una misura di sicurezza, una misura cautelare o una misura di prevenzione, o anche per concedere la sospensione condizionale di una pena o l'affidamento in prova al servizio sociale [30].

Ebbene, a tale fondamentale quesito oggi i nostri giudici forniscono risposte per lo più intuitive, affidate esclusivamente alla loro esperienza personale e al loro buon senso, oppure, quando consentito dalla legge, basate su valutazioni cliniche di periti [31], mentre in futuro (e già nel presente di altri ordinamenti giuridici) siffatte valutazioni prognostiche della pericolosità criminale potrebbero essere affidate a specifici algoritmi (*risk assessment tools*, o algoritmi predittivi), capaci di effettuare valutazioni attuariali, rielaborando quantità enormi di dati al fine di far emergere relazioni, coincidenze, correlazioni, che consentano di profilare una persona e prevederne i successivi comportamenti, anche di rilevanza penale [32].

Negli Stati Uniti, in effetti, già da una decina d'anni sono in fase di diffusione algoritmi predittivi della pericolosità criminale.

Essi sono, ad esempio, usati nella fase del *parole* (per decidere se un individuo, nelle more della celebrazione del processo, possa essere rilasciato dietro il pagamento di una eventuale cauzione), o per misurare il rischio di recidiva del condannato, ai fini della sua ammissibilità al *probation* o ad altra misura alternativa alla detenzione, o infine in sede di *sentencing.*

[30] Sui plurimi ambiti, all'interno dei quali risulta necessario formulare una prognosi di futura commissione di un (nuovo) reato, sia consentito rinviare a F. BASILE, *Esiste una nozione ontologicamente unitaria di pericolosità sociale? Spunti di riflessione, con particolare riguardo alle misure di sicurezza e alle misure di prevenzione*, in *Riv. it. dir. proc. pen.*, 2018, p. 644 ss.

[31] Sul cui grado di affidabilità, tuttavia, la dottrina è fortemente scettica: v., per tutti, J. MONAHAN, *Predicting violent behavior: An assessment of clinical techniques*, Sage Pubns, London, 1981.

[32] L. CASTELLETTI, G. RIVELLINI, E. STRATICÒ, *Efficacia predittiva degli strumenti di Violence Risk Assessment e possibili ambiti applicativi nella psichiatria forense e generale italiana*, in *Journal of Psychopathology*, 2014, p. 153 ss.; G. ROCCA, C. CANDELLI, I. ROSSETTO, F. CARABELLESE, *La valutazione psichiatrico forense della pericolosità sociale del sofferente psichico autore di reato: nuove prospettive tra indagine clinica e sistemi attuariali*, in *Riv. it. med. leg. dir. san.*, 4, 2012, p. 1442 ss.

I sostenitori dell'impiego degli algoritmi predittivi ritengono che questi *software*, grazie all'elaborazione di *big data* e all'apprendimento automatico, rendano le valutazioni di pericolosità criminale più accurate e maggiormente esenti dal rischio di risentire di pregiudizi e condizionamenti culturali.

Tuttavia, ancora una volta non possiamo rilevare anche alcune perplessità, espresse ad esempio in relazione al caso Loomis – in cui, in sede di *sentencing*, aveva trovato applicazione il *software* COMPAS – *Correctional Offender Management Profiling for Alternative Sanctions* – dalla Corte Suprema del Wisconsin:

– trattasi di un *software* coperto da segreto industriale, che impedisce la divulgazione di informazioni relative al suo metodo di funzionamento;

– esso effettua valutazioni su base collettiva, di gruppo, e non individuale;

– esso comporta il rischio di una sovrastima del rischio di commissione di reati a carico di talune minoranze etniche [33].

Ma soprattutto, come ormai gli stessi esperti di IA avvertono, occorre considerare (e ciò vale anche per gli altri due ambiti sopra esaminati) che l'algoritmo – qualsivoglia algoritmo – non è "neutro" [34]: nel concepire l'architettura di un algoritmo, il programmatore fa delle scelte che, necessariamente, influenzano il "risultato" dell'operazione computazionale.

1.6. *Quarto ambito – IA e responsabilità penale: così intelligente da essere "responsabile"?*

Droni che uccidono per le strade urbane [35] o su fronti lontani, impegnati nella lotta al terrorismo, auto senza conducente coinvolte nella causazione di inciden-

[33] Su questi aspetti, v. in particolare K. FREEMAN, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, in *North Carolina Journal of Law & Technology*, 18, 2016, p. 76.

[34] V. di recente le riflessioni del filosofo e psicoanalista M. BENASAYAG, *La tirannia dell'algoritmo*, Vita e Pensiero, Milano, 2020.

[35] N. SHARKEY, *La robotica*, cit., p. 197 riferisce, ad esempio, di un sospetto cecchino ucciso a Dallas tramite l'intervento di un drone (luglio 2016), commentando con le seguenti parole tale episodio: «in quel caso esisteva una chiara giustificazione e gli esperti di diritto hanno asserito che l'azione era stata legittima, resta il fatto che in quella circostanza si è probabilmente varcato un confine. È giusto proteggere la polizia, e la polizia dovrebbe, fintanto che è possibile, utilizzare mezzi non violenti. Quando questi si dimostrino inefficaci, è certamente necessario elevare il livello della forza impiegata, ma in modo graduale e proporzionale al reato che viene commesso: e sono valutazioni decisamente impegnative per un robot che agisce senza il controllo umano».

ti anche a danno di persone (come nel tragico investimento di una ciclista avvenuto nel marzo 2018 in Arizona [36]), *software* che eseguono, in collaborazione o addirittura in sostituzione dell'uomo, compiti sempre più sofisticati, come pilotare un grosso aereo, ma che qualche volta possono interferire negativamente con la condotta umana (come i recenti disastri aerei del Boeing 737 MAX hanno purtroppo dimostrato [37]): chi risponde dei fatti di reato in tal modo eventualmente commessi [38]? il programmatore del *software*? il suo produttore? il suo utilizzatore? o direttamente il sistema di intelligenza artificiale?

1.6.1. *IA strumento del reato*

Lo scenario relativamente più semplice è ovviamente quello in cui il sistema di intelligenza artificiale costituisce lo strumento – in mano a un uomo – attraverso il quale il reato viene commesso [39]. Le enormi potenzialità dell'intelligenza artificiale, infatti, potrebbero – e già lo sono state – essere asservite anche a scopi criminali e quindi essere utilizzate per la commissione di reati attraverso modalità fino a qualche anno fa assolutamente inimmaginabili: solo per fare due esempi, pensiamo a droni e sottomarini senza equipaggio, controllati a distanza, utilizzati per il trasporto di stupefacenti e armi illegali; oppure ai *social BOT*, che possono essere utilizzati come strumenti per realizzare molestie, diffamazioni, abusi della credulità popolare, attraverso *tweet*, *retweet* e altre diavolerie simili.

Dobbiamo, insomma, prepararci a un'era in cui la commissione di reati con lo strumento dell'intelligenza artificiale potrebbe diventare assai frequente e incisiva, anche in considerazione dell'accresciuta vulnerabilità di alcuni aspetti della vita umana connessi ad impieghi dell'intelligenza artificiale, a partire dall'impressionante numero di dati sul comportamento e sullo stile di vita di ciascuno di noi – facilitato da una condizione umana perennemente *Onlife* – che possono

[36] L. BUTTI, *Le auto guideranno da sole, ma con quali responsabilità?*, in *Il Bo Live*, 9 novembre 2018; F. SUMAN, *Dilemmi morali per le auto a guida autonoma*, ivi, 7 novembre 2018. Sul funzionamento delle *self-driving cars* e dei loro problemi tecnici nonché giuridici, v. H. SURDEN, M.A.WILLIAMS, *Technological Opacity, Predictability, and Self-Driving Cars*, in *Cardozo Law Review*, 38, 2016, p. 121 ss.

[37] G.F. ITALIANO, *Intelligenza artificiale*, cit.

[38] Sui cambiamenti che il largo utilizzo dei sistemi di IA potrebbe produrre sul sistema della responsabilità giuridica v. altresì le riflessioni di M.B. MAGRO, *Decisione umana e decisione robotica. Un'ipotesi di responsabilità da procreazione robotica*, in *Legisl. pen.*, 10 maggio 2020.

[39] S. RIONDATO, *Robotica e diritto penale (robot, ibridi, chimere, "animali tecnologici")*, in D. PROVOLO, S. RIONDATO, F. YENISEY (a cura di), *Genetics, Robotics*, cit., p. 600 ss.

essere raccolti tramite i vari canali informatici, fino all'eventuale instaurazione di rapporti di vera dipendenza, talora anche affettiva, da macchine e sistemi di servizio che si muovono per noi, lavorano per noi, custodiscono i nostri anziani e i nostri figli.

L'uomo rischia, insomma, di ritrovarsi in balìa della macchina, sguarnito dei presidi tradizionali di protezione, essendo tali presidi concepiti e strutturati per proteggerlo da "attacchi umani".

Sorge allora un primo interrogativo: abbiamo bisogno di nuove fattispecie di reato? O abbiamo bisogno di rimodellare quelle già esistenti, al fine di renderle applicabili alla realizzazione di condotte criminose attraverso lo strumento dell'intelligenza artificiale, offrendo così tutela ai beni giuridici anche da questa nuova fonte di attacchi?

1.6.2. *IA autore del reato?*

Gli esempi sopra formulati, che finora abbiamo presentato come ipotesi in cui l'intelligenza artificiale è lo strumento in mano all'uomo, potrebbero, però, presentarsi anche in uno scenario in cui la mano dell'uomo scompare, o diventa pressoché impercettibile.

Nel caso, infatti, in cui nella realizzazione del reato sia coinvolto un sistema di intelligenza artificiale di ultima generazione, che risulti fornito di capacità di apprendimento e di autonomia decisionale, potremmo chiederci se non risulti già varcata la frontiera del futuro, tanto da potersi individuare direttamente nel sistema di intelligenza artificiale l'"autore" del reato.

Quando le scelte, le valutazioni, i bilanciamenti sottesi alla commissione di un fatto di reato non sono più opera esclusiva dell'uomo, ma sono quanto meno "condivisi con", se non interamente delegati alla macchina, ecco che il percorso di attribuzione delle responsabilità indubbiamente si complica.

Vengono in mente scenari in parte già noti.

Come si individua il responsabile di un'attività svolta in *équipe*? Come si individua il colpevole in quelle ipotesi in cui il procedimento decisionale ed esecutivo è parcellizzato, frazionato e distribuito in capo a una pluralità di soggetti?

La novità sta però ora nel fatto che tra i membri delle *équipe*, tra i plurimi soggetti coinvolti, non vi sono più solo esseri umani, ma anche sistemi di intelligenza artificiale, col conseguente innesco di un processo di "alienazione della responsabilità" dall'agente umano [40], giacché l'agente umano si colloca lontano – nel tempo, nello spazio e nel processo decisionale – rispetto all'offesa al bene giuridico.

C'è allora il rischio di creare zone franche, sacche di illiceità all'interno delle

[40] C. BAGNOLI, *Teoria della responsabilità*, Il Mulino, Bologna, 2019, p. 77.

quali non è possibile imputare alcuna responsabilità alla persona fisica. E, se del reato non risponde l'uomo, chi ne dovrà rispondere?

Ecco, quindi, che occorre porci un nuovo interrogativo: *machina delinquere potest?* [41]

A dire il vero, la questione della possibile attribuzione di responsabilità ad entità diverse dall'uomo non è una novità assoluta. Platone, nelle Leggi, attribuiva la responsabilità anche ad animali e cose [42]; ancora, alle soglie dell'illuminismo, venivano celebrati processi penali a carico di animali "delinquenti" [43] e, dal 2001, anche in Italia è stata configurata una responsabilità da reato in capo agli enti, a carico quindi di persone che sono tali solo per effetto di una *fictio* giuridica.

L'ultima frontiera è segnata dai sistemi di intelligenza artificiale.

Possono essi essere considerati persone? O, quanto meno, possono essere assimilati alle persone, al fine di un'attribuzione di responsabilità non solo civile, ma anche penale [44]?

[41] La suggestiva formula *machina delinquere non potest* (che noi qui riprendiamo sopprimendo il "non" ed aggiungendo il punto di domanda) – formula la quale a sua volta ricalca l'antico brocardo *societas delinquere non potest*, a lungo invocato per precludere una responsabilità da reato a carico degli enti – è stata coniata da A. CAPPELLINI, *Machina delinquere non potest. Brevi appunti su intelligenza artificiale e responsabilità penale*, in *Criminalia*, 2018, p. 499 ss.

[42] Come ci ricorda, da ultimo, C. BAGNOLI, *Teoria della responsabilità*, cit., p. 72.

[43] Riferimenti in A. CAPPELLINI, *Machina delinquere non potest*, cit., p. 20; C. BAGNOLI, *Teoria della responsabilità*, cit., p. 73.

[44] Il dibattito in materia è stato inizialmente avviato dai filosofi del diritto e dai filosofi dell'informatica (v., tra gli altri, H. JONAS, *The Imperative of Responsibility. In search of an Ethics for the Technological Age*, University of Chicago Press, Chicago, 1984; L.B. SOLUM, *Legal Personhood for Artificial Intelligences*, in *North Carolina Law Review*, 70, 1992, p. 1231, ora in *Illinois Public Law and Legal Theory Research Papers*, 9-13, 20 marzo 2008; L. FLORIDI, J.W. SANDERS, *On the Morality of Artificial Agents*, in *Minds and Machines*, 14/3, 2004, p. 349 ss.; B.C. STAHL, *Information, Ethics, and Computers: The Problem of Autonomous Moral Agents*, ivi, 14, 2004, p. 67 ss.; ID., *Responsible Computers? A Case for Ascribing Quasi-Responsibility to Computers Independent of Personhood or Agency*, in *Ethics and Information Technology*, 8, 2006, p. 205 ss.; G. SARTOR, *Gli agenti software: nuovi soggetti del ciberdiritto*, in *Contratto e impresa*, 2, 2002, p. 57 ss.), e si è di recente acceso anche tra gli studiosi della responsabilità civile (si veda, ad esempio, A. SANTOSUOSSO, C. BOSCARATO, F. COROLEO, *Robot e diritto: una prima ricognizione*, in *Nuova giur. civ. comm.*, II, 2012, p. 497 ss.; A. SANTOSUOSSO, *If the agent is not necessarily a human being. Some legal thoughts*, in D. PROVOLO, S. RIONDATO, F. YENISEY (a cura di), *Genetics, Robotics, Law, Punishment*, cit., p. 545 ss., nonché il volume U. RUFFOLO (a cura di), *Intelligenza artificiale e responsabilità*, Giuffrè, Milano, 2018) e tra i costituzionalisti (si veda, ad esempio, il volume a cura di F. PIZZETTI, *Intelligenza artificiale, protezione dei dati personali e regolazione*, cit.).

La risposta positiva comporterebbe di pagare un prezzo molto alto: la disponibilità ad ammettere una colpevolezza "extra-umana".

Possiamo davvero parlare di un coinvolgimento soggettivo dell'autore-macchina al fatto commesso? Possiamo concepire una rimproverabilità, per l'appunto personale, della macchina? Possiamo parlare di capacità di intendere e di volere, in relazione a una rete neurale? Possiamo configurare una "colpa" o addirittura un "dolo" dell'algoritmo [45]?

C'è chi dice di sì [46], facendo leva sui recenti progressi fatti nella robotica, nella percezione e nel *machine learning*, supportati dai miglioramenti sempre più veloci della tecnologia informatica, al punto che oggi la frase di buonsenso comunemente accettata secondo la quale "i computer fanno solo quello che sono programmati a fare" non sarebbe più vera [47].

Accanto, peraltro, al quesito *machina delinquere potest*?, occorrerà subito dopo porsi anche il connesso quesito: (*quomodo*) *machina puniri potest*?, con quali pene? e perseguendo quale tra le possibili funzioni della pena?

1.6.2.1. *Irriducibilmente umano?*

Eppure, di fronte a questo possibile scenario, qualcosa ci lascia inquieti. La responsabilità penale è personale, cioè "della persona": davvero potremo assimilare, ai fini dell'allocazione della responsabilità penale, la macchina all'uomo? Oppure c'è qualcosa che la persona umana ha e che la macchina non potrà mai avere? [48]

[45] Su quest'ultimo interrogativo, v. D. FALCINELLI, *Il dolo in cerca di una direzione penale. Il contributo della scienza robotica ad una teoria delle decisioni umane*, in *Arch. pen.*, 1, 2018, p. 9.

[46] Tra gli scienziati di IA, fornisce una convinta risposta affermativa alle questioni formulate nel testo, J. KAPLAN, *Intelligenza artificiale*, cit., p. 153: «un sistema di IA può commettere reati? La risposta è sì»; ID., *Le persone non servono. Lavoro e ricchezza nell'epoca dell'intelligenza artificiale*, Luiss University Press, Roma, 2016, p. 80. Tra gli studiosi di diritto penale, la posizione più avanzata è quella sostenuta da Gabriel Hallevy, i cui lavori sono oggetto di una meditata presentazione critica da parte di A. CAPPELLINI, *Machina delinquere non potest*, cit., p. 10 ss., e di M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale, responsabilità e accountability. Verso nuovi paradigmi?*, in F. PIZZETTI (a cura di), *Intelligenza artificiale*, cit., p. 363 ss., ai quali, pertanto, è in questa sede possibile rinviare.

[47] J. KAPLAN, *Intelligenza artificiale*, cit., p. 19.

[48] Vedi le stimolanti riflessioni in proposito di Massimo Cacciari, in M. CACCIARI, S. ARCIERI, F. BASILE, R. BIANCHETTI, P.E. CICERONE, *Alla radice dell'imputabilità e della colpevolezza penali. Conversazione con Massimo Cacciari – pt. 2*, in *Dir. pen. uomo*, 13 gennaio 2021. Inoltre, vedi le acute e ancora attuali riflessioni di E. AGAZZI, *Alcune osservazioni sul*

Forse un'"intelligenza" superiore? Ahi, questo purtroppo no: l'intelligenza dei computer sta ormai superando quella degli esseri umani [49], almeno a livello prestazionale.

Forse la "coscienza del dis-valore sociale" della propria condotta? o i "sentimenti", che le macchine non hanno e che probabilmente mai avranno? Probabilmente no, dal momento che coscienza del dis-valore sociale e sentimenti non sono elementi necessari per fondare una responsabilità penale.

Allora il "libero arbitrio"? Be', le neuroscienze hanno ampiamente messo in discussione il libero arbitrio dell'uomo [50].

Ma se escludiamo l'intelligenza, la coscienza, i sentimenti, il libero arbitrio, cosa rimane ancora di irriducibilmente umano?

Qual è lo *specificum* dell'uomo? Che cosa potrebbe impedire, come ultima Thule, una piena assimilabilità della macchina all'uomo, anche ai fini della responsabilità penale?

problema dell'intelligenza artificiale, in *Riv. fil. neo-scolastica*, 59, 1, 1967, p. 1 ss., il quale – da ottimo filosofo della scienza quale egli è – partendo dal presupposto che nulla è logicamente impossibile, mette in luce come l'uomo sia dotato di un misterioso e indefinibile *quid pluris*, che plasma tutte quelle attività squisitamente umane, noto come "intenzionalità", che ad oggi le macchine non hanno (ancora) replicato. Sulle medesime questioni vedi altresì J.R. SEARLE, *La mente è un programma?*, in *Le scienze*, 259, 1990, p. 16 ss.

[49] V. pure quanto affermato da S. Hawking durante la Conferenza *Zeitgeist*, Londra, maggio 2015: «nell'arco dei prossimi cento anni, l'intelligenza dei computer supererà quella degli esseri umani» [citazione riportata da Redazione (a cura di), *Do You Trust This Computer?*, in *Dir. pen. uomo*, 15 maggio 2019; v. la notizia anche su *Newsweek* (L. WALKER, *Stephen Hawking warns artificial intelligence could end humanity*, 14 maggio 2015)].

[50] In particolare v. J. KAPLAN, *Intelligenza artificiale*, cit., p. 113 ss., il quale mette in discussione la concezione tipicamente occidentale e cartesiana del libero arbitrio, attingendo alle ultime scoperte neuroscientifiche e a conoscenze matematico-informatiche di lunga data (in particolare ai cd. problemi indecidibili).

CAPITOLO II

Reati colposi e tecnologie dell'intelligenza artificiale

ALBERTO CAPPELLINI

SOMMARIO: 2.1. Introduzione. Autorìa "artificiale intelligente" e responsabilità colposa. – 2.2. Macchine intelligenti e imprevedibilità tecnologica. – 2.3. Il "*responsibility gap*" e le problematiche regolative dell'attribuzione di colpa: i riflessi sull'imputazione ai produttori umani. – 2.4. (*Segue*). I riflessi sull'imputazione agli utilizzatori umani. – 2.5. IA, colpa, caso fortuito: la politicità della soglia del rischio consentito e le influenze della precauzione. – 2.6. Riflessioni conclusive: quali prospettive per il futuro?

2.1. *Introduzione. Autorìa "artificiale intelligente" e responsabilità colposa*

Lo sviluppo impetuoso delle tecnologie dell'intelligenza artificiale, e la prospettiva di una penetrazione via via più diffusa di strumenti tecnologici "intelligenti" nelle società contemporanee, ha condotto la riflessione penalistica più recente a interrogarsi sui numerosi profili di intersezione che tali innovazioni hanno con il diritto penale. Uno dei campi forse più futuristici, ma certamente ricco di precipitati in un domani neanche poi così lontano, è quello di come la responsabilità penale più classica – quella individuale, dei soggetti umani – subisca modificazioni quando il reato non sia commesso immediatamente per mano dell'autore, bensì per mezzo di strumenti a carattere artificiale intelligente [1].

[1] Non sono pochi, ormai, i lavori generali sul tema, anche solo limitandosi alla lingua italiana. Per tutti: C. PIERGALLINI, *Intelligenza artificiale: da "mezzo" ad "autore" del reato?*, in *Riv. it. dir. proc. pen.*, 2020, p. 1743; I. SALVADORI, *Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale*, in *Riv. it. dir. proc. pen.*, 2021, p. 83; A. GIANNINI, *Intelligenza artificiale,* human oversight *e responsabilità penale: prove d'impatto a livello europeo*, in *disCrimen*, 21 novembre 2022; B. MAGRO, *Robot, cyborg e intelligenze artificiali*, in A. CADOPPI, S. CANESTRARI, A. MANNA, M. PAPA (diretto da), *Cybercrime*, Utet-Wki, Milano, 2019, p. 1179; EAD., *Biorobotica, robotica e diritto penale*, in D. PROVOLO, S. RIONDATO, F.

Come spesso accade, talvolta la realtà supera la fantasia nel proporre all'attenzione degli operatori una casistica concreta che già oggi è più variegata di quanto si immagini possibile [2].

YENISEY (eds), *Genetics, robotics, law, punishment*, Padova University Press, Padova, 2014, p. 499; S. RIONDATO, *Robotica e diritto penale (robots, ibridi, chimere e "animali tecnologici")*, ivi, p. 599; ID., *Robot: talune implicazioni di diritto penale*, in P. MORO, C. SARRA (a cura di), *Tecnodiritto. Temi e problemi di informatica e robotica giuridica*, Francoangeli, Milano, 2017, p. 85; U. PAGALLO, *Saggio sui robot e il diritto penale*, in S. VINCIGUERRA, F. DASSANO (a cura di), *Scritti in memoria di Giuliano Marini*, Esi, Napoli, 2010, p. 595; ID., *Robotica*, in U. PAGALLO, M. DURANTE (a cura di), *Manuale di informatica giuridica e diritto delle nuove tecnologie*, Utet, Torino, 2012, p. 141; R. BORSARI, *Intelligenza Artificiale e responsabilità penale: prime considerazioni*, in *MediaLaws*, 3/2019, p. 262; V. MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in *disCrimen*, 15 maggio 2020, p. 2; U. RUFFOLO, *Machina delinquere potest? Responsabilità ed "illeciti" (anche penali?) della "persona elettronica" e tutele per gli agenti software autonomi*, in ID. (a cura di), *XXVI Lezioni di Diritto dell'Intelligenza Artificiale*, Giappichelli, Torino, 2021, p. 295; P. SEVERINO, *Intelligenza artificiale e diritto penale*, in U. RUFFOLO (a cura di), *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Giuffrè Francis Lefebvre, Milano, 2020, p. 533; EAD., *Le implicazioni dell'intelligenza artificiale nel campo del diritto con particolare riferimento al diritto penale*, in P. SEVERINO (a cura di), *Intelligenza artificiale, Politica, economia, diritto, tecnologia*, Luiss University Press, Roma, 2022; V. ARAGONA, *I Robot: the criminal liability of artificial intelligences*, in *TransJus Working Papers Publications*, 4/2019, p. 83. Per un inquadramento più ampio del tema entro quello più generale dei rapporti tra IA e giustizia penale, fondamentale F. BASILE, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Dir. pen. uomo*, 29 settembre 2019, oltre al successivo ID., *Intelligenza artificiale e diritto penale: qualche aggiornamento e qualche nuova riflessione*, in F. BASILE, M. CATERINI, S. ROMANO (a cura di), *Il sistema penale ai confini delle* hard sciences*. Percorsi epistemologici tra neuroscienze e intelligenza artificiale*, Pacini, Pisa, 2021, p. 11 (nel cennato volume, esattamente sul tema, va altresì ricordato il lavoro di G.R. MINELLI, *Quando l'autore del reato è un robot: tra vecchi modelli imputativi e nuovi possibili paradigmi di responsabilità penale*, p. 57, oltre a vari interessanti spunti negli altri contributi, fra cui E. LO MONTE, *Intelligenza artificiale e diritto penale: le categorie dommatiche alla prova del futuribile*, p. 41). Più incentrato sulla sola tematica in questione un ulteriore contributo di F. BASILE, *Diritto penale e Intelligenza Artificiale*, in *Giur. it.*, 2019, suppl., p. 67. Sia consentito, infine, un riferimento fin d'ora, una volta per tutte, al nostro A. CAPPELLINI, Machina delinquere non potest*? Brevi appunti su intelligenza artificiale e responsabilità penale*, in *Criminalia*, 2018, p. 499, per una ricostruzione dell'argomento dell'IA come agente del reato, e della possibilità di "punire" quest'ultimo. La letteratura straniera sul tema, invece, è troppo ampia per essere qui sistematicamente ricordata: si rinvia, per tutti, ai richiami di cui alle note che seguono, oltre che ai riferimenti bibliografici di cui già al nostro scritto da ultimo citato.

[2] Fra la casistica "bizzarra" già verificatasi, si può ricordare, solo a titolo di esempio, l'utilizzo di sottomarini robotici per traffici di droga: U. PAGALLO, *The Laws of Robots. Crimes, Contracts and Torts*, Springer, Dordrecht, 2013, p. 65. Per un panorama generale, attuale e possibile futuro, degli *AI crimes*: T.C. KING, N. AGGARWAL, M. TADDEO, L. FLORIDI, *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions*,

Eppure, al di là delle pressoché infinite possibilità con cui i più diversi reati previsti dall'ordinamento possono combinarsi con modalità commissive che passano attraverso l'azione materiale di un soggetto artificiale intelligente, ve ne è una che merita particolare attenzione. Essa, infatti, forse più di tutte, è gravida di importanti conseguenze che già si possono avvertire e che si può probabilmente preconizzare non mancheranno di emergere per frequenza, importanza pratica e rilievo teorico in futuro.

Il riferimento è ai reati colposi contro la vita e l'incolumità individuale [3].

L'orizzonte verso cui si stanno muovendo le tecnologie in questione – come sempre accade nelle moderne società industriali-capitalistiche – è infatti quello che, al di là di iniziali prototipi unici o limitati nel numero, cerca di "confezionare" l'intelligenza artificiale in prodotti standardizzati da assemblare in serie, abbattendo così i costi di produzione, e poi immettere sul mercato. L'introduzione in massa di questi "prodotti artificiali intelligenti" nel tessuto sociale, non può che riproporre all'attenzione degli operatori – *mutatis mutandis* – quel problema fondamentale della sicurezza degli utenti e di terzi che solitamente ricade sotto l'etichetta dei profili di responsabilità per danno da prodotto [4], ma che in tale sce-

in *Science and Engineering Ethics*, 2020, p. 89; M. CALDWELL, J.T.A. ANDREWS, T. TANAY, L.D. GRIFFIN, *AI-enabled future crime*, in *Crime Science*, 2020.

[3] Fra i lavori specificamente dedicati al tema del rapporto tra autorìa artificiale e reati colposi, per tutti: S. BECK, *Intelligent agents and criminal law – Negligence, diffusion of liability and electronic personhood*, in *Robotics and Autonomous Systems*, 2016, p. 138; EAD., *Google Cars, Software Agents, Autonomous Weapons Systems – New Challanges for Criminal Law*, in E. HILGENDORF, U. SEIDEL (eds), *Robotics, Autonomics and the Law*, Nomos, Baden-Baden, 2017, p. 227; A. MORAITI, *AI Crimes and Misdemeanors: Debating the Boundaries of Criminal Liability and Imputation*, in G. VERMEULEN, N. PERŠAK, N. RECCHIA (eds), *Artificial Intelligence, Big Data and Automated Decision-Making in Criminal Justice*, Maklu, Antwerpen, 2021, p. 109 (nel medesimo volume va ricordato altresì B. PANATTONI, *AI and Criminal Law: The Myth of "Control" in a Data-Driven Society*, p. 125, dal taglio tematico invero un po' più ampio).

[4] Nella penalistica italiana, sulla responsabilità per danno da prodotto cfr., per tutti: C. PIERGALLINI *Danno da prodotto e responsabilità penale. Profili dommatici e politico-criminali*, Giuffrè, Milano, 2004; ID., *La responsabilità del produttore: una nuova frontiera del diritto penale?*, in *Dir. pen. proc.*, 2007, p. 1125 (studi dei quali il ricordato recente scritto ID., *Intelligenza artificiale: da "mezzo" ad "autore" del reato?*, cit., è espressamente presentato come una "continuazione"); D. CASTRONUOVO, *Responsabilità da prodotto e struttura del fatto colposo*, in *Riv. it. dir. proc. pen.*, 2005, p. 301; A. BERNARDI, *La responsabilità da prodotto nel sistema italiano: profili sanzionatori*, in *Riv. trim. dir. pen. econ.*, 2003, p. 1. Un nesso tra le due tematiche è sottolineato altresì da S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame? Self-driving cars and criminal liability*, in *New Criminal Law Review*, 2016, p. 426. Sul tema del rapporto tra IA e danno da prodotto cfr. anche R. BERTOLESI, *Intelligenza artificiale e responsabilità penale per danno da prodotto*, Università degli Studi di Milano, Tesi dottorale, a.a. 2018/2019.

nario – a ben vedere – pare estendersi fino a coinvolgere ambiti di responsabilità per colpa più classici e tradizionali [5].

In effetti, la diffusione ampia di tali tipologie nuove di prodotti, anche con l'adozione dei massimi standard di sicurezza possibili, inevitabilmente provocherà un certo numero di sinistri, con danno per l'incolumità umana. Anzi, per la verità simili eventi si sono già verificati: anche solo limitandosi al settore applicativo delle tecnologie dell'IA forse più ampio, quello delle auto a guida autonoma, sono moltissimi gli accadimenti nefasti, anche mortali, di cui si ha ad oggi notizia [6].

Ma, a ben vedere, è la struttura stessa del reato colposo che fa sì che la rilevanza pratica di una prospettiva di rischio vada al di là della mera frequenza statistica – magari in realtà rara – degli incidenti. Certamente il delitto d'evento non può sussistere a prescindere da un accadimento di danno. Ma d'altro canto è vero anche che le regole cautelari, le quali stabiliscono i limiti dell'attività in questione, hanno l'obiettivo di prevenire – a monte – uno spettro di prospettive lesive quanto mai ampio e variegato. Insomma, anche quando il momento "patologico" – l'incidente-evento e, correlativamente, il reato colposo di risultato – sia statisticamente infrequente, è indubbio come l'idea di prevenzione e di tutela dai rischi che informa (in particolare ma non solo) le attività industriali complesse plasmi con prepotenza tutta la disciplina del fenomeno, anche nella sua fisiologia. Così, in breve, l'atteggiamento che l'ordinamento assume nei confronti della prospettiva-limite – l'incidente mortale, il disastro – si ripercuote a ritroso venendo a influenzare in ogni parte e aspetto le varie, singole attività che caratterizzano il vivere sociale. Fino a che punto consentirle, regolamentarle attraverso l'introduzione di cautele e come distribuire tra i vari "partecipanti" le responsabilità conseguenti, sono scelte che dipendono da delicate operazioni di bilanciamento tra interessi in gioco, valutazioni costi/benefici che dipendono in modo ampiamente significativo anche dalla prospettiva dell'eventualità ultima, più recondita e negativa [7].

[5] Quali – come si dirà meglio *infra* – la colpa stradale, la colpa medica, e così via.

[6] Due gli episodi più ricordati: Williston (Florida, USA), 7 maggio 2016, una vettura Tesla modello S si infilava sotto ad un camion bianco, non riuscendo a distinguerlo dal cielo luminoso, distruggendo completamente l'abitacolo e provocando così la morte del conducente; Tempe (Arizona, USA) 18 marzo 2018, una vettura Uber completamente autonoma investiva un pedone, provocandone la morte. Ma ne sono avvenuti altri anche altri. Alcuni mortali: Mountain View (California, USA) 23 marzo 2018; Houston (Texas, USA), 17 aprile 2021. Molti altri – difficili peraltro anche da tracciare in fonti ufficiali e non – con conseguenze meno gravi.

[7] Si tratta di valutazioni che definiscono la soglia di quello che è penalisticamente noto come *rischio consentito*. Nel contesto contemporaneo della *società del rischio* vi è sostanziale convergenza di opinioni circa la natura politico-valutativa di un simile giudizio, e dell'importante ruolo della *paura* (che sovente si presenta nella veste di *precauzione*) nella scelta

In tale quadro, è innegabile la carica sostanziale, di valore, della questione. Essa infatti incrocia, da un lato, l'entità e la portata della penetrazione delle tecnologie dell'IA nel tessuto sociale che si ritenga opportuno autorizzare sul piano della scelta politica; dall'altro, le ataviche paure dell'uomo circa i rischi per la propria sicurezza fisica, nel loro approcciarsi all'ignoto tecnologico che tali innovazioni prospettano di introdurre nella vita di tutti i giorni.

Il discorso sulla colpa penale, insomma, imbattendosi nel *novum* delle tecnologie dell'intelligenza artificiale, non fa che riproporre alcune problematiche di fondo – di matrice forse più antropologica e sociopolitica che giuridica in senso stretto – che lo caratterizzano; che già l'incontro con la modernità tecnologica – quella che in sociologia si è definita "post-modernità", o "società del rischio" – ha acuito e riproposto; e che la prospettiva ulteriore e ultima di evoluzione del contesto in cui può compiersi il reato colposo – ovvero quando ciò accada per mano di un soggetto artificiale intelligente – fa emergere in modo ancor più incisivo [8].

2.2. *Macchine intelligenti e imprevedibilità tecnologica*

Da sempre il diritto penale ha conosciuto la figura del "mezzo", o "strumento", del reato. Le "macchine", *lato sensu* intese, ne hanno finora condiviso appieno lo statuto: anche il più complesso dei computer, se utilizzato per commettere un reato, lascia inalterata la responsabilità del soggetto umano alle sue spalle, al pari del più rudimentale degli utensili di cui quest'ultimo si sia avvalso a fini criminosi [9].

Nei reati colposi, in particolare, qualora l'evento lesivo sia provocato da un

collettiva. Per tutti, nella letteratura sociologica, U. BECK, *La società del rischio. Verso una seconda modernità* (1986), Carocci, Roma, 2000, p. 38; A. GIDDENS, *Le conseguenze della modernità* (1990), Il Mulino, Bologna, 1994, p. 125; N. LUHMANN, *Sociologia del rischio* (1991), Mondadori, Milano, 1996, p. 40; nella letteratura penalistica, oltre a C. PIERGALLINI *Danno da prodotto e responsabilità penale*, cit., p. 16, v. M. DONINI, *Il volto attuale dell'illecito penale. La democrazia penale tra differenziazione e sussidiarietà*, Giuffrè, Milano, 2004, p. 107; J.M. SILVA SÁNCHEZ, *La expansión del derecho penal. Aspectos de la Política criminal en las sociedades postindustriales*, III ed., BdeF, Montevideo-Buenos Aires, 2011, p. 26; B. MENDOZA BUERGO, *El derecho penal en la sociedad del riesgo*, Civitas, Madrid, 2001, p. 24. Più in generale, sul tema del rischio, oltre già a V. MILITELLO, *Rischio e responsabilità penale*, Giuffrè, Milano, 1988, p. 55, C. PERINI, *Il concetto di rischio nel diritto penale moderno*, Giuffrè, Milano, 2010, p. 168.

[8] Sull'IA come "ultimo stadio" del percorso conflittuale della colpa nella società del rischio, C. PIERGALLINI, *Intelligenza artificiale: da "mezzo" ad "autore" del reato?*, cit., p. 1747.

[9] Per tutti: U. PAGALLO, *Saggio sui robot e il diritto penale*, cit., p. 595.

cattivo uso o da un difetto di costruzione o progettazione di un prodotto, di esso ne potrà rispondere – ove ovviamente ricorrano gli estremi della colpa – l'utilizzatore o il produttore umano [10]. La tipologia di prodotto considerato, in caso di particolare complessità del suo funzionamento, o del processo industriale in cui è costruito, al massimo potrà rendere più difficili e articolati i giudizi di causalità materiale, o – soprattutto – di colpa.

Ciò, tuttavia, non muta ancora il ruolo neutrale, silente, che la *res* assume nel fare da tramite fra il soggetto umano cui si giudica se imputare il fatto e l'accadimento lesivo stesso. Tale neutralità, infatti, si radica nella sostanziale *prevedibilità* del "comportamento" del prodotto-oggetto tradizionale. A fronte di determinate situazioni, o di determinati comandi, un prodotto – anche avente la veste di "macchina" complessa, ma non gestito da tecnologie dell'IA – reagirà sempre allo stesso modo, in base alla sua conformazione materiale o programmazione algoritmica; di talché l'uomo che lo gestisca o lo progetti è, a monte, nelle condizioni di rappresentarsi gli effetti provocati dalle proprie azioni per mezzo del prodotto stesso.

È un fatto ormai noto come il carattere "intelligente" dei soggetti artificiali che si avvalgono delle tecnologie in discussione conduca invece all'almeno parziale *imprevedibilità* del loro comportamento, a fronte di analoghi stimoli [11]. Elemento connaturato all'intelligenza artificiale, infatti, è la sua capacità di apprendimento, il *machine learning*. Esso, nel suo strutturale funzionamento, modifica i percorsi decisionali della "macchina" rispetto a quanto originariamente previsto in sede di programmazione [12].

Viene introdotta, così, un'*opacità tecnologica* al percorso imputativo tradizionale dell'evento colposo all'azione del soggetto umano che sta dietro al prodotto intelligente [13]: la "macchina" non è più un tramite neutrale, un puro mez-

[10] U. PAGALLO, *The Adventures of Picciotto Roboto: AI & Ethics in Criminal Law*, in AA.VV., *The Social Impact of Social Computing. Proceedings of the Twelth International Conference ETHICOMP 2011*, Sheffield, 2011, p. 352.

[11] S. BECK, *Google Cars, Software Agents, Autonomous Weapons Systems*, cit., p. 243; U. PAGALLO, *The Laws of Robots*, cit., p. 47.

[12] E. PALMERINI, voce *Robotica*, in E. SGRECCIA, A. TARANTINO (diretta da), *Enciclopedia di bioetica e scienza giuridica*, vol. X, ESI, Napoli, 2016, p. 1106; S. BECK, *Intelligent agents and criminal law*, cit., p. 140; I. SALVADORI, *Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale*, cit., p. 102. Sul *machine learning*, cfr. H. SURDEN, *Machine Learning and Law*, in *Washington Law Review*, 2014, p. 87; J. STILGOE, *Machine learning, social learning and the governance of self-driving cars*, in *Social Studies of Science*, 2018, p. 29.

[13] H. SURDEN, M.A. WILLIAMS, *Technological Opacity, Predictability, and Self-Driving Cars*, in *Cardozo Law Review*, 2016, p. 157; Y. BATHAEE, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, in *Harvard Journal of Law & Technology*, 2018, p. 889.

zo-oggetto, ma diviene almeno in parte soggetto a carattere autonomo e proattivo. Essa è, insomma, un "prodotto soggettivizzato", che non si limita più a realizzare la volontà umana che le sta dietro, ma agisce nel mondo in modo che non è più governato integralmente dalla mano dell'uomo.

2.3. *Il "*responsibility gap*" e le problematiche regolative dell'attribuzione di colpa: i riflessi sull'imputazione ai produttori umani*

Più aumenta il carattere "intelligente" di simili prodotti, più la portata di una simile imprevedibilità tecnologica è destinata ad aumentare. Così, lo scenario prossimo è evidentemente quello di un progressivo aumento del peso di una questione che ad oggi è ancora in parte speculativa, ma che in prospettiva futura parrebbe indirizzata ad assumere un rilievo applicativo affatto trascurabile.

L'opacità tecnologica si ripercuote infatti sul meccanismo tradizionale di addebito del reato colposo d'evento: l'imprevedibilità «*paralizza il giudizio di imputazione per colpa*» [14], generando un *responsibility gap*, un *vuoto di responsabilità* [15]. Via via che i margini di autonomia dei soggetti artificiali intelligenti aumenteranno, saranno infatti sempre più i possibili risultati lesivi dell'incolumità umana, provocati dal comportamento di questi, che rimarranno privi di "copertura" sul piano della responsabilità penale.

Si può accennare, in modo schematico, gli effetti che una simile evoluzione avrebbe sulle due figure umane più tipiche – già ricordate – che vengono in rilievo in relazione ai danni cagionati da prodotto. Da un lato, l'utente, l'*utilizzatore* del prodotto stesso. Dall'altro, a monte, il suo progettatore, programmatore, manufattore: in senso lato, il *produttore*.

Iniziando da quest'ultima figura, si potrebbe dire come le indubbie più ampie difficoltà sul piano imputativo che la riguardano rendono paradossalmente più semplice lo scenario delle scelte regolative rimesse alla politica criminale. L'eventualità di imputare al produttore umano possibili accadimenti lesivi connessi all'autonomia del prodotto appare infatti così complessa da non lasciare sostanzialmente alcun margine a scelte politiche che individuino soggetti umani cui addossare la responsabilità, la "colpa", in modo da occultare il problema del *responsibility gap*.

[14] C. PIERGALLINI, *Intelligenza artificiale: da "mezzo" ad "autore" del reato?*, cit., p. 1760.

[15] B. MAGRO, *Biorobotica, robotica e diritto penale*, cit., p. 515; cfr. altresì, più in generale, P. PAGALLO, S. QUATTROCOLO, *The impact of AI on criminal law, and its twofold procedures*, in W. BARFIELD, U. PAGALLO (eds), *Research Handbook on the Law of Artificial Intelligence*, Elgar, Cheltenham-Northampton, 2018, p. 385. In una prospettiva più ampia, J. DANAHER, *Robots, Law and the Retribution Gap*, in *Ethics and Information Technology*, 2016, p. 299.