

## ANTICIPAZIONI

---

**ALBERTO CAPPELLINI**

### ***Reati colposi e tecnologie dell'intelligenza artificiale\****

L'introduzione nella vita comune di strumenti basati sull'intelligenza artificiale prospetta un grosso impatto sulla dimensione dell'imputabilità, per colpa, di eventi avversi. L'estrema difficoltà del riconoscere profili di negligenza in capo a singoli soggetti umani afferenti alla rete organizzativa complessa che ha progettato l'IA la cui azione è causa dell'avvenimento negativo, in uno con l'insolubile problema di fondo dell'almeno parziale imprevedibilità dei soggetti artificiali intelligenti, generano una tensione a livello politico-criminale sulla figura dell'utilizzatore umano, sospeso tra istanze colpevoliste, di assicurazione sociale, e la pragmatica necessità di non frustrare i benefici connessi al loro uso. Nella parte conclusiva, si delineano alcune prospettive di possibile mitigazione - non soluzione - della problematica, mediante l'utilizzo di collettori di imputazione altri dal singolo agente umano.

*Criminal Negligence and AI technologies*

*The introduction of AI tools into everyday life promises to have a major impact on criminal liability in the event of tragic occurrences. First, it is extremely difficult to account for a negligent behaviour an individual belonging to the complex network that designed the AI whose action provoked the negative event. Second, the behaviour of AI subjects seems partially unpredictable, due to machine learning. So, the human user is suspended between the tendency to be blamed for the event, only for social reassurance, and the pragmatic need not to frustrate the benefits associated with AI technologies. In the final section, the article explores some ways in which the outlined problems could be alleviated (not resolved at all), with respect to AI's direct responsibility and corporate liability.*

**SOMMARIO:** 1. Introduzione. Autorità "artificiale intelligente" e responsabilità colposa. - 2. Macchine intelligenti e imprevedibilità tecnologica. - 3. Il "responsibility gap" e le problematiche regolative dell'attribuzione di colpa: i riflessi sull'imputazione ai produttori umani. - 4. (segue): i riflessi sull'imputazione agli utilizzatori umani. - 5. IA, colpa, caso fortuito: la politicità della soglia del rischio consentito e le influenze della precauzione. - 6. Riflessioni conclusive: quali prospettive per il futuro?

1. *Introduzione. Autorità "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

---

\* Il presente scritto trae origine dalla relazione tenuta il 29.09.2021 nell'ambito del webinar *La ricerca al servizio della complessità giuridica e dell'innovazione tecnologica*, organizzata dal Dipartimento di Giurisprudenza dell'Università degli Studi della Campania Luigi Vanvitelli, ed è destinato alla pubblicazione nei relativi atti. Si ringrazia l'editore Giappichelli per averne consentito l'anticipazione nella presente sede.

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<sup>1</sup>.

---

<sup>1</sup>Non sono pochi, ormai, i lavori generali sul tema, anche solo limitandosi alla lingua italiana. Per tutti: PIERGALLINI, *Intelligenza artificiale: da 'mezzo' ad 'autore' del reato?*, in *Riv. it. dir. proc. pen.*, 2020, 1743; SALVADORI, *Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale*, in *Riv. it. dir. proc. pen.*, 2021, 83; GIANNINI, *Intelligenza artificiale, human oversight e responsabilità penale: prove d'impatto a livello europeo*, in *Criminalia*, 2021, 249; MAGRO, *Robot, cyborg e intelligenze artificiali*, in *Cybercrime*, diretto da Cadoppi-Canestrari-Manna-Papa, Milano, 2019, 1179; EAD., *Biorobotica, robotica e diritto penale*, in *Genetics, robotics, law, punishment*, eds. by Provolo-Riondato-Yenisey, Padova, 2014, 499; RIONDATO, *Robotica e diritto penale (robots, ibridi, chimere e "animali tecnologici")*, ivi, 599; ID., *Robot: talune implicazioni di diritto penale*, in *Tecnodiritto. Temi e problemi di informatica e robotica giuridica*, a cura di Moro-Sarra, Milano, 2017, 85; PAGALLO, *Saggio sui robot e il diritto penale*, in *Scritti in memoria di Giuliano Marini*, a cura di Vinciguerra-Dassano, Napoli, 2010, 595; ID., *Robotica*, in *Manuale di informatica giuridica e diritto delle nuove tecnologie*, a cura di Pagallo-Durante, Torino, 2012, 141; BORSARI, *Intelligenza Artificiale e responsabilità penale: prime considerazioni*, in *MediaLaws*, 3/2019, 262; MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in *www.discrimen.it*, 15 maggio 2020, 2; RUFFOLO, *Machina delinquere potest? Responsabilità ed "illeciti" (anche penali?) della "persona elettronica" e tutele per gli agenti software autonomi*, in *XXVI Lezioni di Diritto dell'Intelligenza Artificiale*, a cura di Ruffolo, Torino, 2021, 295; SEVERINO, *Intelligenza artificiale e diritto penale*, in *Intelligenza artificiale. Il diritto, i diritti, l'etica*, a cura di Ruffolo, Milano, 2020, 533; EAD., *Le implicazioni dell'intelligenza artificiale nel campo del diritto con particolare riferimento al diritto penale*, in *Intelligenza artificiale, Politica, economia, diritto, tecnologia*, a cura di Severino, Roma, 2022; ARAGONA, *I Robot: the criminal liability of artificial intelligences*, in *TransJus Working Papers Publications*, 4/2019, 83; MINELLI, *La responsabilità "penale" tra persona fisica e corporation alla luce della Proposta di Regolamento sull'Intelligenza Artificiale*, in *Dir pen. cont.*, 2022, 2, 50. Per un inquadramento più ampio del tema entro quello più generale dei rapporti tra IA e giustizia penale, fondamentale BASILE, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *www.dirittopenaleuomo.org*, 29 settembre 2019, oltre al successivo ID., *Intelligenza artificiale e diritto penale: qualche aggiornamento e qualche nuova riflessione*, in *Il sistema penale ai confini delle hard sciences. Percorsi epistemologici tra neuroscienze e intelligenza artificiale*, a cura di Basile-Caterini-Romano, Pisa, 2021, 11 (nel cennato volume, esattamente sul tema va altresì ricordato il lavoro di MINELLI, *Quando l'autore del reato è un robot: tra vecchi modelli imputativi e nuovi possibili paradigmi di responsabilità penale*, 57, oltre a vari interessanti spunti negli altri contributi, fra cui LO MONTE, *Intelligenza artificiale e diritto penale: le categorie dominiche alla prova del futuribile*, 41). Più incentrato sulla sola tematica in questione un ulteriore contributo di BASILE, *Diritto penale e Intelligenza Artificiale*, in *Giur.it.*, 2019, suppl., 67. Sia consentito, infine, un riferimento fin d'ora, una volta per tutte, al nostro CAPPELLINI, *Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale*, in *Criminalia*, 2018, 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 sistematica-

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<sup>2</sup>.

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<sup>3</sup>.

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<sup>4</sup>, ma

---

mente 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.

<sup>2</sup>Fra la casistica “bizzarra” già verificatasi, si può ricordare, solo a titolo di esempio, l'utilizzo di sottomarini robotici per traffici di droga: PAGALLO, *The Laws of Robots. Crimes, Contracts and Torts*, Dordrecht, 2013, 65. Per un panorama generale, attuale e possibile futuro, degli *AI crimes*: KING-AGGARWAL-TADDEO-FLORIDI, *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions*, in *Science and Engineering Ethics*, 2020, 89; CALDWELL-ANDREWS-TANAY-GRIFFIN, *AI-enabled future crime*, in *Crime Science*, 2020.

<sup>3</sup>Fra i lavori specificamente dedicati al tema del rapporto tra autorità artificiale e reati colposi, per tutti: BECK, *Intelligent agents and criminal law – Negligence, diffusion of liability and electronic personhood*, in *Robotics and Autonomous Systems*, 2016, 138; EAD., *Google Cars, Software Agents, Autonomous Weapons Systems – New Challenges for Criminal Law*, in *Robotics, Autonomics and the Law*, eds. by Hilgendorf-Seidel, Baden-Baden, 2017, 227; MORATTI, *AI Crimes and Misdemeanors: Debating the Boundaries of Criminal Liability and Imputation*, in *Artificial Intelligence, Big Data and Automated Decision-Making in Criminal Justice*, eds. by Vermeulen-Peršak-Recchia, Antwerpen, 2021, 109 (nel medesimo volume va ricordato altresì PANATTONI, *AI and Criminal Law: The Myth of 'Control' in a Data-Driven Society*, 125, dal taglio tematico invero un po' più ampio).

<sup>4</sup>Nella penalistica italiana, sulla responsabilità per danno da prodotto cfr., per tutti: PIERGALLINI, *Danno*

che in tale scenario – a ben vedere – pare estendersi fino a coinvolgere ambiti di responsabilità per colpa più classici e tradizionali<sup>5</sup>.

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<sup>6</sup>.

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

---

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

<sup>5</sup>Quali – come si dirà meglio *infra* – la colpa stradale, la colpa medica, e così via.

<sup>6</sup>Due gli episodi più ricordati: Williston (Florida, USA), 7.5.2016, 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) 18.3.2018, una vettura Uber completamente autonoma investiva un pedone, provocandone la morte. Ma ne sono avvenuti altri anche altri. Alcuni mortali: Mountain View (California, USA) 23.3.2018; Houston (Texas, USA), 17.4.2021. Molti altri – difficili peraltro anche da tracciare in fonti ufficiali e non – con conseguenze meno gravi.

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 tengono ampiamente conto anche dalla prospettiva dell'eventualità ultima, più recondita e negativa<sup>7</sup>.

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 avvicinarsi 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

---

<sup>7</sup>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 collettiva. Per tutti, nella letteratura sociologica, BECK, *La società del rischio. Verso una seconda modernità* (1986), Roma, 2000, 38; GIDDENS, *Le conseguenze della modernità* (1990), Bologna, 1994, 125; LUHMANN, *Sociologia del rischio* (1991), Milano, 1996, 40; nella letteratura penalistica, oltre a PIERGALLINI, *Danno da prodotto e responsabilità penale*, cit., 16, v. DONINI, *Il volto attuale dell'illecito penale. La democrazia penale tra differenziazione e sussidiarietà*, Milano, 2004, 107; SILVA SÁNCHEZ, *La expansión del derecho penal. Aspectos de la Política criminal en las sociedades postindustriales*<sup>8</sup>, Montevideo-Buenos Aires, 2011, 26; MENDOZA BUERGO, *El derecho penal en la sociedad del riesgo*, Madrid, 2001, 24. Più in generale, sul tema del rischio, oltre già a MILITELLO, *Rischio e responsabilità penale*, Milano, 1988, 55, PERINI, *Il concetto di rischio nel diritto penale moderno*, Milano, 2010, 168.

per mano di un soggetto artificiale intelligente – fa emergere in modo ancor più incisivo<sup>8</sup>.

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<sup>9</sup>.

Nei reati colposi, in particolare, qualora l’evento lesivo sia provocato da un 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<sup>10</sup>. 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.

---

<sup>8</sup>Sull’IA come “ultimo stadio” del percorso conflittuale della colpa nella società del rischio, PIERGALLINI, *Intelligenza artificiale: da ‘mezzo’ ad ‘autore’ del reato?*, cit., 1747.

<sup>9</sup>Per tutti: PAGALLO, *Saggio sui robot e il diritto penale*, cit., 595.

<sup>10</sup>PAGALLO, *The Adventures of Picciotto Roboto: AI & Ethics in Criminal Law*, in AA.VV., *The Social Impact of Social Computing. Proceedings of the Twelfth International Conference ETHICOMP 2011*, Sheffield, 2011, 352.

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<sup>11</sup>. 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<sup>12</sup>.

Viene introdotta, così, un’*opacità tecnologica* al percorso imputativo tradizionale dell’evento colposo all’azione del soggetto umano che sta dietro al prodotto intelligente<sup>13</sup>: la “macchina” non è più un tramite neutrale, un puro mezzo-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.

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.

---

<sup>11</sup>BECK, *Google Cars, Software Agents, Autonomous Weapons Systems*, cit., 243; PAGALLO, *The Laws of Robots*, cit., 47.

<sup>12</sup>PALMERINI, voce *Robotica*, in *Enciclopedia di bioetica e scienza giuridica*, diretta da Sgreccia-Tarantino, X, Napoli, 2016, 1106; BECK, *Intelligent agents and criminal law*, cit., 140; SALVADORI, *Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale*, cit., 102. Sul *machine learning*, cfr. SURDEN, *Machine Learning and Law*, in *Washington Law Review*, 2014, 87; STILGOE, *Machine learning, social learning and the governance of self-driving cars*, in *Social Studies of Science*, 2018, 29.

<sup>13</sup>SURDEN-WILLIAMS, *Technological Opacity, Predictability, and Self-Driving Cars*, in *Cardozo Law Review*, 2016, 157; BATHAEE, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, in *Harvard Journal of Law & Technology*, 2018, 889.

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*»<sup>14</sup>, generando un *responsibility gap*, un *vuoto di responsabilità*<sup>15</sup>. 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, manifattore: in senso lato, il *produttore*.

Iniziando da quest'ultima figura, si potrebbe dire come le indubbe 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*.

In effetti, la tipologia di eventi avversi evocata non riguarda quelle tipologie di colpa del "produttore" che si sostanziano in un difetto di manifattura, o di manutenzione<sup>16</sup>, di vero e proprio errore nella programmazione<sup>17</sup>, o anche di mancato aggiornamento della stessa<sup>18</sup>; le quali potrebbero essere ancora inquadrati

<sup>14</sup>PIERGALLINI, *Intelligenza artificiale: da 'mezzo' ad 'autore' del reato?*, cit., 1760.

<sup>15</sup>MAGRO, *Biorobotica, robotica e diritto penale*, cit., 515; cfr. altresì, più in generale, PAGALLO-QUATTROCOLO, *The impact of AI on criminal law, and its twofold procedures*, in *Research Handbook on the Law of Artificial Intelligence*, a cura di Barfield-Pagallo, Cheltenham-Northampton, 2018, 385. In una prospettiva più ampia, DANAHER, *Robots, Law and the Retribution Gap*, in *Ethics and Information Technology*, 2016, 299.

<sup>16</sup>Da intendersi ovviamente come esecuzione della manutenzione, rientrando la sua omissione, ove doverosa (se periodicamente richiesta, oppure al manifestarsi di un guasto o difetto prevedibilmente pericoloso) nelle eventuali responsabilità del proprietario-utilizzatore.

<sup>17</sup>Ad esempio, mediante inserimento di dati errati.

<sup>18</sup>È il caso dell'omessa considerazione di nuovi dati emersi dalla ricerca o sviluppo, oppure dalla casistica



secondo le (invero già assai complesse) cadenze “ordinarie” di responsabilità del produttore.

La casistica che qui rileva, piuttosto, riguarda quelle ipotesi in cui il danno derivi da un comportamento – come detto – *imprevedibile* del soggetto artificiale, attribuibile a quel margine di autonomia lui riservato, in cui è chiamata in causa la responsabilità del produttore nella persona del *programmatore*.

In tali ipotesi, se anche solo fossero superabili le difficoltà connesse all’individuazione del singolo (o dei singoli) soggetti umani responsabili all’interno di una rete plurisoggettiva di produzione-programmazione la cui estensione e complessità è ben intuibile, residuerebbe comunque il problema di fondo per cui il comportamento *imprevedibile* della macchina, come tale, per definizione non può certo essere rimproverato *per colpa* al programmatore<sup>19</sup>. E anche il dubbio circa il carattere più o meno prevedibile di tale azione dovrebbe far propendere per una soluzione liberatoria della figura umana astrattamente individuabile.

L’effetto complessivo, così, è che tale problematica, solo apparentemente circoscritta, peraltro installandosi in un quadro già di elevata complessità quanto all’attribuzione di colpevolezza, conduce alla sostanziale impossibilità di individuare soggetti umani la cui “colpa” possa colmare il vuoto di responsabilità di matrice tecnologica.

4. (*segue*): *i riflessi sull’imputazione agli utilizzatori umani*. Diverso è lo scenario che riguarda la figura dell’utilizzatore umano delle tecnologie dell’IA, la quale – in ragione della sua prossimità rispetto al fatto – conserva la possibilità di essere responsabilizzata sul piano penalistico anche in ipotesi di elevata automazione, e rimane pertanto aperta, in punto di disciplina, a subire le influenze derivanti dalla sensibilità politico-criminale sull’argomento.

In generale, la fattispecie dell’utilizzatore si caratterizza per il fatto che la “novità” rispetto al passato recente, che è data dal carattere intelligente del prodotto

---

concreta (ad esempio, il verificarsi di un sinistro da parte di un soggetto IA di analogo modello).

<sup>19</sup>BECK, *Google Cars, Software Agents, Autonomous Weapons Systems*, cit., 244; PAGALLO, *The Laws of Robots*, cit., 72.

di cui questi fa uso, si innesta su un genere di attività che invece, tradizionalmente, era in tutto e per tutto umana (la guida su strada<sup>20</sup>, l'attività medico-chirurgica<sup>21</sup>, e così via).

Lo specifico fine dell'introduzione della tecnologia dell'intelligenza artificiale è evidentemente quello di affiancare, coadiuvare chi compia tale attività, permettendogli di farla con minor impegno e fatica, oppure con maggiore accuratezza, o comunque utilizzando risorse e capacità inaccessibili a un soggetto umano.

Almeno in un primo momento, un simile ausilio non è progettato per giungere a sostituire *tout court* l'uomo. Sviluppi tecnologici ulteriori, con l'approdo a un'autonomia del soggetto artificiale piena e completa, rendono possibile l'eventualità di far transitare il ruolo dell'uomo a quello di mero controllore di un'attività della "macchina" del tutto indipendente; se non addirittura a quella di mero fruitore passivo di un servizio automatizzato, privo financo della possibilità di interferire con l'azione artificiale.

L'esempio più immediato di tale paradigma è quello delle cosiddette auto a guida autonoma, i cui principali prototipi – ad oggi – in sviluppo presentano livelli di automazione tali da consentire una circolazione su strada in parte anche senza l'intervento diretto dell'uomo, benché ancora (almeno in Italia) sia

---

<sup>20</sup>La letteratura penalistica sulle auto a guida autonoma, in ambito internazionale e in special modo statunitense, è sterminata. Senza la minima pretesa di completezza, per tutti: HILGENDORF, *Automated Driving and the Law*, in *Robotics, Autonomics and the Law*, eds. by Hilgendorf-Seidel, Baden-Baden, 2017, 171; DOUMA-PALODICHUK, *Criminal Liability Issues Created by Autonomous Vehicles*, in *Santa Clara Law Review*, 2012, 1157; GURNEY, *Driving into the Unknown: Examining the Crossroads of Criminal Law and Autonomous Vehicles*, in *Wake Forest Journal of Law & Policy*, 2015, 393; WESTBROOK, *The Google Made Me Do It: The Complexity of Criminal Liability in the Age of Autonomous Vehicles*, in *Michigan State Law Review*, 2017, 97. Nella letteratura italiana, PICOTTI, *Profili di responsabilità penale per la circolazione di veicoli a guida autonoma*, in *Studi in onore di Antonio Fiorella*, I, Roma, 2021, 813; sia poi consentito il rinvio, una volta per tutte, al nostro CAPPELLINI, *Profili penalistici delle self-driving cars*, in *Dir pen. cont.*, 2019, 2, 325, anche per più ampi riferimenti bibliografici.

<sup>21</sup>Sul tema dei rapporti tra attività terapeutiche e impiego dell'IA in medicina (in ottica tuttavia prevalentemente civilistica), la letteratura è ormai ampia: per tutti, COLLETTI, *Intelligenza artificiale e attività sanitaria. Profili giuridici dell'utilizzo della robotica in medicina*, in *Rivista di Diritto dell'Economia, dei Trasporti e dell'Ambiente*, 2021, 201; HARNED-LUNGREN-RAJPURKAR, *Machine Vision, Medical AI, and Malpractice*, in *Harvard Journal of Law & Technology Digest*, 2019; GRIFFIN, *Artificial Intelligence and Liability in Health Care*, in *Health Matrix: The Journal of Law-Medicine*, 2019, 65; in un'ottica penalistica, PERIN, *Standardizzazione, automazione e responsabilità medica. Dalle recenti riforme alla definizione di un modello d'imputazione solidaristico e liberale*, in *Rivista di BioDiritto*, 2019, 1, 229.

imposta la costante sorveglianza dell'“utilizzatore”<sup>22</sup>. Non mancano, tuttavia, modelli ancor più avanzati, dotati di un tale livello di autonomia da poter fare a meno degli stessi comandi di guida.

Ora, fintantoché permane una figura umana di “utilizzatore” con un potere materiale di intervento, di governo, sull'azione autonoma del soggetto artificiale, tale da sovrascriverne gli schemi algoritmici e correggerne il comportamento anche “in corsa”, durante il suo svolgimento, il problema del *responsibility gap* generato dalle tecnologie dell'IA rimane ancora in larga misura nascosto.

Infatti, la persistenza di un potere materiale ed effettivo di intervento dell'uomo fa scivolare la questione sul piano tutto politico di se, come ed eventualmente in che misura sia opportuno gravare questi di un vero e proprio *obbligo giuridico* di controllo e intervento rispetto all'attività del soggetto artificiale intelligente.

Ove si ritenga in via interpretativa che un simile obbligo sussista<sup>23</sup>, l'utilizzatore umano potrà essere ritenuto responsabile degli eventi avversi derivati dal malfunzionamento dell'IA, avendo tenuto un comportamento, *lato sensu* omissivo<sup>24</sup>, integrante gli estremi della colpa penale. E ciò può valere - al limite -

---

<sup>22</sup> Il Codice della strada italiano, al momento, non ha subito modifiche, di talché dovranno ancora ritenersi “veicoli” - a mente dell'art. 46 - “*tutte le macchine di qualsiasi specie, che circolano sulle strade guidate dall'uomo*”. Sono pertanto da ritenersi finora ammissibili solo dei meccanismi minori di automazione, sostanzialmente di “guida assistita”, tali che l'uomo permanga saldamente al governo del mezzo.

<sup>23</sup> Desumendolo dalle norme che attribuiscono all'utilizzatore umano il governo dell'IA (ad esempio, il citato art. 46 C.d.S. per le auto a guida autonoma). Oppure, più radicalmente, si giunga a imporlo espressamente a livello di diritto positivo.

<sup>24</sup> A stretto rigore, la condotta potrà dirsi omissiva, in senso proprio, solo quando nella fattispecie concreta l'autonomia dell'IA sia sostanzialmente integrale, limitandosi l'utilizzatore a sorvegliare un'attività “altrui”, mentre dovrà ancora qualificarsi come commissiva quando l'uomo conservi, nella fisiologia del fenomeno, almeno in parte la necessità di “portare avanti” l'azione con ragionevole continuità, pur agevolato dal mezzo artificiale intelligente. Ma anche in tale ultimo caso, il fatto che qui interessa rimane pur sempre l'omissione del controllo e/o dell'intervento correttivo rispetto a un agire di mano del soggetto artificiale, anche se inserito in un “contesto” di azione complessivamente commissiva. Del resto, i confini tra azione commissiva e omissiva tendono a sfumare in ambito colposo, considerato il comune carattere normativo della colpa e dell'omissione, nonché la presenza di un momento omissivo in ogni condotta colposa: GIUNTA, *Illiceità e colpevolezza nella responsabilità colposa*, Padova, 1993, 90; VENEZIANI, *Regole cautelari “proprie” ed “improprie” nella prospettiva delle fattispecie colpose causalmente orientate*, Padova, 2003, 44.

anche ove l'attività del soggetto artificiale giunga *davvero* ad essere sostanzialmente autonoma - ovvero senza necessità di intervento pratico, anche solo saltuario, dell'utilizzatore - ma tuttavia permanga, anche solo in astratto, la possibilità di intervento sostitutivo dell'uomo.

Lo scenario muta soltanto dove - come accennato - la tecnologia utilizzata raggiunga un livello di sviluppo tale da escludere del tutto la stessa materiale possibilità di intervento dell'uomo: si pensi all'ipotesi di un *robotaxi* con clienti che, come in tutti i taxi, siano meri passeggeri, o comunque a vetture private a guida autonoma prive del volante e dei pedali, il cui proprietario non possa definirsi conducente ma anch'egli - appunto - mero passeggero.

In tale prospettiva-limite, il *responsibility gap* si mostra senza veli: venendo meno la possibilità di imputare un eventuale risultato lesivo a un utilizzatore che non partecipa più in alcun modo dell'azione pericolosa, addirittura ormai privato della capacità di governarla, non rimane nient'altro che un problematissimo "fatto proprio" del soggetto artificiale. Oppure - dipende dai punti di vista - un caso fortuito: un fatto che la vittima subisce, senza che vi sia la possibilità di attribuirne la responsabilità a qualcuno, e che rimane pertanto un frutto della casualità.

5. *IA, colpa, caso fortuito: la politicità della soglia del rischio consentito e le influenze della precauzione.* In un simile scenario, non sorprende come le "ragioni" della nuova tecnologia dell'IA non siano le sole a giocare un ruolo di primo piano in quel complesso e articolato giudizio politico di bilanciamento tra interessi che individua, a livello sociale prima e giuridico poi, la soglia del rischio consentito, ripartendo le aree da un lato del caso fortuito (o, che dir si voglia, del "fatto" che ricade solo sulla vittima) e, dall'altro, dell'attribuibilità a qualcuno per colpa.

Ciò accade nonostante, su un piano strettamente oggettivo, siano vari e validi gli argomenti che depongono in favore di una politica quanto più ampia possibile di autorizzazione di simili tecnologie e di sollevazione degli utenti dalle responsabilità penali conseguenti al loro maneggio. Statisticamente parlando,

in effetti, si potrebbe dire come le “macchine” – già adesso – siano spesso *più sicure* della mano dell’uomo che mirano a sostituire.

Anzitutto, c’è da dire che, se a livelli di automazione più basilari la sorveglianza e l’eventuale intervento umano possono ancora garantire dei livelli di sicurezza oggettivamente più elevati, a stadi tecnologici più avanzati l’interferenza umana può rivelarsi addirittura pericolosa, se non essa stessa possibile autonoma fonte di incidenti.

Più la macchina è capace di fare da sé, più il controllo umano si riduce ad impedire una prospettiva, sempre più improbabile e remota, di malfunzionamento. Ma allora, in certi casi saranno inevitabili i cali di attenzione e di prontezza dell’utente umano in un contesto di prolungata mera sorveglianza passiva (si pensi al settore, ancora, delle macchine a guida autonoma<sup>25</sup>). Similmente, è irrealistico pensare che un uomo sempre più disabituato a compiere da solo le attività in questione sia efficacemente in grado di intervenire proprio in momenti di crisi, a elevata difficoltà<sup>26</sup>. E ancora, rispetto ad altri settori di attività (ad esempio, in campo medico), è irrealistico immaginare che l’uomo possa davvero sopperire al soggetto artificiale dove il *quid pluris* di quest’ultimo sia una precisione o accuratezza fisiologicamente impossibili per la mano umana.

Poi, rimane sempre il problema della riconoscibilità, da parte del controllore, dell’“avaria”, del momento in cui è necessario intervenire. È facile immaginare come la difficile “leggibilità” del comportamento artificiale agli occhi umani possa condurre a interventi correttivi – istintivi, dettati da paura o altro – non necessari o errati, talvolta essi stessi forieri di pericolo<sup>27</sup>.

Ma anche al di là della sostanziale inefficacia oggettivo-preventiva della sorveglianza umana, va infine considerato come la *ratio* stessa dell’innovazione tecnologica, in molti settori, verrebbe sostanzialmente frustrata ove la figura umana conservasse una responsabilità *omnibus* per l’attività artificiale. Ciò non

---

<sup>25</sup> ARIA-OLSTAM-SCHWIETERING, *Investigation of Automated Vehicle Effects on Driver’s Behaviour and Traffic Performance*, in *Transportation Research Procedia*, 2016, 761.

<sup>26</sup> DOUMA-PALODICHUK, *Criminal Liability Issues Created by Autonomous Vehicles*, cit., 1164.

<sup>27</sup> HEVELKE-NIDA-RÜMELIN, *Responsibility for Crashes of Autonomous Vehicles: An Ethical Analysis*, in *Science and Engineering Ethics*, 2015, 624; GURNEY, *Driving into the Unknown*, cit., § 16.

consentirebbe, nella sostanza, il minimo sgravio di impegno per l'uomo, limitandosi a trasformare un dovere di attenzione nell'esecuzione proattiva di un'attività pericolosa in un non meno capillare e continuo dovere di controllo rispetto al compimento di un'azione "altrui".

Eppure, il carattere - come detto - tutto politico della questione fa sì che gli argomenti ripercorsi siano solo alcuni di quelli che entrano davvero in gioco nella scelta regolativa circa fino a che punto consentire l'utilizzo pubblico di soggetti artificiali intelligenti e - soprattutto - *se e in che misura* esonerare da responsabilità gli utilizzatori umani - gli unici soggetti, come visto, davvero strutturalmente capaci di caricarsi della "colpa" - non imponendo loro un obbligo di controllo onnicomprensivo.

Non si può non tenere conto, infatti, del ruolo capitale che la paura dell'ignoto tecnologico ha nel determinare le scelte compiute a livello regolativo, e in particolare del potere di piegarle a logiche ispirate a criteri, anche estremi, di precauzione<sup>28</sup>.

In ogni società l'imputazione di un evento avverso a taluno, l'individuazione di un colpevole, svolge un ruolo di rassicurazione del gruppo collettivo. È un esorcismo antropologico-giudiziario il quale, così spiegando il verificarsi di una sciagura, dà conto del fatto che essa si è verificata per "colpa" di qualcuno, presupponendo che senza tale "colpa" nulla di tragico sarebbe accaduto, e che la realtà avrebbe preservato intatta la sua illusione di sicurezza<sup>29</sup>. Senza la spiegazione data dell'attribuzione di "colpa", invece, quell'evento, quel male, diventa il frutto beffardo del caso, privo di spiegazioni, indominabile: come tale, fonte di angoscia per l'uomo<sup>30</sup>.

---

<sup>28</sup> Sul principio di precauzione in ambito penalistico, per tutti: CASTRONUOVO, *Principio di precauzione e diritto penale: paradigmi dell'incertezza nella struttura del reato*, Roma, 2012; FORTI, "Accesso" alle informazioni sul rischio e responsabilità: una lettura del principio di precauzione, in *Criminalia*, 2006, 155; GIUNTA, *Il diritto penale e le suggestioni del principio di precauzione*, in *Criminalia*, 2006, 227.

<sup>29</sup> La letteratura extragiuridica sul punto è sterminata. Per tutti, DOUGLAS, *Purezza e pericolo. Un'analisi dei concetti di contaminazione e tabù* (1970), Bologna, 2014; ID., *Rischio e colpa* (1992), Bologna, 1996; REMOTTI, *Maleficio*, in *I concetti del male*, a cura di Portinaro, Torino, 2002, 148; BAUMAN, *Paura liquida*, Roma-Bari, 2006, 69; SOFSKY, *Rischio e sicurezza*, Torino, 2005, 12.

<sup>30</sup> In dettaglio, nella letteratura penalistica, CASTRONUOVO, *La colpa penale*, Milano, 2009, 86.

Tale atteggiamento, intimamente connesso alla natura umana più profonda, nel confronto storico con nuove tecnologie e correlati pericoli tende a perpetuarsi, individuando – quando non sia possibile neutralizzare i nuovi rischi con divieti *tout court* – potenziali colpevoli cui addossare gli inevitabili conseguenti risultati di danno, rifuggendo così il timore connesso al vuoto di responsabilità<sup>31</sup>. Nel caso delle tecnologie dell'IA, pertanto, sulla falsariga di quanto appena detto, il rischio tangibile è dunque duplice. Anzitutto, quello di attendismo, in particolare con riferimento al consentire l'utilizzo di quei soggetti artificiali con livelli di autonomia particolarmente elevati, nelle scelte regolatorie che i singoli Stati sono chiamati a compiere in merito alla portata e ai limiti che simili tecnologie risconteranno nel mercato; con evidente pregiudizio dei benefici che pure le stesse prospettano all'uomo.

Ma soprattutto, è evidente il rischio di scelte politiche “ultraresponsabiliste”, volte a prorogare il ruolo e i doveri di controllo ad ampio spettro, sussistenti in capo a utilizzatori umani, in maniera del tutto irrealistica: considerato come, in un simile contesto, è palese come gli stessi abbiano ormai perduto il reale governo dell'attività in questione. E ciò, sia con potenziali derive – sul piano penalistico – di inammissibili responsabilità oggettive, da posizione (*rectius*, da prossimità, *ictu oculi* evidente, al pericolo); sia – più in generale – ostacolando i benefici per l'uomo connessi a tali tecnologie, la cui diffusione viene evidentemente disincentivata da simili politiche.

6. *Riflessioni conclusive: quali prospettive per il futuro?*<sup>9</sup> Come detto, il timore di vuoti di tutela per i beni primari della vita, sicurezza e salute umana può ben alimentare la permanenza di politiche precauzionistiche e ultraresponsabiliste a danno dei protagonisti umani più immediatamente coinvolti, gli utilizzatori. Così, a fronte di tale problema, allo stato apparentemente irrisolvibile, è imme-

---

<sup>31</sup> STORTONI, *Angoscia tecnologica ed esorcismo penale*, in *Riv. it. dir. proc. pen.*, 2004, 71; STELLA, *Giustizia e modernità. La protezione dell'innocente e la tutela delle vittime*<sup>3</sup>, Milano, 2003, 552; CENTONZE, *La normalità dei disastri tecnologici. Il problema del congedo dal diritto penale*, Milano, 2004, 35; CIVELLO, *La “colpa eventuale” nella società del rischio. Epistemologia dell'incertezza e “verità soggettiva” della colpa*, Torino, 2013, 197.

diato interrogarsi circa le prospettive – non certo di superamento, ma quantomeno di mitigazione – di tali criticità, che potrebbero, in futuro, essere esplorate.

La via indubbiamente più suggestiva, all'apparenza fantasiosa ma – in particolare in prospettiva futura – forse non poi così tanto, è quella dell'introdurre delle forme di responsabilità dirette del soggetto artificiale, ponendo in dubbio l'assioma, finora incontestato, per il quale *machina delinquere non potest*. E' una prospettiva con tutta evidenza non immediatamente percorribile nelle sue forme più radicali, giacché le macchine di oggi non posseggono ancora quella "libertà del volere" che è il presupposto di un giudizio di colpevolezza, facendo perdere così di significato e di efficacia una pena criminale che sia loro comminata, ed eventualmente irrogata, in caso di integrazione dell'illecito<sup>32</sup>.

Tuttavia – oltre a non chiudere la porta ad eventuali sviluppi tecnologici futuri, in realtà non del tutto prevedibili, che potrebbero condurre alla creazione di IA ancor più simili all'uomo – la riflessione può forse già oggi utilmente svolgersi, agganciandosi a una prospettiva che sia diversa da quella della penalità "classica": la quale è inevitabilmente legata, ancora oggi, ai concetti di colpevolezza e retribuzione.

---

<sup>32</sup>Nella letteratura internazionale, la principale voce favorevole alla configurabilità di una responsabilità penale diretta dell'IA è Gabriel HALLEVY: *Liability for Crimes Involving Artificial Intelligence Systems*, Dordrecht, 2015; *Virtual Criminal Responsibility*, in *Original Law Review*, 2010, 6; "I, Robot – I, Criminal" – *When Science Fiction Becomes Reality: Legal Liability of AI Robots committing Criminal Offences*, in *Syracuse Science & Technology Law Reporter*, 2010, 1. Scettica la letteratura assolutamente prevalente. Sul tema, per tutti, HILDEBRANDT, *Ambient Intelligence, Criminal Liability and Democracy*, in *Criminal Law and Philosophy*, 2008, 163; ASARO, *A Body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics*, in *Robot Ethics*, eds. by Lin-Abney-Bekey, Cambridge (Massachusetts), 2012, 169; FREITAS-ANDRADE-NOVAIS, *Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible*, in *AI Approaches to the Complexity of Legal Systems*, eds. by Casanovas-Pagallo-Palmirani-Sartor, Berlin-Heidelberg, 2014, 145; LIMA, *Could AI Agents Be Held Criminally Liable? Artificial Intelligence and the Challenges for Criminal Law*, in *South Carolina Law Review*, 2018, 677; SIMMLER-MARKWALDER, *Guilty Robots? – Rethinking the Nature of Culpability and Legal Personhood in an Age of Artificial Intelligence*, in *Criminal Law Forum*, 2019, 1; ABBOTT-SARCH, *Punishing Artificial Intelligence: Legal Fiction or Science Fiction*, in *University of California Davis Law Review*, 2019, 323; LEMLEY-CASEY, *Remedies for Robots*, in *The University of Chicago Law Review*, 2019, 1311; OSMANI, *The Complexity of Criminal Liability of AI Systems*, in *Masaryk University Journal of Law and Technology*, 2020, 53. Cfr. altresì – nella letteratura italiana – l'ampia trattazione del tema a penna di PIERGALLINI, *Intelligenza artificiale: da 'mezzo' ad 'autore' del reato?*, cit., 1764.



La storia, invero, ha conosciuto in epoche preilluministiche delle penalità diverse – degli *animali* prima di tutto, ma anche delle *cose* – che evidentemente assolvevano a funzioni del tutto differenti rispetto a quella di punire il male, liberamente commesso, attraverso un analogo male<sup>33</sup>. Si trattava, probabilmente, di funzioni sinceramente *vendicatorie*, di vera e propria rappresaglia; ma anche – sotto altro angolo visuale – *consolatorie*, nella misura in cui miravano a ristorare la vittima del dolore subito, mostrando plasticamente, attraverso la natura *lato sensu* rituale, solenne, della giurisdizione penale, la solidarietà della comunità di fronte all’evento nefasto, al ricadere di un male in capo a un proprio membro<sup>34</sup>.

Analogamente, si è suggerito da parte di alcuni come l’esercizio di una “punizione” nei confronti delle macchine che abbiano cagionato un male possa in qualche modo apportare un beneficio psicologico alla vittima (o alle persone a lei vicine), dando voce alla loro irrazionale brama di vendetta<sup>35</sup>. Si tratterebbe – ci pare – di una sorta di penalità “decolpevolizzata”, intermedia tra la (ritualmente neutra) gestione amministrativa di interessi e la penalità classica, colpevole: costituzionalmente possibile, nei confronti delle “macchine”, perché non ancora soggetti di diritto al pari dell’uomo, dunque “sacrificabili” anche in mancanza di colpevolezza allo specifico fine di apportare un beneficio a un soggetto di diritto “pieno”.

Un’altra via, certamente un po’ più immediata, che potrebbe valer la pena esplorare è quella della responsabilità da reato dei soggetti collettivi<sup>36</sup>. Dietro

---

<sup>33</sup> Sul tema, per tutti, cfr. il classico EVANS, *The Criminal Prosecution and Capital Punishment of Animals* (1906), trad. it. *Animali al rogo*, Milano, 2012.

<sup>34</sup> Si tratterebbe di una sorta di versione primigenia della moderna prevenzione generale positiva. O, forse, di una sua declinazione – legata alla forza del *rito*, della gestualità sociale – che, in modo sotterraneo, ha invero sempre pervaso la penalità. In questo senso, essa potrebbe svolgersi, un domani, anche nei confronti di potenziali “autori artificiali intelligenti”, offrendo ristoro alle vittime di questi.

<sup>35</sup> MULLIGAN, *Revenge Against Robots*, in *South Carolina Law Review*, 2018, 579. Non si è mancato tuttavia di contestare simili impostazioni, su un piano anzitutto etico, ritenendo che istituzionalizzare simili “vendette” educerebbe alla violenza, con il rischio che questa finisca per coinvolgere anche vittime umane: DARLING, *Extending Legal Protection to Social Robots: The Effects of Antropomorphism, Empathy, and Violent Behaviour Towards Robotic Objects*, in *Robot Law*, eds. by Calo-Froomkin-Kerr, Cheltenham-Northampton (Massachusetts), 2016, 213; *contra*, SWEENEY, *Why Indirect Harms do not Support Social Robot Rights*, in *Minds & Machines*, 2022.

<sup>36</sup> Sull’implementazione di una “responsabilità 231” per i fatti colposi commessi da IA, sia sotto un profilo dogmatico che politico-criminale, PIERGALLINI, *Intelligenza artificiale: da ‘mezzo’ ad ‘autore’ del reato?*,

ogni (futura) IA di larga diffusione sociale c'è infatti inevitabilmente una società commerciale che riassume l'insieme delle competenze tecnico-specialistiche necessarie a "tenere in piedi" progetti di sviluppo e produzione così complessi. La natura collettiva di tale soggetto, peraltro, parrebbe permettere di bypassare il problema altrimenti insolubile – poc'anzi evocato – dell'individuazione dei singoli programmatori, nell'enorme rete plurisoggettiva del circuito di produzione, che non hanno previsto ciò avrebbero dovuto prevedere. Il nome del "colpevole" si ritroverebbe, banalmente, inciso sulla carrozzeria del soggetto robotico...

Le cose, tuttavia, sono più complesse di quanto simili tentazioni semplificatorie facciano sperare. Certamente, un errore di programmazione – non univocamente riconducibile a singoli soggetti determinati – potrà ben essere attribuito alla responsabilità dell'ente quando lo stesso si radichi in un deficit organizzativo del soggetto collettivo: una *colpa* della *corporation* che, così, rimpiazza l'impraticabile colpa penale del singolo. Ciò, tuttavia, non elimina affatto il problema principale: il *responsibility gap*. *Il comportamento imprevedibile della macchina*, come tale, è e rimane non imputabile anche all'ente: la cui colpevolezza si deve radicare in una mancanza organizzativa relativa a dei rischi che siano comunque *prevedibili ex ante*.

Insomma: la responsabilità da reato degli enti sarà sì uno strumento indispensabile della *governance* pubblica dei prodotti IA (a fianco di quello risarcitorio e, ovviamente, di quello strettamente amministrativo, regolatorio/sanzionatorio) per correggere molte carenze e difetti sul piano dell'organizzazione, programmazione e produzione. Essa, tuttavia, risulta al contempo insufficiente a fornire una copertura di tutela, con finalità strettamente *preventive*, in relazione a *tutti* i rischi: e in particolare a quelli – massimamente problematici – di derivazione tecnologica.

---

cit., 1753. Cfr. altresì FED. MAZZACUVA, *The Impact of AI on Corporate Criminal Liability: Algorithmic Misconduct in the Prism of Derivative and Holistic Theories*, in *Artificial Intelligence, Big Data and Automated Decision-Making in Criminal Justice*, eds. by Vermuelen-Peršak-Recchia, Antwerpen, 2021, 143; DIAMANTIS, *The Extended Corporate Mind: When Corporations Use AI to Break the Law*, in *North Carolina Law Review*, 2020, 893.

E in ogni caso, anche volendo mettere da parte una simile insufficienza sul piano della prevenzione, residua comunque il diverso problema sopra accennato: ovvero se l'alternativa della responsabilità collettiva fornisca un valido sostituto all'imputazione del singolo sotto il diverso angolo visuale della necessità di *rassicurazione sociale*, di irrazionale esorcismo collettivo, connessa all'evento avverso di matrice tecnologica. Può una sanzione sostanzialmente pecuniaria all'ente surrogare la classica penalità individuale, con la sua "rassicurante" definitività nel "dare le colpe" e spiegare, così, il perché dell'evento avverso?

È un interrogativo aperto, cui forse si è istintivamente portati a dare risposta del tutto negativa, lasciando così irrisolto il *responsibility gap*; ma che, al contrario, meriterebbe probabilmente un più approfondito esame, al pari dei molti altri interrogativi che popolano un'area di frontiera solo apparentemente lontana e futuristica.