

INTRODUZIONE

1. Il dibattito sulla regolazione dell'intelligenza artificiale (i.a.) attraversa una molteplicità di questioni che coinvolgono le più disparate dimensioni del *vivere sociale*, interrogando nel profondo l'*idea di società* che, come collettività, vogliamo costruire per il futuro. Secondo alcuni studiosi, infatti, l'intelligenza artificiale potrebbe avere un impatto trasformativo sulla nostra società paragonabile a quello generato dalle rivoluzioni industriali, segnando una svolta profonda nei modelli economici, nei rapporti sociali e nelle strutture istituzionali¹. Di qui la necessità, per i decisori pubblici, di guidare lo sviluppo tecnologico e di bilanciare i benefici dell'innovazione con i rischi che si stagliano all'orizzonte; rischi che coinvolgono un'amplissima gamma di sfide cruciali per la nostra società.

Da un lato, infatti, l'intelligenza artificiale promette di offrire grandi vantaggi in termini di benessere sociale, potendo svolgere in maniera più rapida, e talvolta più efficace, attività che fino a qualche tempo fa potevano essere realizzate soltanto dagli esseri umani: si pensi, ad esempio, agli algoritmi di i.a. utilizzati in ambito medico-sanitario, che si sono già rivelati estremamente accurati nella formulazione di diagnosi; o alle macchine a guida autonoma, che, secondo alcune stime, se adottate da gran parte degli utenti della strada, potrebbero ridurre significativamente il numero di incidenti stradali – oltre che liberare uomini e donne da un'attività spesso percepita come noiosa e stressante.

Dall'altro lato, tuttavia, gli studiosi hanno da tempo puntato il riflettore

¹ V., in questo senso, S. ABIS-L. VELDKAMP, *The Changing Economics of Knowledge Production Get access Arrow*, in *The Review of Financial Studies*, 37, 1, 2024, p. 91 ss.; H. DEVLIN, *AI 'could be as transformative as Industrial Revolution'*, in *The Guardian* (online), 3 maggio 2023. Cfr. anche L. FLORIDI, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Raffaello Cortina, 2017, secondo cui le nuove tecnologie dell'informazione (tra le quali è inclusa anche l'intelligenza artificiale) si apprestano a realizzare una trasformazione della società paragonabile a quelle indotte dagli studi di Copernico, Darwin e Freud.

sul “lato oscuro” dell’i.a.². Inquinamento ambientale³; massiccio utilizzo di dati personali, con conseguente crescita delle capacità di sorveglianza delle *big tech*⁴; discriminazione algoritmica⁵; calo occupazionale⁶; approccio monopolistico dei *player* globali del settore, ormai talmente influenti da poter condizionare il funzionamento delle istituzioni democratiche⁷: sono soltanto alcuni dei *rischi* che stanno emergendo con sempre più evidenza dallo sviluppo e dalla diffusione dell’intelligenza artificiale.

A tale elenco esemplificativo si aggiunge un ulteriore elemento di criticità, relativo alla *difficoltà di controllare e di prevedere il “comportamento” e i risultati dell’intelligenza artificiale*. All’incremento delle capacità operative dei sistemi di i.a. corrisponde infatti, parallelamente, una crescita di impercetrabilità e di imprevedibilità dei dispositivi intelligenti, che spesso sfuggono alle capacità di comprensione e di controllo degli esseri umani – tanto dei produttori, quanto delle persone che utilizzano o sono chiamate a supervisionare gli algoritmi.

Ebbene, la *perdita di controllo dell’essere umano sui sistemi di i.a.* non può che avere ripercussioni anche sul piano dell’imputazione penale delle offese derivanti dall’intelligenza artificiale, che potrebbero mettere in discussione la tenuta dei classici meccanismi ascrittivi del diritto penale, strutturalmente fondati sul *mancato dominio*, da parte dell’agente, di *fatti offensivi effettivamente dominabili*⁸. Insomma, chi risponderà degli incidenti stradali derivanti dalle *self driving cars*, o delle errate diagnosi effettuate dai sistemi di supporto all’attività medico-sanitaria, o, ancora, delle

² K. CRAWFORD, *Né intelligente né artificiale. Il lato oscuro dell’IA*, Il Mulino, 2021.

³ Cfr. G. SISSA, *Le emissioni segrete. L’impatto ambientale dell’universo digitale*, Il Mulino, 2024. Si consideri, a tal proposito, che un solo ciclo di apprendimento di un sistema di intelligenza artificiale può arrivare a produrre più di 626.000 libbre di anidride carbonica (corrispondenti a 284 tonnellate), circa cinque volte le emissioni di un’automobile americana nel suo ciclo di vita; v. K. HAO, *Training a single AI model can emit as much carbon as five cars in their lifetimes*, in *MIT Technology Review*, 6 giugno 2019.

⁴ Cfr. S. ZUBOFF, *Il capitalismo della sorveglianza. Il futuro dell’umanità nell’era dei nuovi poteri*, Luiss University Press, 2023.

⁵ Cfr. S.U. NOBLE, *Algorithms of Oppression: How Search Engines Reinforce Racism*, NYU Press, 2018.

⁶ Cfr. D. SUSSKIND, *Un mondo senza lavoro: Come rispondere alla disoccupazione tecnologica*, Bompiani, 2022.

⁷ Cfr. O. POLLICINO-P. DUNN, *Intelligenza artificiale e democrazia. Opportunità e rischi di disinformazione e discriminazione*, Bocconi University Press, 2024.

⁸ Così A. FIORELLA, (voce) *Responsabilità penale*, in *Enc. dir.*, XXXIX, Giuffrè, 1988, p. 1289.

manipolazioni del mercato provocate dagli algoritmi di *trading*, che ormai dominano i mercati finanziari?

Le fattispecie astrattamente configurabili, in relazione alle offese derivanti dallo sviluppo o dall'utilizzo dell'intelligenza artificiale, sono le più varie, così come varie sono ormai le applicazioni dell'i.a., in ambito industriale, professionale e ricreativo. L'imprevedibilità dell'intelligenza artificiale, tuttavia, potrebbe determinare una *crisi delle categorie penalistiche*, sia sul terreno colposo, sia sul terreno doloso: laddove non c'è controllo sulle fonti del rischio, infatti, non può esservi nemmeno rimproverabilità per il risultato algoritmico lesivo o pericoloso. È quello che, nella letteratura di area giuridica e filosofica che si è occupata di intelligenza artificiale, è ormai comunemente noto come il problema del *responsibility gap*⁹ – che può tradursi come “il problema del vuoto di responsabilità”, determinato dall'impossibilità di individuare un soggetto al quale sia rimproverabile la commissione del fatto offensivo.

È su questo aspetto, relativo all'imputazione penale delle offese “da intelligenza artificiale”, che intendiamo soffermarci in questo lavoro. Ad una prospettiva più strettamente dogmatica, che si interroga sull'impatto dell'i.a. sulle categorie penalistiche, uniremo un approccio politico-criminale, volto a identificare possibili soluzioni – penalistiche ed extrapenalistiche – idonee a prevenire e a sanzionare i fatti offensivi emergenti dalla produzione e dall'utilizzo dei sistemi di i.a.

2. Sebbene il tema sia tutto sommato *nuovo* – tanto che la stessa tecnologia che sta alla base dell'i.a. è in divenire e non è ancora afferrabile compiutamente sul piano fenomenologico – la letteratura penalistica, italiana e straniera, è già di ampiezza e di approfondimento considerevole. Limitandoci soltanto alle monografie pubblicate in Italia, si segnalano già alcuni recenti lavori che affrontano il tema dell'imputazione della responsabilità penale in relazione a specifiche applicazioni dell'i.a., e, in particolare, nell'ambito della circolazione stradale¹⁰, della sanità¹¹ e della gestione

⁹ V., per tutti, A. MATTHIAS, *The responsibility gap: Ascribing responsibility for the actions of learning automata*, in *Ethics and Information Technology*, 6, 2004, p. 175; sul tema v. *infra*, cap. II, par. 1 ss.

¹⁰ M. LANZI, *Self-driving cars e responsabilità penale. La gestione del “rischio stradale” nell'era dell'intelligenza artificiale*, Giappichelli, 2023; R. COMPOSTELLA, *Auto a guida autonoma e diritto penale. Profili di responsabilità individuale e collettiva*, Editoriale Scientifica, 2024.

¹¹ L.A. TERRIZZI, *Medical devices e diritto penale. Profili di responsabilità del produttore e dell'utilizzatore*, Giuffrè, 2023; N. AMORE, *L'effetto della robotica e dell'IA nell'imputazione*

d'impresa¹². Si segnala, inoltre, un lavoro monografico in lingua inglese di un'Autrice italiana, che esplora i possibili profili di responsabilità penale diretta dei sistemi di i.a.¹³.

Manca, invece, nella letteratura italiana, una riflessione di parte generale sull'impatto dell'intelligenza artificiale sulle categorie e sui principi del diritto penale. È proprio questo vuoto che il nostro lavoro intende contribuire a colmare, nella convinzione che l'i.a. possa incidere sulla struttura del reato in maniera *trasversale* rispetto alle singole fattispecie di parte speciale, e che, di conseguenza, possa essere opportuna l'elaborazione di un quadro interpretativo in grado di orientare l'analisi dei più svariati fenomeni criminali che possono vedere coinvolta l'i.a.¹⁴: dalle lesioni cagionate da un *healthcare robot*, all'utilizzo di un sistema di i.a. per pubblicare sui *social network* commenti o post diffamatori, fino all'incidente stradale causato da un veicolo a guida autonoma.

Siamo consapevoli che, nel momento in cui scriviamo questo libro, sono ben poche le certezze circa l'effettivo funzionamento dell'intelligenza artificiale e circa l'impatto di tale tecnologia sulle capacità di controllo umane. Lungi dallo scoraggiarci dall'intraprendere questo studio, tale constatazione non ha fatto che rafforzare i nostri intenti. Quella che ci proponiamo di offrire al lettore è infatti una riflessione sul rapporto tra diritto penale e *incertezza*, tra diritto penale e *imprevedibilità* – in un contesto in cui, per la prima volta nella storia della tecnologia, l'imprevedibilità non è un *effetto collaterale* dell'evoluzione tecnologica, una deviazione rispetto al mito positivista della predicibilità, ma è *l'inevitabile portato della capacità di apprendimento* dei sistemi di i.a. (c.d. *machine learning*). Ebbene, a noi pare che, in questo momento di transizione, vi sia la necessità di *mettere in ordine i problemi*, tentando di inquadrarli nella griglia concettuale del diritto penale.

Il momento, in particolare, ci pare propizio anche per altre due ragioni.

giuridica degli eventi infausti, in E. ROSSERO-N. AMORE, *Robotica e intelligenza artificiale nell'attività medica. Organizzazione, autonomia, responsabilità*, il Mulino, 2023, p. 101 ss.

¹² A. MANGIONE, *Intelligenza artificiale, attività d'impresa e diritto penale. La «funzione di garanzia» nell'organizzazione e dell'organizzazione per la «sorveglianza dell'AI»*, Giappichelli, 2024.

¹³ A. GIANNINI, *Criminal behavior and Accountability of Artificial Intelligence Systems*, Eleven, 2023.

¹⁴ Segnaliamo, a tal proposito, che durante la fase di correzione delle bozze di questo lavoro è stata pubblicata una monografia che è mossa da obiettivi simili ai nostri, e che si propone di vagliare l'applicabilità dei meccanismi di imputazione della responsabilità penale nei casi in cui nella dinamica criminosa sia coinvolto un sistema di i.a.; v. L. D'AMICO, *La misura della (im)prevedibilità. Modelli d'imputazione della responsabilità al tempo dell'intelligenza artificiale*, ESI, 2025.

Da un lato, nonostante la casistica sulle offese derivanti dall'intelligenza artificiale sia già ampia e ben documentata, la giurisprudenza penale non ha ancora avuto occasione di pronunciarsi: ci muoviamo, dunque, su un terreno vergine, nell'auspicio che le nostre riflessioni possano risultare utili nel momento in cui tribunali e corti dovessero iniziare a confrontarsi con il fenomeno degli *AI-crimes*. Dall'altro lato, proprio mentre scriviamo, in Parlamento si discute una legge recante «disposizioni e delega al governo in materia di intelligenza artificiale»¹⁵, che interviene, tra l'altro, nella materia penale. Anche da questo punto di vista, l'auspicio è che il nostro lavoro possa fornire utili spunti per il dibattito in sede parlamentare e per l'attuazione delle deleghe da parte del Governo.

3. Dal punto di vista dogmatico, l'obiettivo del nostro lavoro, come già accennavamo, è quello di indagare l'impatto dell'intelligenza artificiale sulle categorie penalistiche e sull'imputazione della responsabilità.

Tale obiettivo ha imposto, innanzitutto, di individuare uno o più criteri tassonomici in grado di ordinare, in maniera chiara e coerente, i problemi giuridici da affrontare. Tra le varie opzioni prospettabili, la nostra scelta è ricaduta sull'utilizzo di due criteri: il primo basato sull'elemento soggettivo (*reati colposi vs reati dolosi*), il secondo basato sulla "posizione" dell'agente rispetto al "comportamento" del sistema di i.a. (*produttore vs supervisore-utilizzatore*). Dall'intersezione di questi due assi derivano quattro diverse aree d'indagine, dotate, al loro interno, di una certa omogeneità, non soltanto sul piano dogmatico, ma anche sul piano criminologico:

- (i) *responsabilità colposa del produttore* di sistemi di i.a. (ad es., per un omicidio colposo: si pensi al produttore di una *self driving car*, nel caso in cui quest'ultima, non riconoscendo un semaforo rosso, investa un pedone, cagionandone la morte);
- (ii) *responsabilità colposa dell'utilizzatore-supervisore* (ad es., per un omicidio colposo: un paziente muore a causa della terapia prescritta da un medico, che si era affidato all'indicazione terapeutica, poi rivelatasi errata, fornita da un sistema di i.a.);
- (iii) *responsabilità dolosa dell'utilizzatore* (ad es., per una manipolazione del mercato: un investitore attiva un sistema di *trading* algoritmico, che

¹⁵D.d.l. n. S. 1146 del 2024, recante «disposizioni e delega al governo in materia di intelligenza artificiale», presentato dal Presidente del Consiglio dei Ministri (Meloni) e dal Ministro della giustizia (Nordio), comunicato alla Presidenza del Senato il 20 maggio 2024, così come approvato dal Senato il 20 marzo 2025 e trasmesso alla Camera dei Deputati (n. 2316).

- effettua operazioni tali da determinare una sensibile alterazione dei prezzi degli strumenti finanziari);
- (iv) *responsabilità dolosa del produttore* (ad es., per la diffusione di codici *malware*: si pensi al fornitore di un *chatbot* come ChatGPT, nel caso in cui il *chatbot* fornisca ad un utente il testo di un codice attraverso il quale realizzare un cyberattacco).

Una simile ripartizione corrisponde, da un lato, ad una distinzione strutturale tra reati colposi e reati dolosi, che richiedono necessariamente una separata trattazione; dall'altro lato, tale suddivisione corrisponde a *quattro aree criminologiche ben distinte*, che, di conseguenza, possono dare luogo ad esigenze politico-criminali diverse. Per ciascuna area d'indagine, cercheremo di valutare le possibili forme di attrito tra imprevedibilità algoritmica e categorie penalistiche (causalità, colpa, dolo, posizione di garanzia, etc.), cercando, infine, di identificare gli ambiti in cui sia necessario un intervento normativo.

Anticipiamo fin da subito che, se per alcune aree – e, in particolare, per quel che riguarda il produttore –, emergerà l'esigenza di evitare forme di deresponsabilizzazione, in altri casi si risconterà l'esigenza opposta: è il caso, in particolare, della persona incaricata della supervisione dei sistemi di i.a., in relazione alla quale potrebbe rendersi necessaria una limitazione della responsabilità penale, al fine di evitare che le vengano imputati anche gli eventi lesivi derivanti da comportamenti incontrollabili e imprevedibili dell'i.a.

4. La seconda direttrice che ha guidato il nostro lavoro – oltre a quella, di cui abbiamo già dato conto, che concerne il tentativo di affrontare il problema *de qua* con un'impostazione di parte generale – consiste nell'idea di inserire la riflessione penalistica in *un orizzonte interdisciplinare*. Ciò, in particolare, è avvenuto in due diverse direzioni.

Da un lato, abbiamo tentato di comprendere, immergendoci nella letteratura scientifica (e di divulgazione scientifica), quali sono le caratteristiche tecniche che rendono l'intelligenza artificiale un fenomeno *dirompente* per il diritto penale. Si tratta di uno studio che – privi di un *background* in materie informatiche o ingegneristiche – abbiamo condotto con gli strumenti limitati e con le lenti del giurista, e che ha avuto come obiettivo quello di delineare le coordinate socio-tecnologiche di riferimento per comprendere l'impatto dell'intelligenza artificiale sulle capacità di controllo dell'essere umano.

Dall'altro lato, il nostro lavoro si caratterizza per una marcata attenzione ai profili di normazione e di disciplina *extrapenalistica*: ci soffermeremo, in

particolare, sulla regolazione relativa alla sicurezza dei sistemi di i.a. – contenuta, per quanto riguarda l’Unione europea, nel regolamento sull’intelligenza artificiale (c.d. AI Act) –, e sui profili di responsabilità extracontrattuale per i danni derivanti dai sistemi di i.a. La scelta, che potrebbe apparire inusuale nell’ambito di una monografia di diritto penale, è in realtà funzionale agli obiettivi della nostra indagine – che, lo accennavamo, ha un taglio, al contempo, dogmatico e politico-criminale.

Prima di tutto, il nostro obiettivo è quello di introdurre i penalisti ai punti nevralgici del dibattito, internazionale e italiano, sulla regolazione dell’intelligenza artificiale e sulla disciplina della responsabilità extracontrattuale per i danni da i.a. Già da diversi anni, infatti, la dottrina civilistica, e quella che si occupa di *law & technology*, si interrogano sulle modalità con le quali il mondo del diritto debba rispondere alle trasformazioni indotte dall’intelligenza artificiale. Tali riflessioni, a nostro avviso, costituiscono le premesse indispensabili per chi intenda indagare i rapporti tra diritto penale e intelligenza artificiale in una cornice teorica, e di “politica del diritto”, di più ampio respiro.

D’altra parte, l’esame della normativa e della disciplina extrapenalistica ha precise ricadute anche sul piano dogmatico e politico-criminale.

Quanto all’AI Act, l’analisi delle disposizioni contenute nel regolamento consentirà, innanzitutto, di delineare l’area di rischio lecito per lo sviluppo e la messa in commercio di sistemi di i.a., che ha evidenti implicazioni sul piano dell’accertamento della colpa del produttore, in caso di eventi lesivi scaturenti dall’i.a. In secondo luogo, l’esame del regolamento mira a far emergere l’assetto di *governance* europea sulla sicurezza dei sistemi di i.a., con particolare riferimento al ruolo attribuito alle *autorità di vigilanza nazionali*, che costituiranno il perno attorno al quale ruota l’intero sistema del controllo sulla sicurezza del prodotto.

Per altro verso, l’*excursus* sui profili di responsabilità extracontrattuale ci consentirà di valutare l’adeguatezza, di fronte ai nuovi rischi tecnologici, dei meccanismi ascrittivi della *responsabilità civile* – soprattutto alla luce della nuova direttiva europea sul danno da prodotto, che estende il tradizionale regime di responsabilità oggettiva della *product liability* anche ai danni derivanti dai sistemi di i.a.¹⁶ L’interesse penalistico di tale disamina si coglie, a nostro avviso, in un’ottica di complessiva funzionalità del sistema di prevenzione (e di compensazione) dei danni da i.a. Un rafforzamento della tutela compensativa offerta dalla responsabilità aquiliana, infatti, può avere effetti

¹⁶ Direttiva (UE) 2024/2853 del 23 ottobre 2024 sulla responsabilità per danno da prodotti difettosi, che abroga la direttiva 85/374/CEE del Consiglio. In argomento v. diffusamente *infra*, cap. V, par. 2 ss.

benefici anche sul diritto penale, contribuendo a ridimensionare le aspettative simboliche e soddisfattive che i consociati sempre più spesso nutrono, illusoriamente, nei confronti dello *ius terribile*.

Si tratta, a ben vedere, di una riflessione di carattere generale – che, tuttavia, nel settore dell'intelligenza artificiale può trovare un ambito privilegiato di sviluppo: sia perché, come già accennato, la strategia dell'Unione europea sembra quella di voler rafforzare, proprio in relazione ai danni da i.a., la funzione compensativa del diritto civile; sia perché l'imprevedibilità dell'intelligenza artificiale rischia di mettere in crisi i meccanismi ascrittivi della responsabilità penale, con la conseguenza che il diritto punitivo finirà, sempre più spesso, per rivelarsi inadeguato a soddisfare le aspettative “di giustizia” espresse dai consociati.

5. Già si intravede, in queste prime pagine introduttive, il taglio spiccatamente politico-criminale che impronta la nostra indagine. Se tale approccio attraverserà integralmente il lavoro – talvolta esplicitamente, in altri casi in maniera più implicita –, è soprattutto nel capitolo conclusivo che tireremo le somme circa le possibili soluzioni alla “questione criminale” legata alla produzione e all'utilizzo di sistemi di i.a. In quella sede, in particolare, cercheremo di individuare alcune possibili soluzioni alle esigenze politico-criminali emerse nel corso della trattazione – e lo faremo con uno sguardo che travalica la sola tutela penalistica, cercando di identificare, seppur soltanto nelle sue linee generali, un *sistema di tutela integrato*, tra diritto penale, diritto amministrativo, e diritto civile, volto a prevenire (e a compensare) le offese da intelligenza artificiale.

Tale prospettiva nasce dai risultati dell'indagine sui profili dogmatici svolta nella Parte III del nostro lavoro, che rivela come il diritto penale potrebbe rivelarsi inadatto – o comunque insufficiente – a svolgere una funzione general-preventiva rispetto ai produttori (e, in parte, anche nei confronti di coloro che utilizzano i sistemi di i.a. per compiere attività illecite). Il diritto penale è infatti uno strumento per sua natura *rigido*, inidoneo a contenere quei fenomeni sistemici, globalizzati ed incerti che caratterizzano la postmodernità, e di cui l'intelligenza artificiale costituisce l'ultima, dirompente, espressione. In questo contesto, l'alternativa che si presenta all'interprete e allo studioso è radicale: bisogna accettare un certo grado di *ineffettività del diritto penale*¹⁷ – inadatto ad attagliarsi a quei fenomeni imprevedibili ed imperscrutabili frutto del progresso

¹⁷ Imprescindibile il riferimento a C.E. PALIERO, *Il principio di effettività del diritto penale*, in *Riv. it. dir. proc.*, 1990, p. 430 ss.

tecnologico –, oppure è necessario che *anche il diritto penale sia oggetto di una trasformazione*, che lo renda “adeguato” rispetto alle nuove forme di rischio derivanti dall’evoluzione tecnologica?¹⁸. Insomma, quando c’è di mezzo l’i.a. *abbiamo necessariamente bisogno di trovare un colpevole*, anche al costo di *superare i principi di garanzia del diritto penale?*

Le domande, ovviamente, sono retoriche – ed è, d’altra parte, evidente che un certo grado di ineffettività del diritto penale costituisce una caratteristica *fisiologica* di un sistema punitivo fondato sui principi di legalità, colpevolezza e offensività.

In questo contesto, diventa dunque essenziale, nell’ambito della regolazione dell’intelligenza artificiale, volgere lo sguardo anche a possibili forme di intervento e di tutela *extrapenalistica*. Ciò non significa – può essere utile sottolinearlo fin da subito – che il diritto penale è destinato a non avere alcun ruolo della prevenzione delle offese da i.a.; dedicheremo, anzi, ampi sforzi a tentare di tracciare alcune possibili soluzioni penalistiche che possano attagliarsi al fenomeno dell’intelligenza artificiale. Significa, piuttosto, che il diritto penale è chiamato a operare all’interno di un più ampio sistema di prevenzione, del quale rappresenta solo uno degli strumenti possibili – lo strumento, oltretutto, *residuale*. Di qui l’ampio spazio dedicato, nella prima parte del lavoro, ai sistemi di regolazione extrapenale: senza un viaggio in quei sistemi, nel quale accompagneremo il lettore, la prospettiva penalistica, soprattutto in chiave politico-criminale, sarebbe non solo parziale, ma anche inevitabilmente viziata da grandi angoli ciechi.

6. Chiarite le direttrici di fondo della nostra indagine, ci pare ora opportuno delineare sinteticamente l’impianto complessivo del lavoro, che è articolato in quattro Parti: la prima dedicata alla *descrizione del funzionamento dell’intelligenza artificiale* e dei possibili fatti offensivi che possono derivarne; la seconda alla regolazione e alla disciplina *extrapenalistica*; la terza alla *responsabilità penale*; la quarta alle *conclusioni e alle prospettive sul rapporto tra i.a. e responsabilità penale*.

La *prima Parte* consta di tre capitoli.

Il *primo capitolo* illustra le principali fasi di sviluppo storico dell’intelligenza artificiale, nell’ottica di mettere a fuoco le ragioni per le quali gli

¹⁸ Cfr. L.A. ZAPATERO, *Introduzione*, in L. STORTONI-L. FOFFANI (a cura di), *Critica e giustificazione del diritto penale nel cambio di secolo. L’analisi critica della Scuola di Francoforte*, Giuffrè, 2004, p. 17: «Si pongono due opzioni di fondo: o scavare trincee contro il “moderno” diritto penale, o tentare di montare e cavalcare questo cavallo bizzoso dei moderni fenomeni materiali che premono sul sistema penale».

attuali modelli di *machine learning* realizzano un cambiamento di paradigma epocale nel rapporto tra essere umano e tecnologia, che ha ripercussioni anche sull'imputazione della responsabilità penale.

Il *secondo capitolo* si concentra invece sul “problema del controllo”, cercando di mettere in luce le principali caratteristiche del *machine learning* che incidono sulla capacità di controllo umano, e, di conseguenza, sulla possibilità di individuare una persona *responsabile* per le offese derivanti dall'intelligenza artificiale.

Il *terzo capitolo*, infine, si propone di fornire alcuni esempi concreti, tratti dalla casistica ad oggi esistente, di come i sistemi di i.a. possano realizzare *comportamenti offensivi di beni giuridici penalmente rilevanti*. Si tratta di una panoramica che non ha alcuna pretesa di esaustività, ma che intende offrire al lettore un'esemplificazione delle molteplici forme di offese che possono derivare dallo sviluppo e dall'utilizzo dei sistemi di i.a.

La *seconda Parte* segna il passaggio dalla *descrizione della tecnologia* alla sua *disciplina giuridica*, con un ampio sguardo interdisciplinare alla regolazione dell'intelligenza artificiale.

Il *quarto capitolo*, in particolare, è dedicato ad un'analisi dei primi tentativi di regolazione a livello internazionale, europeo, e domestico: un'attenzione specifica, come già evidenziato, sarà dedicata al regolamento europeo sull'intelligenza artificiale (AI Act), così come al d.d.l. n. S. 1146 del 2024 in materia di intelligenza artificiale, attualmente oggetto di discussione in Parlamento.

Nel *quinto capitolo*, invece, ci soffermeremo sui profili di responsabilità extracontrattuale per i danni da i.a., concentrandoci, in particolare, sui modelli di imputazione oggettiva, tra i quali spicca, come già si è detto, la nuova direttiva europea sul danno da prodotto¹⁹.

Con la *terza Parte* entreremo nel vivo della riflessione penalistica. Dopo un *sesto capitolo* volto a delimitare concettualmente e metodologicamente il perimetro dell'indagine, seguirà una Sezione dedicata ai reati colposi e una Sezione dedicata ai reati dolosi. Ciascuna delle due Sezioni contiene, a sua volta, due capitoli, uno dedicato ai profili di responsabilità del produttore, e l'altro ai profili di responsabilità dell'utilizzatore-supervisore. In breve, si succederanno quattro capitoli, ciascuno dedicato ad una delle *quattro aree di indagine* delle quali abbiamo già dato conto: ci occuperemo, innanzitutto, dei profili di responsabilità colposa del produttore di sistemi di i.a. (*settimo capitolo*) e della responsabilità colposa dell'utilizzatore o, comunque, della persona chiamata a supervisionare il sistema di i.a. (*ottavo capitolo*);

¹⁹Direttiva (UE) 2024/2853 del 23 ottobre 2024 sulla responsabilità per danno da prodotti difettosi, cit.