UNIVERSITÀ DEGLI STUDI DI MESSINA

DIPARTIMENTO DI GIURISPRUDENZA

CORSO DI DOTTORATO DI RICERCA IN "SCIENZE GIURIDICHE" – XXXV CICLO

Curriculum: *Tutela penale e garanzie della persona nel diritto interno, comparato, europeo ed internazionale: profili sostanziali e processuali.*

SSD: IUS/17

Scienza, precauzione e responsabilità penale. Premesse ad uno studio sull'intelligenza artificiale.

Dottoranda

Dott.ssa Laura D'Amico

Tutor

Chiar.ma Prof.ssa Lucia Risicato

Coordinatrice del Corso di Dottorato

Chiar.ma Prof.ssa Concetta Parrinello

ANNO ACCADEMICO 2021/2022

INDICE

CAPITOLO I

LA NORMATIVA SOVRANAZIONALE IN MATERIA DI INTELLIGENZA ARTIFICIALE E I SUOI POSSIBILI SVILUPPI

CAPITOLO II

LE ANTINOMIE TRA INTELLIGENZA ARTIFICIALE E CATEGORIE PENALISTICHE

CAPITOLO III

DALLA TEORIA ALLA PRATICA. ALCUNE CONCRETE IMPLICAZIONI PENALISTICHE DELL'INTELLIGENZA ARTIFICIALE

CAPITOLO I
LA NORMATIVA SOVRANAZIONALE IN MATERIA DI INTELLIGENZA ARTIFICIALE E I SUOI POSSIBILI SVILUPPI

SOMMARIO: SEZIONE PRIMA. DE IURE CONDITO – 1. Introduzione: la normativa esistente e gli scopi dell'indagine. 2. La Risoluzione del Parlamento europeo recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica. – 3. L'intelligenza artificiale per l'Europa. – 4. La Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi. – 5. Il Piano coordinato sull'intelligenza artificiale. – 6. La Risoluzione del Parlamento europeo su una politica industriale europea globale in materia di robotica e intelligenza artificiale. – 7. Gli Orientamenti etici per un'IA affidabile. – 8. Il Libro Bianco sull'intelligenza artificiale. – 8.1. La Relazione sulle implicazioni dell'intelligenza artificiale, dell'Internet delle cose e della robotica in materia di sicurezza e di responsabilità. – SEZIONE SECONDA. DE IURE CONDENDO – 9. Prospettive *de iure condendo*. – 9.1. La Proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'unione. – 9.2. La Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale. – 10. Conclusioni.

SEZIONE PRIMA
DE IURE CONDITO

1. *Introduzione: la normativa esistente e gli scopi dell'indagine.*

Il tema dell'intelligenza artificiale, allo stato attuale, si connota per una serie di incertezze e di dubbi interpretativi. L'assenza di un quadro normativo armonico e solido non aiuta certo il lavoro dello studioso che intenda approfondire questa materia. Nel presente capitolo ci si propone di prendere in esame alcuni dei più significativi interventi dell'Unione Europea sul punto e di estrapolare, ove possibile, spunti di riflessione dai connotati penalistici.

Dal punto di vista metodologico questo primo capitolo intende, da un lato, analizzare i testi selezionati, evidenziando regole e approcci di carattere generalizzante e, dall'altro, cogliere i possibili collegamenti con la materia penalistica.

Tale scelta strutturale origina dalla convinzione della difficoltà (se non forse proprio dell'impossibilità) per il giurista di potersi avvicinare allo studio di una data materia o di un determinato fenomeno senza avere una base di appoggio normativa, per quanto disarmonica possa essere. Ci si propone dunque in tal sede di tentare di ricondurre a sistema i seguenti interventi (perlopiù "paranormativi")[1] dell'Unione, non soltanto al fine di comprendere l'approccio che quest'ultima intende adottare nella regolamentazione della materia *de qua* ma anche al fine di cominciare a familiarizzare con concetti tecnici tipici dell'Intelligenza Artificiale (d'ora in avanti, anche IA).

[1] A. D'ALOIA, *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale,* in *BioLaw Journal*, 2019, p. 6.

2. La Risoluzione del Parlamento europeo recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica.

Nonostante la presente Risoluzione limiti il proprio campo di operatività esclusivamente alla materia civilistica, essa appare ai nostri occhi ugualmente meritevole di analisi, essendo ricca di interessanti riferimenti (anche) di carattere squisitamente penalistico che possono costituire, in prospettiva *de iure condendo*, spunti di riflessione e di opportuno approfondimento.

Il 16 febbraio 2017 il Parlamento europeo ha approvato una Risoluzione con la quale ha rivolto alla Commissione una serie di raccomandazioni concernenti norme di diritto civile sulla robotica[2].

Il suddetto Parlamento, in tal sede, prende in considerazione le potenzialità dell'intelligenza artificiale[3], ritenendo imprescindibile considerarne le relative implicazioni dal punto di vista legislativo ed etico, anche al fine di non ostacolare la ricerca scientifica in materia[4]. Il Parlamento europeo evidenzia altresì la necessità di approcciarsi al tema garantendo «la non discriminazione, il giusto processo, la trasparenza e la comprensibilità dei processi decisionali»[5] e preservando «la dignità, l'autonomia e l'autodeterminazione degli individui»[6].

Enunciando i principi generali che dovrebbero ispirare l'approccio dei soggetti a vario titolo coinvolti nel corso della vita di un sistema intelligente (segnatamente progettisti, fabbricanti ed utilizzatori) il Parlamento europeo richiama le leggi di Asimov, ai sensi delle quali:

1) un robot non può recar danno a un essere umano né può permettere che, a causa del proprio mancato intervento, un essere umano riceva danno;

2) un robot deve obbedire agli ordini impartiti dagli esseri umani, purché tali ordini non contravvengano alla Prima Legge;

3) un robot deve proteggere la propria esistenza, purché questa autodifesa non contrasti con la Prima o con la Seconda Legge;

---

[2] *Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*, (2015/2103(INL)), 16.2.2017. Tale Risoluzione è stata preceduta dal *Progetto di relazione recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*, 31.5.2016, Relatore: Mady Delvaux. Il testo si caratterizza per una commistione di elementi di natura giuridica ed etica che, invero, risulta spesso presente nella regolamentazione europea del settore della robotica, E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale: il dibattito, le proposte, le prospettive. Alcuni spunti di riflessione*, in *MediaLaws*, 10.3.2019, p. 82. Dedicano la propria attenzione alla presente Risoluzione anche F. CAROCCIA, *Soggettività giuridica dei robot?*, in G. ALPA, *Diritto e intelligenza artificiale*, Pisa, 2020, pp. 222 ss. e A. TURANO, *Robotica e roboetica: questioni e prospettive nazionali ed europee*, in G. ALPA, *Diritto e intelligenza artificiale*, cit., pp. 146 ss.

[3] Al Considerando E il Parlamento europeo ritiene l'intelligenza artificiale in grado di modificare le abitudini di vita e di lavoro dei consociati, aumentando il livello dei servizi, portando benefici in termini di risparmio economico ed efficienza, in particolare in ambito commerciale, sanitario e manifatturiero, come anche nel settore dei trasporti, dell'istruzione e dell'agricoltura. Più avanti, al Considerando G, il Parlamento prende atto della tendenza attuale a sviluppare macchine autonome in grado di apprendere ed assumere decisioni indipendentemente, evidenziando i potenziali vantaggi da ciò derivanti, uniti alle preoccupazioni ad essa connessi.

[4] *Risoluzione del Parlamento europeo*, cit., Considerando B.

[5] *Risoluzione del Parlamento europeo*, cit., Considerando H.

[6] *Risoluzione del Parlamento europeo*, cit., Considerando O.

0) un robot non può recare danno all'umanità, né può permettere che, a causa del proprio mancato intervento, l'umanità riceva danno[7].

Ferme restando le linee guida generali ed astratte poste dalle succitate leggi, il Parlamento evidenzia la necessità di stabilire una serie di norme concernenti la trasparenza e l'assunzione di responsabilità che siano, al contempo, in grado di riflettere i valori europei e di non influenzare (negativamente) la ricerca nel settore della robotica[8].

Una simile considerazione appare evocativa del principio di precauzione, ben noto alla materia penalistica e, in particolare, all'ambito della ricerca scientifica[9].

Tale principio, negli anni 2000, è stato fatto oggetto di una specifica comunicazione della Commissione[10] al fine di delineare una proficua strategia europea nel fare ricorso a siffatto principio.

In tal sede la Commissione chiarisce come la scelta di ricorrere o meno al principio di precauzione avvenga, generalmente, in contesti in cui le informazioni scientifiche a disposizione siano insufficienti e sussistano fondati motivi per ritenere che i potenziali effetti di una determinata azione possano incidere negativamente sulla salute umana o sull'ambiente, in modo incompatibile con il prescelto livello di protezione.

Posto che un immotivato ricorso al principio di precauzione porterebbe con sé il rischio di un protezionismo dissimulato, la materia *de qua* impone la costante ricerca di un (non semplice) equilibrio tra i diritti delle industrie e quelli degli individui. Per far ciò occorre muoversi attraverso le tre fasi della gestione del rischio:

- valutazione del rischio;
- scelta della strategia più adeguata per gestire il rischio medesimo;
- comunicazione del suddetto rischio[11].

---

[7] *Risoluzione del Parlamento europeo*, cit., Considerando T. Parte della dottrina non ha accolto favorevolmente il riferimento alle leggi di Asimov ritenendo che esse esauriscano la loro portata nella letteratura fantascientifica senza aggiungere alcunché, in punto di giuridicità, al dibattito in questione, così A. ZORNOZA, M. LAUKYTE, *Robotica e diritto: riflessioni critiche sull'ultima iniziativa di regolamentazione in Europa*, in *Contr. Impr. Eur.*, 2/2016, p. 814. A sostegno della loro tesi gli AA. richiamano S.L. ANDERSON, *The Unacceptability of Asimov's Three Laws of Robotics as a Basis for Machine Ethics*, in M. ANDERSON, S.L. ANDERSON (a cura di) *Machine Ethics*, Cambridge, 2011, pp. 285 ss. Della stessa A. v. anche S.L. ANDERSON, *Asimov's "Three Laws of Robotics" and Machine Metaethics*, in *AI & Society*, 22/2008, 10.3.2007, pp. 477-493.

[8] *Risoluzione del Parlamento europeo*, cit., Considerando U. Tale principio è ribadito anche al Considerando X ove si propone un approccio graduale alle future iniziative in materia di robotica ed intelligenza artificiale, proprio al fine di non soffocare la relativa innovazione.

[9] Il principio di precauzione è normativamente richiamato dall'art. 174 della Versione consolidata del Trattato che istituisce la Comunità Europea (Gazzetta ufficiale delle Comunità europee, C 325/33, 24.12.2002, art. 174 comma 2) il quale, riferendosi in particolare alle politiche ambientali, stabilisce come l'azione europea in tale ambito debba mirare ad un elevato livello di tutela, fondandosi sui principi di precauzione e di azione preventiva. Sul rapporto tra principio di precauzione e diritto penale ambientale I. SALVEMME, *Il ruolo del principio di precauzione nel "nuovo" diritto penale dell'ambiente*, in *dirittopenalecontemporaneo.it*, 25.5.2018; C. RUGA RIVA, *Dolo e colpa nei reati ambientali. Considerazioni su precauzione, dolo eventuale ed errore*, in *dirittopenalecontemporaneo.it*, 19.1.2015; L. TROYER, *I nuovi reati ambientali "abusivi": quando la rinuncia alla legalità penale diviene un illusorio instrumentum regni*, in *Criminalia*, 2015, pp. 332 ss.

[10] *Comunicazione della Commissione sul principio di precauzione*, Bruxelles, 2.2.2000.

[11] *Comunicazione della Commissione sul principio di precauzione*, cit., pp. 7-8.

Lo studio del principio di precauzione mette in correlazione la scelta politica di agire o meno (in considerazione dei fattori che connotano il rischio) ed, eventualmente, la determinazione delle misure da adottare, quindi del "come" agire[12].

Al fine di ricorrere al principio di precauzione (il quale si colloca inevitabilmente in un contesto di incertezza scientifica)[13] è necessario, *in primis*, che sia stata effettuata una valutazione dei dati scientifici relativi ai rischi – con conseguente identificazione dei potenziali effetti negativi derivanti da un dato fenomeno – ed, *in secundis*, che sia stata svolta (sulla base delle conoscenze scientifiche disponibili in quel dato momento storico) una valutazione scientifica in ordine a tali potenziali effetti negativi[14].

La Commissione delinea nel prosieguo le Linee direttrici che dovrebbero ispirare il ricorso al principio di precauzione, chiarendo come l'attuazione di una strategia fondata su questo principio dovrebbe basarsi su una valutazione scientifica quanto più completa possibile, coinvolgendo tutte le parti in causa per vagliare le possibili opzioni di gestione del rischio e prendendo in considerazione le conseguenze di un'eventuale inazione, concentrandosi da ultimo sui principi ispiratori della precauzione e sul relativo onere della prova[15].

Secondo la dottrina penalistica il principio di precauzione[16] trova cittadinanza in tutti quei settori scientifici in cui le conoscenze disponibili in un dato momento

---

[12] *Comunicazione della Commissione sul principio di precauzione*, cit., pp. 12-13. In tal sede la Commissione raccomanda di non sovrapporre il principio di precauzione alla c.d. strategia di prudenza: «La *strategia di prudenza* è iscritta nella politica di *valutazione dei rischi* che è determinata prima di qualunque valutazione dei rischi stessi (…) Essa fa quindi parte integralmente del parere scientifico espresso da coloro che valutano il rischio. L'applicazione del *principio di precauzione* appartiene, invece, alla *gestione del rischio*, quando l'incertezza scientifica non consente una valutazione completa di tale rischio e i responsabili ritengono che il livello prescelto di protezione dell'ambiente o della salute umana, animale o vegetale possa essere minacciato», corsivo nostro.

[13] «L'incertezza scientifica deriva di solito da cinque caratteristiche del metodo scientifico: le variabili prescelte, le misurazioni effettuate, i campioni individuati, i modelli utilizzati e le relazioni causali impiegate. L'incertezza scientifica può derivare inoltre da controversie sui dati esistenti o dalla mancanza di dati. L'incertezza può riguardare elementi qualitativi o quantitativi dell'analisi» *Comunicazione della Commissione sul principio di precauzione*, cit., p. 14.

[14] *Comunicazione della Commissione sul principio di precauzione*, cit., p. 13. La valutazione del rischio consta di quattro componenti: l'identificazione del pericolo, la caratterizzazione del pericolo, la valutazione dell'esposizione e la caratterizzazione del rischio, concetti approfonditi nell'Allegato III della Comunicazione in esame, p. 29.

[15] I principi cui ci si riferisce sono, nello specifico: la proporzionalità; la non discriminazione; la coerenza; l'esame dei vantaggi e degli oneri derivanti dall'azione o dalla mancanza di azione; l'esame dell'evoluzione scientifica. *Comunicazione della Commissione sul principio di precauzione*, cit., pp. 18 ss.

[16] Si veda in generale sul tema R. TITOMANLIO, *Il principio di precauzione fra ordinamento europeo e ordinamento italiano*, Torino, 2018; E. CORN, *Il principio di precauzione nel diritto penale. Studio sui limiti all'anticipazione della tutela penale*, Torino, 2013; F. CONSORTE, *Tutela penale e principio di precauzione. Profili attuali, problematicità, possibili sviluppi*, Torino, 2013; D. CASTRONUOVO, *Principio di precauzione e diritto penale. Paradigmi dell'incertezza nella struttura del reato*, Roma, 2012; ID., *Principio di precauzione e beni legati alla sicurezza. La logica precauzionale come fattore espansivo del "penale" nella giurisprudenza della Cassazione*, in *dirittopenalecontemporaneo.it*, 21.7.2011; C. BRUSCO, *Rischio e pericolo, rischio consentito e principio di precauzione. La c.d. "flessibilizzazione delle categorie del reato"*, in *Criminalia*, 2012, pp. 383 ss.; A. MASSARO, *Principio di precauzione e diritto penale:* nihil novi sub sole*? Funzioni e limiti del principio di precauzione de iure condito e condendo*, in

storico non appaiono sufficienti al fine di poter determinare con certezza la pericolosità o la dannosità di un determinato fenomeno, potendone da ciò scaturire nuovi ed imponderabili rischi.

Ci si chiede pertanto, in una situazione di incertezza scientifica e di dubbio circa i danni o i pericoli verificabili, entro che limiti ed al ricorrere di quali presupposti sia possibile imporre coattivamente misure preventive in grado di limitare il progresso della ricerca scientifica.

Secondo autorevole dottrina determinate forme di tutela anticipata possono considerarsi legittime al ricorrere dei medesimi presupposti giustificativi delle fattispecie di pericolo astratto.

In primo luogo, pertanto, deve trattarsi di rischi che, per quanto non siano riscontrabili con assoluta certezza, non siano mero frutto di congetture, dovendo pur sempre poggiare su seri fondamenti.

In secondo luogo, deve sussistere un rapporto di proporzione tra i beni giuridici da proteggere ed i potenziali costi – segnatamente in termini di limitazione della ricerca scientifica – che deriverebbero dall'anticipazione della tutela conseguente all'applicazione del principio di precauzione[17].

Fatta questa corposa – quanto necessaria – incidentale sul principio di precauzione (il quale trova piena cittadinanza in materia di intelligenza artificiale al punto da essere oggetto di attenzione anche nel prosieguo della trattazione) pare ora opportuno tornare sulla Risoluzione in esame, ove il Parlamento europeo definisce come "essenziale" la disamina della questione concernente la responsabilità giuridica scaturente dall'azione dannosa di un'entità dotata di intelligenza artificiale[18].

A tal riguardo, dopo aver fornito una definizione del concetto di "autonomia di un robot"[19], il Parlamento evidenzia come ad una maggiore autonomia di tali macchine intelligenti corrisponda una minore possibilità di considerarle come meri strumenti nelle mani dell'uomo, facendo da ciò discendere un naturale interrogativo, ossia se le attuali norme in materia di attribuzione della responsabilità siano sufficienti oppure se vi sia la necessità di stabilire nuove regole per individuare il responsabile di un'azione (od omissione) dannosa posta

---

dirittopenalecontemporaneo.it, 9.5.2011; C. RUGA RIVA, *Principio di precauzione e diritto penale. Genesi e contenuto della colpa in contesti di incertezza scientifica*, in E. DOLCINI, C.E. PALIERO (a cura di), *Studi in onore di Giorgio Marinucci*, Milano, 2006, pp. 1743 ss.; G. FORTI, *"Accesso" alle informazioni sul rischio e responsabilità: una lettura del principio di precauzione*, in *Criminalia*, 2006, pp. 155 ss.; F. GIUNTA, *Il diritto penale e le suggestioni del principio di precauzione*, in *Criminalia*, 2006, pp. 227 ss. Specificamente in ambito medico si veda B. BERTARINI, *Tutela della salute, principio di precauzione e mercato del medicinale. Profili di regolazione giuridica europea e nazionale*, Torino, 2016. In materia di sicurezza alimentare si veda L. TUMMINELLO, *Sicurezza alimentare e diritto penale: vecchi e nuovi paradigmi tra prevenzione e precauzione*, in *dirittopenalecontemporaneo.it*, 15.10.2013. Nell'ambito della sicurezza sul lavoro v. I. SCORDAMAGLIA, *Il diritto penale della sicurezza del lavoro tra i principi di prevenzione e di precauzione*, in *dirittopenalecontemporaneo.it*, 23.11.2012.

[17] G. FIANDACA, E. MUSCO, *Diritto Penale. Parte generale*, Bologna, 2019, pp. 222-223.

[18] *Risoluzione del Parlamento europeo*, cit., Considerando Z.

[19] *Risoluzione del Parlamento europeo*, cit., Considerando AA: «considerando che l'autonomia di un robot può essere definita come la capacità di prendere decisioni e metterle in atto nel mondo esterno, indipendentemente da un controllo o un'influenza esterna; che tale autonomia è di natura puramente tecnologica e il suo livello dipende dal grado di complessità con cui è stata progettata l'interazione di un robot con l'ambiente».

in essere da un sistema autonomo e non riconducibile ad un soggetto umano specificatamente individuabile[20].

Anticipando un concetto che sarà approfondito nel prosieguo, posto che attualmente i succitati sistemi non possono essere ritenuti responsabili[21] per una condotta omissiva o commissiva recante danni a terzi e che le norme ad oggi esistenti riguardano l'ipotesi in cui l'agire dell'entità dotata di intelligenza artificiale possa comunque farsi risalire ad uno specifico agente umano, potrebbe da ciò derivare una forma di responsabilità oggettiva per i comportamenti realizzati da un robot in capo al suo utilizzatore, programmatore o proprietario[22], con il conseguente rischio che si vada giocoforza consolidando una sorta di responsabilità per posizione di tali soggetti[23].

Il Parlamento europeo in tal sede giunge a prendere in considerazione l'ipotesi in cui un sistema intelligente assuma decisioni autonomamente, evidenziando come in siffatta circostanza le norme attualmente vigenti non sarebbero sufficienti a fronteggiare la responsabilità per danni cagionati dall'intelligenza artificiale in quanto non sarebbe individuabile il soggetto su cui incomberebbe l'obbligo del risarcimento[24]. L'insufficienza dell'attuale assetto normativo dipenderebbe anche dall'evolversi delle capacità robotiche e dal

---

[20] *Risoluzione del Parlamento europeo*, cit., Considerando AB. Il successivo Considerando sottopone all'attenzione del lettore un ulteriore problematica, ossia che «l'autonomia dei robot solleva la questione della loro natura alla luce delle *categorie giuridiche esistenti* e dell'eventuale necessità di creare una *nuova categoria con caratteristiche specifiche* e implicazioni proprie», corsivo nostro.

[21] I. SALVADORI, *Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale*, in *Riv. it. dir. proc. pen*., 2021, p. 97. «È sempre necessario che qualcuno, o qualcosa, sia responsabile per i danni arrecati (…) E sulla base degli attuali sistemi di responsabilità, quel qualcuno è il proprietario, o, se del caso, il programmatore» A. ZORNOZA, M. LAUKYTE, *Robotica e diritto*, cit., pp. 811-812.

[22] *Risoluzione del Parlamento europeo*, cit., Considerando AD. Nel Considerando successivo il Parlamento europeo evidenzia come le norme in materia di responsabilità da prodotto e di responsabilità per azioni dannose siano applicabili ai danni cagionati dall'intelligenza artificiale. Segnatamente ci si riferisce alla responsabilità del produttore in caso di malfunzionamento del prodotto ed alla responsabilità dell'utente nel caso di un comportamento che conduca al danno.

[23] Anche su questo punto si è registrata una duplicità di vedute. Da un lato vi è chi ritiene che il livello di autonomia potenzialmente raggiungibile dai sistemi autonomi sia tale da rendere iniquo attribuire all'uomo dietro la macchina la responsabilità per i danni cagionati da quest'ultima. Ritenere automaticamente responsabili gli sviluppatori di questi sistemi intelligenti per i danni da questi causati potrebbe, da un lato, disincentivare gli investimenti nel settore e, dall'altro, ostacolare la ricerca tecnologica: investitori e ricercatori sarebbero infatti demotivati dall'approcciarsi a tale ambito a fronte del potenziale rischio di incorrere in forme di responsabilità oggettiva per danni causati da un sistema talmente autonomo da non essere controllabile. Di tale avviso A. ZORNOZA, M. LAUKYTE, *Robotica e diritto*, cit., p. 811. Di contro vi è chi ritiene che, per quanto tali sistemi possano essere autonomi, la responsabilità per i danni da loro causati non potrà che essere attribuita ai progettisti e ai costruttori che li hanno dotati di quelle specifiche peculiarità. Tale ultima tesi è stata sostenuta da F. RODI, *Gli interventi dell'Unione Europea in materia di intelligenza artificiale e robotica: problemi e prospettive*, in G. ALPA, *Diritto e intelligenza artificiale*, cit., p. 194.

[24] *Risoluzione del Parlamento europeo*, cit., Considerando AF. Nel Considerando AH viene sottolineato come «per quanto riguarda la responsabilità extracontrattuale, la direttiva 85/374/CEE riguarda solamente i danni causati dai difetti di fabbricazione di un robot e a condizione che la persona danneggiata sia in grado di dimostrare il danno effettivo, il difetto nel prodotto e il nesso di causalità tra difetto e danno e che pertanto la responsabilità oggettiva o la responsabilità senza colpa potrebbero non essere sufficienti».

conseguente grado di imprevedibilità connesso alla loro capacità di autoapprendimento ed autoadattamento la quale, inevitabilmente, finirebbe per sfuggire al controllo umano[25].

Pur avendo evidenziato i punti critici della materia, l'Unione europea predilige un approccio propositivo alla materia, ponendo l'attenzione sull'importanza di fornire definizioni comuni dei sistemi autonomi in considerazione, proponendo l'introduzione di un sistema globale per la registrazione dei robot maggiormente avanzati e chiarendo, altresì, come lo sviluppo di simili tecnologie debba avere come obiettivo l'integrazione delle capacità umane e non la loro sostituzione, essendo sempre necessario che gli uomini mantengano il controllo di tali macchine intelligenti[26].

Nel prosieguo della Risoluzione si evidenzia la necessità di dar vita ad un quadro etico da osservare durante le fasi di sviluppo, progettazione, produzione, uso e modifica dei robot, che sia altresì fondato sui principi di autonomia, giustizia, dignità, eguaglianza, non discriminazione, consenso informato, rispetto della vita privata, protezione dei dati personali e trasparenza. In particolare, secondo tale ultimo principio, dovrebbe essere sempre possibile risalire alla logica sottesa ad ogni decisione assunta dall'intelligenza artificiale, ricostruendone i passaggi[27].

Adottando un approccio maggiormente concreto, la Risoluzione si sofferma poi sugli aspetti relativi alla materia della proprietà intellettuale, del rispetto alla vita privata e della protezione dei dati, vagliando – seppur sinteticamente – le potenzialità dell'intelligenza artificiale in settori di particolare interesse quali veicoli autonomi[28], droni[29] e robot medici[30].

Il Parlamento europeo si concentra, in ultima analisi, sulla materia della responsabilità civile per i danni causati dai robot, svolgendo però una serie di considerazioni da reputarsi valevoli anche in ambito penalistico e che, in prospettiva *de iure condendo*, sarebbe opportuno tenere presenti.

---

[25] *Risoluzione del Parlamento europeo*, cit., Considerando AI.

[26] *Risoluzione del Parlamento europeo*, cit., Punti 1, 2 e 3 i quali aprono la Sezione relativa ai "Principi generali riguardanti lo sviluppo della robotica e dell'intelligenza artificiale per uso civile".

[27] *Risoluzione del Parlamento europeo*, cit., Punti 11, 12 e 13.

[28] *Risoluzione del Parlamento europeo*, cit., Punti 24 ss. A tal proposito il Parlamento europeo chiarisce come all'interno della nozione di trasporto autonomo debbano annoverarsi tutte le forme di pilotaggio automatizzate (avvengano esse su strada, ferrovie, spazio aereo o marittimo), evidenziando altresì come un approccio normativo frammentario, specie nel settore automobilistico, rischierebbe di ostacolare la realizzazione di tali sistemi di trasporto.

[29] *Risoluzione del Parlamento europeo*, cit., Punti 30 ss. Relativamente a tale materia il Parlamento europeo esorta la Commissione a vagliare la possibilità di introdurre un meccanismo di tracciabilità e identificazione obbligatorio per i sistemi aerei a pilotaggio remoto (RPAS), anche noti come veicoli aerei senza equipaggio (UAV).

[30] *Risoluzione del Parlamento europeo*, cit., Punti 33 ss. In ordine a tale aspetto la Risoluzione evidenzia l'importanza di una formazione adeguata da parte del personale sanitario, altresì sancendo il principio di "autonomia supervisionata dei robot" secondo il quale la scelta finale in ordine alle pratiche sanitarie spetterebbe sempre al medico umano. L'impiego dell'intelligenza artificiale in medicina può tornare utile non soltanto nel campo dell'autodiagnosi, mediante l'ausilio di robot mobili, ma anche durante la fase di assistenza e cura, così riducendo il margine di errore dell'uomo, garantendo una maggiore personalizzazione del trattamento sanitario ed aumentando conseguentemente la qualità della vita, grazie anche ad esempio agli sviluppi nel campo delle protesi robotiche.

In tal sede il Parlamento considera imprescindibile l'esigenza di affrontare la materia in modo armonico al fine di garantire un medesimo livello di certezza giuridica in tutta l'Unione[31], chiarendo come lo strumento legislativo che verrà prescelto per far fronte a questa innovativa tipologia di responsabilità non dovrebbe limitare la tipologia o l'ammontare dei danni risarcibili alla persona offesa sol perché il danno non è stato cagionato da un soggetto umano[32].

Il Parlamento europeo propone poi alla Commissione di operare una scelta di fondo e decidere se strutturare tale strumento normativo in termini di *responsabilità oggettiva* – la quale richiede la prova del danno e del nesso di causalità intercorrente tra quest'ultimo e il funzionamento del robot – oppure secondo l'approccio della *gestione dei rischi*, il quale si concentra sulla responsabilità del soggetto in grado di minimizzare i rischi derivanti dall'agire dell'intelligenza artificiale contrastandone l'impatto lesivo[33].

In conclusione, ricollegandosi alle precedenti considerazioni svolte sul punto, la Risoluzione ritiene necessario che la responsabilità[34] di coloro ai quali sia possibile imputare l'evento dannoso vada proporzionata non soltanto a livello di indicazioni impartite al robot ma anche al suo grado di autonomia[35].

---

[31] *Risoluzione del Parlamento europeo*, cit., Punto 49.

[32] *Risoluzione del Parlamento europeo*, cit., Punto 52.

[33] *Risoluzione del Parlamento europeo*, cit., Punto 53, 54 e 55.

[34] Al Punto 56 la Risoluzione specifica che, conseguentemente, quanto maggiore sarà la durata della formazione di un robot, tanto maggiore dovrebbe essere la responsabilità del suo formatore. Nel prosieguo il Parlamento europeo chiarisce che «nella determinazione della responsabilità reale per il danno causato, le competenze derivanti dalla "formazione" di un robot non dovrebbero essere confuse con le competenze che dipendono strettamente dalle sue abilità di autoapprendimento», osservando inoltre che, allo stato attuale, la responsabilità per l'agire del robot non può che essere imputata all'uomo. Ai Punti 57 e 58 la Risoluzione paventa altresì la possibilità di istituire un regime di assicurazione obbligatoria per risolvere il problema dell'attribuzione della responsabilità per il danno cagionato da forme di intelligenza artificiale sempre più autonome, nonché di uno specifico fondo per garantire la possibilità di risarcire i danni nel caso in cui manchi la copertura assicurativa. «La creazione di un'assicurazione obbligatoria può essere certamente utile per consentire di affrontare, almeno temporaneamente, gli aspetti connessi alla responsabilità dal punto di vista economico» A. ZORNOZA, M. LAUKYTE, *Robotica e diritto*, cit., pp. 812-813. Al Punto 59 (lett. c, e ed f) il Parlamento europeo, infine, invita la Commissione a valutare la possibilità che gli addetti ai lavori (produttori, programmatori, proprietari, utilizzatori ecc.) possano usufruire di una responsabilità limitata qualora sottoscrivano la succitata assicurazione e costituiscano un fondo per i risarcimenti; la realizzazione di un numero di immatricolazione in modo da associare ogni robot al relativo fondo; la possibilità di istituire uno status giuridico, specie per le forme di intelligenza artificiale più sofisticate, in modo da poterli considerare come persone elettroniche cui imputare il risarcimento per i danni da loro causati; l'istituzione di una personalità elettronica (status giuridico) per quei robot talmente tanto sviluppati da essere in grado di interagire autonomamente con altre persone, fino a prevederne una responsabilità a titolo risarcitorio per i danni da questi ultimi cagionati. Sul tema del riconoscimento di una specifica personalità giuridica in capo ai robot si vedano M. LAUKYTE, *Artificial agents among us: Should we recognize them as agents proper?*, in *Ethics and Information Technology*, 19/2017, pp. 1-17; L.B. SOLUM, *Legal Personhood for Artificial Intelligences*, in *North Carolina Law Review*, 70, 1992, pp. 1231-1287. Ad avviso di parte della dottrina l'introduzione di una personalità elettronica deriverebbe principalmente da un'istanza di rassicurazione dei cittadini, ravvisabile nella garanzia del risarcimento dei danni cagionati dal sistema intelligente. A. ZORNOZA, M. LAUKYTE, *Robotica e diritto*, cit., pp. 811-812. Il tema verrà approfondito al Cap. II, Sez. III, Par. 17.3.

[35] L'appena esaminata Risoluzione è accompagnata da un Allegato di particolare interesse del quale pare opportuno prendere brevemente in esame i punti nodali. Il Parlamento europeo propone

Come è possibile evincere dalla disamina appena svolta, numerosi sono i riferimenti valevoli anche in ambito penale contenuti nella presente Risoluzione, la quale merita attenzione – anche al di fuori del campo civilistico – non soltanto in quanto rappresenta uno dei primi approcci dell'Unione Europea alla regolamentazione di un fenomeno nuovo e delicato come l'intelligenza artificiale, ma anche in ragione del fatto che le considerazioni in essa contenute hanno portata generalizzante, rappresentando un prezioso apripista per cominciare a riflettere sull'argomento[36].

3. L'intelligenza artificiale per l'Europa.

L'impegno europeo in materia prosegue con l'emanazione, il 25 aprile 2018, della Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni dal titolo "L'intelligenza artificiale per l'Europa".

---

di fornire una definizione comune di "robot autonomo intelligente" che tenga conto di alcuni requisiti: «la capacità di acquisire autonomia grazie a sensori e/o mediante lo scambio di dati con il proprio ambiente (interconnettività) e l'analisi di tali dati; la capacità di apprendimento attraverso l'esperienza e l'interazione; la forma del supporto fisico del robot; la capacità di adeguare il suo comportamento e le sue azioni all'ambiente».
In tal sede il Parlamento individua i principi cui dovrebbero ispirarsi i ricercatori che si occupano di robotica (segnatamente beneficenza, non malvagità, autonomia e giustizia) e stabilisce che la ricerca effettuata in tale ambito debba svolgersi nel rispetto dei diritti fondamentali, del principio di precauzione, della trasparenza, dell'inclusione, della rendicontabilità, della sicurezza, della reversibilità e del diritto alla privacy.
Nel corpo dell'Allegato in esame vengono stilati il codice etico-deontologico degli ingegneri robotici, il codice per i comitati etici di ricerca (CER), la licenza per progettisti e la licenza per gli utenti. Le indicazioni ivi contenute sembrano a tutti gli effetti costituire adempimenti preventivi da parte dei soggetti coinvolti che, ove venissero definitivamente normativizzati, potrebbero assurgere al rango di fonte giuridica il cui mancato rispetto o la cui inosservanza potrebbe comportare l'applicazione della responsabilità omissiva *ex* art. 40 cpv. c.p., anche secondo una declinazione colposa nel caso in cui l'omesso adempimento di uno di questi obblighi si fosse verificato per negligenza (la quale, per l'appunto, ricorre nel caso in cui la regola di condotta violata prescriva il compimento di un'attività positiva). Attira l'attenzione del penalista un inciso contenuto nella sezione concernente gli utenti dei sistemi intelligenti, ove viene stabilito che essi hanno diritto di attendersi dal robot che questo svolga i compiti per i quali è stato creato, non essendo autorizzati ad utilizzarlo in modo contrario alle norme etiche e giuridiche. Il possibile uso distorto di un sistema intelligente non deve, difatti, costituire un limite alla ricerca scientifica e tecnologica: «Ad esempio, i droni possono essere utilizzati per attività di topografia aerea e per questo scopo sono dotati di telecamere. Se un individuo utilizza un drone per spiare i suoi vicini, la sua attività determina una violazione della privacy, ma questo non è un valido motivo per limitare l'innovazione, impedendo che i droni siano dotati di telecamere» A. ZORNOZA, M. LAUKYTE, *Robotica e diritto*, cit., p. 813. Si comincia in tale prospettiva ad avere per la prima volta contezza della possibilità che lo strumento dotato di intelligenza artificiale venga utilizzato per scopi illeciti, costituendo mero strumento nelle mani dell'uomo al fine di delinquere.

[36] Critica sul punto parte della dottrina ad avviso della quale, iniziando a intervenire normativamente sul punto, si andrebbero ad «anticipare le eventuali conseguenze giuridiche che potrebbero emergere dall'interazione con i (e l'uso dei) robot intelligenti ed autonomi» A. ZORNOZA, M. LAUKYTE, *Robotica e diritto*, cit., p. 808. Approccio di maggiore apertura ha, invece, altra parte della dottrina la quale, riferendosi al Rapporto Delvaux che precede l'entrata in vigore della Risoluzione in esame, ritiene che esso si limiti «a fissare principi, a compiere valutazioni, a motivare la necessità di un intervento in materia di IA, senza però offrire e predisporre definitive soluzioni; esse saranno elaborate solo nel prossimo futuro», così F. RODI, *Gli interventi dell'Unione Europea in materia di intelligenza artificiale e robotica*, cit., p. 193.

La Comunicazione che ci si accinge ad esaminare non spicca per i riferimenti penalistici ma riteniamo ugualmente utile svolgerne una sintetica analisi in quanto tale testo ha il pregio di tracciare le linee programmatiche del lavoro da svolgere nella materia dell'intelligenza artificiale, fornendo una serie di input che ritroveremo nelle pagine successive.

Prendendo atto della necessità di un approccio coordinato in materia, la Commissione chiarisce che un'iniziativa europea in tema di intelligenza artificiale[37] dovrebbe essere volta a:

- dare impulso all'uso di siffatte tecnologie in ogni settore economico;
- prepararsi ai relativi mutamenti socio-economici;
- garantire un adeguato quadro etico-giuridico fondato sui valori dell'Unione europea, inclusi i futuri orientamenti in materia di responsabilità per danno da prodotti difettosi[38].

La Commissione mette in luce le potenzialità dell'Europa in tale settore[39] rendendosi però conto di essere comunque in ritardo, specie nel settore degli investimenti privati, rispetto a Paesi come l'Asia o l'America del Nord. In tal senso l'Unione si propone di aumentare i propri investimenti in materia (specie nel settore della sanità, della guida automatizzata, dell'attività industriale e della giustizia), anche al fine di evitare una cospicua fuga di cervelli e di diventare mero consumatore di forme di intelligenza artificiale sviluppate altrove[40].

La Commissione evidenzia la necessità che l'intelligenza artificiale sia quanto più possibile accessibile e disponibile, nonché l'importanza di dar vita, da un lato, ad una "piattaforma di IA on demand" per fornire agli utilizzatori un unico punto di accesso alle risorse di intelligenza artificiale dell'Unione e, dall'altro, ai c.d.

---

[37] Nel definire cosa si intenda per Intelligenza Artificiale la Commissione si esprime nei seguenti termini: «"Intelligenza artificiale" (IA) indica sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere specifici obiettivi. I sistemi basati sull'IA possono consistere solo in *software* che agiscono nel mondo virtuale (per esempio assistenti vocali, software per l'analisi delle immagini, motori di ricerca, sistemi di riconoscimento vocale e facciale); oppure incorporare l'IA in dispositivi *hardware* (per esempio in robot avanzati, auto a guida autonoma, droni o applicazioni dell'Internet delle cose). Utilizziamo l'IA quotidianamente, per esempio per tradurre le lingue, generare sottotitoli nei video o bloccare lo spam delle email. Molte tecnologie di IA richiedono dati per migliorare le loro prestazioni. Raggiunto un buon livello di prestazioni, esse possono contribuire a migliorare e automatizzare il processo decisionale nello stesso campo. Per esempio, un sistema di IA verrà addestrato e in seguito utilizzato per rilevare gli attacchi informatici sulla base dei dati provenienti dal sistema o dalla rete interessati» *Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni, L'intelligenza artificiale per l'Europa*, COM(2018) 237 Bruxelles, 25.4.2018, p. 1, corsivo nostro.

[38] *L'intelligenza artificiale per l'Europa,* cit. pp. 3-4.

[39] Grazie ad alcuni progetti finanziati dall'Unione oggi è stato sviluppato un veicolo agricolo senza pilota, un progetto pilota di autostrada che fornisce raccomandazioni per una guida sicura, una ortoprotesi robotizzata per soggetti che hanno subito l'amputazione di un arto, nonché robot che svolgono mansioni ripetitive al posto dei lavori negli stabilimenti industriali, specie automobilistici. La Commissione ha altresì lanciato importanti iniziative in materia, come chip neuromorfici realizzati al fine di eseguire operazioni di IA, computer ad elevate prestazioni e tecnologie quantistiche e per la mappatura del cervello umano, *L'intelligenza artificiale per l'Europa,* cit., pp. 5-6.

[40] *L'intelligenza artificiale per l'Europa,* cit., p. 7.

"poli d'innovazione digitale" che aiutino le imprese a sfruttare le opportunità digitali, anche fornendo consulenze su come profittare di tali opportunità[41].

Nel prosieguo della Comunicazione in esame si sottolinea l'importanza della sperimentazione dei prodotti dotati di intelligenza artificiale (in particolar modo nei settori della sanità, dei trasporti e della digitalizzazione delle industrie), specie al fine di garantire la conformità a standard di sicurezza adeguati e di ideare quadri normativi appropriati.

La Commissione chiarisce che, al fine di un compiuto sviluppo dell'IA, è necessario accedere ad un ingente quantitativo di dati (ecco la ragione dell'importanza del ruolo rivestito dal Regolamento generale sulla protezione dei dati[42], noto anche come GDPR, nella materia *de qua*)[43] in quanto «l'apprendimento automatico, un tipo di IA, opera mediante l'individuazione di modelli a partire dai dati disponibili e la successiva applicazione di questa conoscenza ai dati nuovi. Quanto più è grande il set di dati, tanto più accurata sarà l'individuazione delle relazioni anche impercettibili tra i dati»[44].

Nel prosieguo la Commissione individua le tre sfide principali da affrontare in materia di intelligenza artificiale:

- preparare la società ai cambiamenti che l'IA porterà con sé;
- fornire assistenza a quei lavoratori che rivestono mansioni che subirebbero le maggiori trasformazioni o che, financo, scomparirebbero;
- formare un maggior numero di specialisti nel settore dell'intelligenza artificiale in modo tale da renderla inclusiva, interdisciplinare e non discriminatoria.

Ci si rende altresì conto di quanto sia importante, per lo sviluppo e l'utilizzo delle entità dotate di intelligenza artificiale, creare un ambiente improntato su fiducia e responsabilità, in ossequio ai diritti sanciti dall'art. 2 del Trattato sull'Unione Europea[45] e dalla Carta dei diritti fondamentali dell'UE.

---

[41] *L'intelligenza artificiale per l'Europa,* cit., p. 8.

[42] Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016. Sul tema v. F. PIZZETTI, *Protezione dei dati personali in Italia tra GDPR e codice novellato*, Torino, 2021.

[43] Al fine incrementare lo spazio dei dati europeo la Commissione ha proposto una serie di iniziative, tra cui si ricordano una direttiva aggiornata sulle informazioni del settore pubblico, orientamenti riguardanti la condivisione dei dati del settore privato nell'economia, una versione aggiornata della raccomandazione sull'accesso all'informazione scientifica e sulla sua conservazione, nonché una comunicazione sulla trasformazione digitale della sanità e dell'assistenza, così *L'intelligenza artificiale per l'Europa,* cit., p. 12.

[44] *L'intelligenza artificiale per l'Europa,* cit., pp. 10-11. In tal sede la Commissione coglie l'occasione per svolgere qualche ulteriore considerazione sul *deep learning*, chiarendo che: «L'apprendimento profondo è stato un elemento rivoluzionario per l'IA, che ha comportato un incredibile miglioramento delle prestazioni per compiti specifici quali il riconoscimento di immagini o vocale o la traduzione automatica. Addestrare un algoritmo di apprendimento profondo a classificare gli oggetti significa fornirgli una grande quantità di esempi etichettati (per esempio immagini) che sono correttamente categorizzati (per esempio immagini di aeroplani). Una volta addestrati, gli algoritmi possono classificare correttamente oggetti che non hanno mai visto, in alcuni casi con una precisione che supera quella umana. Progressi significativi in queste tecnologie sono stati ottenuti mediante l'impiego di grandi set di dati e una potenza di elaborazione senza precedenti». Il tema dei dati sarà approfondito al Cap. II, Sez. I, Par. 6.

[45] Trattato sull'Unione Europea, C 326/13, 26.10.2012, art. 2: «L'Unione si fonda sui valori del rispetto della dignità umana, della libertà, della democrazia, dell'uguaglianza, dello Stato di diritto e del rispetto dei diritti umani, compresi i diritti delle persone appartenenti a minoranze. Questi

Come anticipato, particolare rilevanza assume in tale settore il Regolamento generale sulla protezione dei dati personali[46], il quale garantisce un elevato livello di tutela, conferendo inoltre agli individui il diritto a non essere sottoposti ad una decisione fondata esclusivamente sul trattamento automatizzato[47], salvo i casi stabiliti espressamente dall'art. 22 del Regolamento medesimo[48].

La Commissione propone inoltre in tal sede una strategia per dar vita a un "mercato unico digitale"[49], chiarendo come simili proposte debbano essere adottate appena possibile «in quanto sia cittadini sia le imprese devono poter avere fiducia nella tecnologia con cui interagiscono, disporre di un contesto normativo prevedibile e contare su efficaci misure di salvaguardia che proteggano i loro diritti e le loro libertà fondamentali»[50].

L'intelligenza artificiale appare indubbiamente foriera di considerevoli opportunità ma anche di non secondari rischi (specie in materia di sicurezza, responsabilità e discriminazioni), per questo la Commissione si propone, nel documento in esame, di dar vita ad un progetto di orientamenti etici per

---

valori sono comuni agli Stati membri in una società caratterizzata dal pluralismo, dalla non discriminazione, dalla tolleranza, dalla giustizia, dalla solidarietà e dalla parità tra donne e uomini».

[46] Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016.

[47] Sul punto C. CASONATO, *Potenzialità e sfide dell'intelligenza artificiale*, in *BioLaw Journal*, 2019, p. 180.

[48] Il secondo comma dell'art. 22 stabilisce che il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato non si applica nei casi in cui: «a) sia necessaria per la conclusione o l'esecuzione di un contratto tra l'interessato e un titolare del trattamento; b) sia autorizzata dal diritto dell'Unione o dello Stato membro cui è soggetto il titolare del trattamento, che precisa altresì misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato; c) si basi sul consenso esplicito dell'interessato». Con particolare riferimento alla materia penalistica, analogo principio è sancito dall'art. 8 del d.lgs. n. 51 del 18.5.2018 il quale attua la Direttiva (UE) 2016/680 del Parlamento Europeo e del Consiglio del 27 aprile 2016 relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, nonché alla libera circolazione di tali dati e che abroga la decisione quadro 2008/977/GAI del Consiglio (G.U. n. 119 del 24.5.2018). Il succitato art. 8, rubricato "Processo decisionale automatizzato relativo alle persone fisiche", sancisce che: «1. Sono vietate le decisioni basate unicamente su un trattamento automatizzato, compresa la profilazione, che producono effetti negativi nei confronti dell'interessato, salvo che siano autorizzate dal diritto dell'Unione europea o da specifiche disposizioni di legge. 2. Le disposizioni di legge devono prevedere garanzie adeguate per i diritti e le libertà dell'interessato. In ogni caso è garantito il diritto di ottenere l'intervento umano da parte del titolare del trattamento. 3. Le decisioni di cui al comma 1 non possono basarsi sulle categorie particolari di dati personali di cui all'articolo 9 del regolamento UE, salvo che siano in vigore misure adeguate a salvaguardia dei diritti, delle libertà e dei legittimi interessi dell'interessato. 4. Fermo il divieto di cui all'articolo 21 della Carta dei diritti fondamentali dell'Unione europea, è vietata la profilazione finalizzata alla discriminazione di persone fisiche sulla base di categorie particolari di dati personali di cui all'articolo 9 del regolamento UE».

[49] Tra le proposte avanzate dalla Commissione in questo settore si ricordano il regolamento sulla libera circolazione dei dati non personali, il regolamento sulla e-privacy e la legge sulla sicurezza informatica. *L'intelligenza artificiale per l'Europa,* cit., p. 15.

[50] *L'intelligenza artificiale per l'Europa,* cit., p. 16. Nel prosieguo la Commissione chiarisce un passaggio importante, ossia il fatto che «Per rafforzare ulteriormente la fiducia è anche necessario che le persone comprendano come funziona la tecnologia, il che spiega l'importanza della ricerca nella *spiegabilità dei sistemi di IA*. In effetti, allo scopo di *aumentare la trasparenza* e *minimizzare il rischio di condizionamenti o errori*, i sistemi di IA dovrebbero essere sviluppati in modo da *permettere agli esseri umani di comprendere le loro azioni e la logica sottostante*», corsivo nostro.

l'Intelligenza Artificiale[51], con l'obiettivo di affrontare temi come il futuro del lavoro, l'equità, la sicurezza, l'inclusione sociale e la trasparenza degli algoritmi, esaminando l'impatto su una serie di diritti fondamentali tra cui la vita privata, la dignità, la tutela dei consumatori e la non discriminazione[52].

Nel prosieguo della Comunicazione in esame la Commissione evidenzia come la comparsa di strumenti di IA dotati di processi decisionali autonomi imponga una riflessione sulle attuali norme di diritto civile in materia di responsabilità[53], non comprendendosi del tutto come mai non ci si soffermi anche sugli aspetti concernenti la responsabilità penale.

Sul finire la Commissione incentiva la condivisione delle buone pratiche tra gli Stati membri, sostenendo una sinergia operativa e una cooperazione volta a coinvolgere partecipanti con ruoli molto diversi in grado di evitare la frammentazione del mercato unico. Per far ciò la Commissione si dice disponibile ad agevolare la creazione dell'"Alleanza europea per l'IA", una piattaforma multiculturale volta alla condivisione di informazioni utili e all'incentivazione di investimenti privati nel settore dell'intelligenza artificiale[54].

La Comunicazione in esame chiosa con l'intenzione di dar vita, entro il 2018, non soltanto alla summenzionata Alleanza europea per l'IA ma anche ad un "piano coordinato" con gli Stati membri volto alla condivisione di informazioni utili e ad affrontare le nuove implicazioni etiche e giuridiche che l'evolversi della materia *de qua* è destinata a portare con sé.

---

[51] *L'intelligenza artificiale per l'Europa,* cit., p. 16.

[52] La Commissione in tal sede si dichiara altresì disponibile a rivalutare i quadri giuridici attualmente esistenti per adattarli alle nuove esigenze dell'IA e per garantire il rispetto dei summenzionati diritti. Nel prosieguo della trattazione la Commissione si attribuisce altresì i seguenti compiti: istituire un quadro normativo per esperti e soggetti coinvolti (denominata "Alleanza europea per l'IA") per elaborare il succitato progetto di orientamenti etici; emanare un documento di orientamento sull'interpretazione della direttiva sulla responsabilità per danno da prodotti difettosi; pubblicare una relazione sui quadri normativi in materia di responsabilità e sicurezza in relazione all'IA, indicando implicazioni, lacune e orientamenti; sostenere la ricerca sullo sviluppo dell'IA spiegabile; sostenere le organizzazioni dei consumatori e le autorità garanti della protezione dei dati a livello nazionale ed europeo per sensibilizzare le implicazioni connesse all'utilizzo dell'IA. *L'intelligenza artificiale per l'Europa,* cit., p. 18.

[53] Invero il "quadro normativo dell'UE sulla sicurezza" è costituito da una solida base normativa che si adatta al progresso tecnologico. Si ricorda in tal sede la direttiva in materia di responsabilità per danno da prodotti difettosi, la quale «stabilisce che se un prodotto difettoso causa danni al consumatore o ai suoi beni, il produttore deve provvedere al risarcimento indipendentemente dal fatto che vi sia o meno negligenza o colpa del produttore», e la direttiva macchine, nonché una recente valutazione iniziale dell'attuale quadro normativo in materia di responsabilità in rapporto all'IA e alle tecnologie emergenti. Ad essere particolarmente esposti, in ragione della diffusione di strumenti dotati di intelligenza artificiale, sono proprio i consumatori, i quali hanno il diritto di ricevere informazioni chiare e complete sul funzionamento dell'IA ed, in particolare, su come contattare un soggetto umano e su come garantire la controllabilità delle decisioni dello strumento medesimo. *L'intelligenza artificiale per l'Europa,* cit., pp. 17-18.

[54] In tal sede la Commissione richiama alcune delle esperienze già realizzate da taluni Stati: «Il 29 marzo 2018, la Francia ha presentato la propria strategia nazionale per l'IA, basata sulla relazione Villani. La Germania, seguendo l'esempio dell'"Industria 4.0", ha istituito una piattaforma sui sistemi di apprendimento per attivare un dialogo strategico tra il mondo accademico, l'industria e il governo, e ha proposto una relazione sull'etica della guida automatizzata e connessa. La Finlandia ha proposto la sua strategia "Tekoälyaika" per essere all'avanguardia nel settore. Ogni Stato membro è incoraggiato a predisporre una strategia sull'IA e sui relativi investimenti» *L'intelligenza artificiale per l'Europa,* cit., p. 19.

In conclusione, la Commissione si dice consapevole delle potenzialità dell'Unione europea in materia di intelligenza artificiale, non soltanto in ragione del primato da quest'ultima rivestito nel settore in esame ma anche alla luce del solido quadro normativo posto a protezione dei consumatori, evidenziando da ultimo la necessità di adottare un approccio inclusivo che tenga sempre come punto fermo il rispetto dei diritti fondamentali dell'Unione.

4. La Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi.

Proseguendo la nostra disamina sugli interventi sovranazionali di regolamentazione dei sistemi intelligenti, non si può far a meno di prendere in esame la Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi.

Il 3 dicembre 2018 la Commissione Europea per l'efficienza della giustizia (anche nota come CEPEJ) ha pubblicato la succitata Carta etica, indirizzandola agli operatori pubblici e privati su cui ricada il compito di creare strumenti dotati di intelligenza artificiale in materia di trattamento dei dati giudiziari e relative decisioni al fine di orientarli in questo delicato compito e di migliorare la qualità e l'efficienza dei servizi giudiziari, sempre nel rispetto dei diritti fondamentali e dei principi di trasparenza, imparzialità ed equità.

Come si avrà modo di evincere nel prosieguo, il contenuto della presente Carta assume rilievo principalmente nella materia processualpenalistica, ponendo le basi per il futuribile utilizzo dei sistemi intelligenti nelle corti di giustizia.

In tal sede il CEPEJ ha individuato i cinque principi cardine concernenti l'utilizzo dell'intelligenza artificiale nei sistemi giudiziari che di seguito si elencano:

1. *Principio del rispetto dei diritti fondamentali*: la Commissione ritiene imprescindibile assicurare l'elaborazione e l'attuazione di strumenti e servizi di intelligenza artificiale che siano compatibili con i diritti fondamentali, in particolar modo con il diritto di accesso al giudice e con il diritto ad un processo equo (tanto in termini di parità delle armi quanto nell'ottica del rispetto del contraddittorio)[55];

2. *Principio di non discriminazione*: alla luce della classificazione dei dati effettuata da tali sistemi e della possibilità che il loro uso produca effetti discriminatori, si evidenzia l'importanza di prevenire lo sviluppo o l'intensificazione di discriminazioni tra persone o gruppi di persone, vigilando sull'elaborazione e sull'uso di questi meccanismi, specie quando essi si servano per il loro funzionamento di dati sensibili[56];

3. *Principio di qualità e sicurezza*[57]: per quanto riguarda il trattamento delle decisioni e dei dati giudiziari appare necessario servirsi di fonti certificate e dati

[55] Il principio dell'equo processo è sancito, a livello sovranazionale, dall'art. 6 della Convenzione Europea dei diritti dell'Uomo e, a livello interno, dall'art. 111 della Costituzione italiana.

[56] Parte della dottrina parla di «*human right by design*» evidenziando la necessità di evitare un uso discriminatorio dei dati utilizzati dall'IA sin dalla progettazione del sistema intelligente, così C. CAVACEPPI, *L'intelligenza artificiale applicata al diritto penale: criticità attuali e prospettive future*, in G. TADDEI ELMI, A. CONTALDO, *Intelligenza artificiale. Algoritmi giuridici,* Ius condendum *o "fantadiritto"?*, Pisa, 2020, p. 121.

[57] La dottrina evidenzia l'importanza di garantire l'uso di dati qualitativamente elevati, in particolar modo, con riguardo al processo penale, sede in cui è possibile incidere sulla libertà

intangibili con modelli elaborati multidisciplinarmente ed in un ambiente tecnologico sicuro, garantendo la tracciabilità del processo e la collaborazione dei professionisti del settore (coinvolgendo in egual misura dottrina e giurisprudenza);

4. *Principio di trasparenza, imparzialità ed equità*: la Commissione evidenzia la necessità di rendere accessibili e comprensibili le metodologie di trattamento dei dati, ricercando un equilibrio tra il diritto alla proprietà intellettuale di alcune di queste metodiche ed i principi di trasparenza, imparzialità ed equità, specie in ragione delle conseguenze giuridiche che dall'uso di tali metodologie possono derivare sui consociati;

5. *Principio del "controllo da parte dell'utilizzatore"*[58]: gli operatori del settore giustizia devono sempre avere la possibilità di rivedere le decisioni assunte dai sistemi intelligenti[59]. Si mira pertanto ad evitare un approccio prescrittivo e a garantire che gli utilizzatori siano attori informati e abbiano il controllo delle loro scelte. Il mero utilizzatore, di contro, deve essere reso edotto del carattere vincolante o meno delle soluzioni proposte dal meccanismo di IA e del suo diritto a ricevere assistenza legale e di opporsi ad un procedimento giudiziario automatizzato, potendo sempre adire un tribunale.

La Carta evidenzia nel prosieguo come, allo stato attuale, la maggior parte delle iniziative nel settore si registrino in ambito privato, specie nel campo dei settori legali ed assicurativi, al fine di ridurre l'incertezza e l'imprevedibilità delle decisioni giudiziarie[60].

In testa ai Paesi nei quali si registra un utilizzo diffuso di algoritmi dotati di intelligenza artificiale nel settore giustizia troviamo gli Stati Uniti, i quali hanno molto investito in tal strumenti, tanto in ambito civile quanto in ambito penale[61].

La CEPEJ individua come condizione essenziale per lo sviluppo dell'intelligenza artificiale la disponibilità di un'ingente mole di dati[62] ed affronta

---

personale degli individui: «più il dato è trasparente e intellegibile maggiore è la sua qualità». Con riferimento, invece, al requisito della sicurezza del dato si evidenzia come quest'ultimo si riferisca tanto alla provenienza del dato medesimo quanto alle sue modalità di trattamento, così C. CAVACEPPI, *L'intelligenza artificiale applicata al diritto penale*, cit., pp. 122-123.

[58] Definito anche come principio «under user control» da A. TRAVERSI, *Intelligenza artificiale applicata alla giustizia: ci sarà un giudice* robot*?*, in *Questione Giustizia*, 10.4.2019, p. 3.

[59] Ciò al fine di «evitare il rischio della c.d. *dittatura del precedente*», così C. CAVACEPPI, *L'intelligenza artificiale applicata al diritto penale*, cit., p. 104, corsivo dell'A.

[60] I principali settori operativi in cui l'intelligenza artificiale può trovare spazio possono essere così sintetizzati: «Motore di ricerca giurisprudenziale avanzato; Risoluzione delle controversie online; Assistenza nella redazione di atti; Analisi (predittiva, tabelle); Categorizzazione dei contratti secondo criteri diversi e individuazione delle clausole contrattuali divergenti o incompatibili; "Chatbots" per informare le parti in lite o sostenerle nel procedimento giudiziario» *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, Strasburgo, 3.12.2018, Punto 15, pp. 15-16.

[61] Viene riportata, in termini comparatistici, l'esperienza francese la quale, con la propria legislazione concernente la diffusione dei dati aperti delle decisioni giudiziarie, ha accolto il principio secondo il quale ogni decisione giudiziaria è pubblicabile, eccetto casi specifici legislativamente individuati e purché ciò avvenga nel rispetto della vita privata. Ad avviso della Corte di Cassazione francese un simile meccanismo garantisce una maggiore conoscibilità degli orientamenti giurisprudenziali ed una più elevata qualità delle decisioni del sistema giudiziario che «sa di essere osservato» *Carta etica europea,* cit., Punto 33, p. 19.

[62] Dopo aver distinto l'accesso alle informazioni (ossia informazioni già pubbliche e divulgate) dall'accesso ai dati (ossia dati aperti contenuti in specifiche banche dati) la Carta distingue i dati

il problema dell'anonimizzazione (ossia il trattamento dei dati funzionale all'impedimento totale e irreversibile dell'identificazione di una persona fisica o giuridica)[63] e della pseudonimizzazione dei dati medesimi (ossia il trattamento dei dati personali con modalità tali da rendere impossibile (ri)attribuirli ad un soggetto specifico se non con l'utilizzo di informazioni aggiuntive)[64].

Proprio con riferimento al rapporto tra anonimizzazione e pseudonimizzazione vengono distinte le ipotesi in cui a dover essere cancellati siano i nomi delle parti o i nomi dei professionisti coinvolti nel processo.

La rimozione dei nominativi delle parti o dei testimoni presenti all'interno delle decisioni pubblicate può avvenire con procedure automatizzate e costituisce un corretto equilibrio tra i diritti fondamentali degli attori del processo e l'esigenza di rendere pubbliche le relative decisioni giudiziarie.

Maggiori problemi pone invece la cancellazione dei nomi dei professionisti contenuti all'interno di tali decisioni, in particolar modo quelli dei giudici. Tale opzione consente di evitarne la profilazione e, conseguentemente, il fenomeno del c.d. *forum shopping*[65] ma potrebbe, di contro, porsi in contrasto con il principio della pubblicità del processo sancito dall'art. 6 della Convenzione europea dei diritti dell'uomo[66].

Un buon contemperamento potrebbe essere quello prescelto dalla Corte europea dei diritti dell'uomo, la quale consente la ricerca delle decisioni giurisprudenziali in base ai nominativi dei membri dell'organo giudicante ma, di contro, non consente la redazione di statistiche relative al singolo giudice[67].

Nel prosieguo la Carta Etica si concentra sulle caratteristiche operative dell'intelligenza artificiale applicata alle decisioni giudiziarie, chiarendo già in incipit che le due tecniche base nell'ambito in esame possono individuarsi nel trattamento del linguaggio naturale e nell'apprendimento automatico[68].

---

medesimi dalle loro modalità di trattamento (note anche come "scienza dei dati"), *Carta etica europea*, cit., Punti 22-23-25, pp. 17-18.

[63] Definizione fornita dal Glossario contenuto al termine della Carta etica in esame, p. 45.

[64] Art. 4 del *Regolamento Generale sulla Protezione dei Dati*, Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016.

[65] Tattica consistente nella scelta del tribunale che ha dimostrato una propensione particolare verso un certo orientamento giurisprudenziale o che appare maggiormente incline a concedere risarcimenti più elevati, *Carta etica europea*, cit., Punto 46, pp. 21-22.

[66] «La risposta alla domanda della legittimità o meno della pubblicazione dei nominativi dei professionisti tra i dati aperti non ha pertanto niente a che fare con l'obbligo di pubblicare i nominativi dei professionisti nelle decisioni. Sembra piuttosto che la questione concerna spesso la conciliazione di requisiti contrastanti: da una parte rendere trasparenti le attività pubbliche permettendo ai cittadini di conoscere e valutare i loro giudici, e dall'altra proteggere la privacy dei professionisti (le cui funzioni non dovrebbero limitare le loro garanzie fondamentali in tale campo). (…) Tali domande non hanno la stessa forma in ogni luogo in Europa e dipendono dalle specifiche caratteristiche del sistema giudiziario interessato (e dalla natura dell'organo di gestione delle carriere dei giudici), dal carattere collegiale o meno della sentenza e dal grado del tribunale interessato. In Svizzera, per esempio, dove i giudici sono eletti, la pubblicazione è una garanzia della trasparenza e della responsabilità sociale dei giudici rispetto ai cittadini e ai gruppi politici» *Carta etica europea*, cit., Punti 50-51, p. 22.

[67] *Carta etica europea*, cit., Punto 52, p. 22.

[68] Tali metodiche vengono considerate come forme di intelligenza artificiale debole, ossia intelligenze artificiali in grado di rendere «alte prestazioni nel loro ambito di addestramento», definizione fornita dal Glossario contenuto al termine della Carta etica in esame, p. 47. In tal sede la CEPEJ distingue le intelligenze artificiali deboli da quelle forti, ossia forme di IA «capaci di contestualizzare problemi specializzati di varia natura in maniera completamente autonoma». Per

Prendendo in esame le caratteristiche dei c.d. "software predittivi" si evidenzia come il loro scopo sia quello di stabilire, mediante un'analisi statistica delle precedenti decisioni, il possibile esito di una causa innanzi ad un tribunale, potendo in ciò coadiuvare il lavoro degli avvocati (nello strutturare la loro linea difensiva) e dei giudici (per fornir loro un supporto nell'assumere le loro determinazioni).

Si tratta di software che costruiscono modelli statistici basati sull'analisi di dati del passato, in grado di apprendere da tale mole di dati e di estrarne modelli complessi dotati di una considerevole precisione predittiva[69].

Ci si è voluti concentrare, seppur brevemente, su queste considerazioni di carattere tecnico in quanto, a ben vedere, nasce tra le loro pieghe un interrogativo strettamente connesso alla giustizia penale: prendendo in esame prevalentemente i dati del passato, gli esiti forniti dal software in termini predittivi (*rectius*, di previsione) rischiano di essere sempre uguali a sé stessi e di non incoraggiare le interpretazioni evolutive della giurisprudenza che, ove ossequiasse pedissequamente gli esiti suggeriti dall'IA, finirebbe per adagiarsi sempre sul precedente.

L'intento di conciliare la certezza e la prevedibilità delle decisioni con l'evoluzione dell'interpretazione giudiziaria porta però con sé, inoltre, una serie di rischi. *In primis* – come anticipato – quello di una giurisprudenza fondata esclusivamente sul precedente (logica in astratto incompatibile con i sistemi di *civil law*, quali il nostro)[70] e, *in secundis*, quello di una omogeneizzazione delle decisioni dei giudici, i quali non sarebbero più chiamati a decidere in funzione della sola legge ma dovrebbero, altresì, adattarsi agli indirizzi giurisprudenziali che si possono trarre dalle statistiche prodotte dal sistema intelligente.

---

meglio spiegare il funzionamento di tali meccanismi la CEPEJ chiarisce che «Nella maggior parte dei casi, l'obiettivo di tali sistemi non è la riproduzione di un ragionamento giuridico, bensì l'individuazione delle correlazioni tra i diversi parametri di una decisione (per esempio, in una domanda di divorzio, la durata del matrimonio, il reddito dei coniugi, la presenza di adulterio, l'importo dell'assegno di mantenimento concesso, ecc.) e, mediante l'utilizzo dell'apprendimento automatico, dedurne uno o più modelli. Tali modelli sarebbero successivamente utilizzati per "predire" o "prevedere" una futura decisione giudiziaria» *Carta etica europea,* cit., p. 23.

[69] *Carta etica europea,* cit., Punti 56-58, p. 23. In tal sede viene tracciata una netta linea di demarcazione tra il concetto di predizione e quello di previsione: «La predizione è l'atto di annunciare anticipatamente (*prae*, prima - *dictare*, dire) gli avvenimenti futuri (per ispirazione sovrannaturale, chiaroveggenza o premonizione). La previsione, d'altra parte, è il risultato dell'osservazione (*visere*, vedere) di un insieme di dati al fine di prevedere una situazione futura» *Carta etica europea,* cit., Punto 60, p. 24.

[70] «Premesso che l'accessibilità digitale a tutte le sentenze pronunciate in un ordinamento (open data) – con possibilità di analisi e ricerca attraverso parole chiave e raggruppamenti di parole – non equivale e non sostituisce il tradizionale principio di pubblicità delle decisioni giudiziarie, diversamente garantito, è opportuno domandarsi quale effetto produrrà tale diffusa e illimitata accessibilità sul valore del 'precedente', soprattutto negli ordinamenti che non sono su di esso basati. La forte correlazione tra un certo gruppo di fattori e una determinata decisione giudiziaria, rivelata da un sistema computazionale (non necessariamente molto sofisticato), può determinare uno 'scivolamento normativo' verso una maggiore vincolatività del precedente, imponendo un onere motivazionale rafforzato al giudice che se ne voglia distaccare?» S. QUATTROCOLO, *Intelligenza artificiale e giustizia: nella cornice della Carta Etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche,* in *La Legislazione Penale,* 22.3.2018, pp. 4-5.

Al fine di fornire maggiore concretezza alla disamina in corso la Carta precisa che tali tecnologie si fondano su tre principi fondamentali:

- l'intelligenza artificiale non può essere ridotta ad un singolo concetto omogeneo, costituendo piuttosto un insieme di diverse scienze;
- i vari modelli di intelligenza artificiale adottano un approccio induttivo, associando input (dati inseriti dall'uomo) e possibili output (esiti forniti dall'IA). Applicando tale ragionamento al settore giustizia i dati in ingresso sono rappresentati dai fatti e dalle motivazioni, i dati in uscita invece sono rappresentati dai dispositivi delle sentenze;
- l'affidabilità della risposta fornita dall'IA dipende integralmente dalla tecnica di apprendimento automatico prescelta e, soprattutto, dalla qualità dei dati inseriti dall'uomo[71].

La Commissione chiarisce nel prosieguo che una completa modellizzazione informatica del ragionamento giuridico appare di difficile realizzazione – anche da parte dei c.d. meccanismi di apprendimento automatico (i quali imparano dall'esperienza)[72] – ciò in considerazione del fatto che l'iter logico seguito dal giudice dipende da una moltitudine di fattori non predeterminabili, come la soggettiva interpretazione del concetto di equità, l'applicazione del principio di proporzionalità o l'applicazione del proprio potere discrezionale[73].

---

[71] *Carta etica europea*, cit., Punto 64, p. 24. Sul rischio di *implicit bias* si esprime efficacemente S. QUATTROCOLO, *Intelligenza artificiale e giustizia*, cit., p. 6, chiarendo che «Plurimi sono i livelli ai quali il rischio si può verificare. In primo luogo, se l'*input* non è completamente neutro, l'*output* dell'interrogazione rischia di essere influenzato da un pregiudizio, che può portare alla discriminazione di singoli individui o di gruppi sociali. In secondo luogo, l'algoritmo – che è concepito e interpretato da un umano – può banalmente riprodurre ingiustificati preconcetti sociali».

[72] La CEPEJ distingue i succitati meccanismi di apprendimento automatico dai vecchi sistemi esperti, ossia sistemi fondati su regole scritte da un informatico, programmati originariamente per riprodurre il ragionamento giuridico. La CEPEJ chiarisce che i metodi di apprendimento automatico includono al loro interno le c.d. reti neurali, definite dalla stessa come sistemi che «"apprendono" a svolgere dei compiti prendendo in considerazione degli esempi, generalmente senza essere programmati con regole specifiche per il compito (…) Le reti artificiali neurali (ANN artificial neural networks) si basano su un insieme di unità connesse o nodi, denominati neuroni artificiali, modellati genericamente sulla falsariga dei neuroni di un cervello biologico. Ciascuna connessione, come le sinapsi di un cervello biologico, può trasmettere un segnale da un neurone artificiale a un altro. Un neurone artificiale che riceve un segnale può trattarlo e inoltrarlo successivamente ad altri neuroni artificiali cui è connesso», definizione fornita dal Glossario contenuto al termine della Carta etica in esame, p. 48.

[73] La Carta chiarisce efficacemente nel prosieguo che «tale lavoro di interpretazione è esattamente quello che al giorno d'oggi le tecniche di apprendimento automatizzato non svolgono - e non cercano di svolgere - in quanto esse, come abbiamo visto, effettuano elaborazioni automatizzate basate sul presupposto che la correlazione di grandi volumi di informazioni possa sostituire la comprensione dei veri nessi causali di una decisione. Non tentano di formalizzare il ragionamento giuridico, ma sperano che i modelli registrati da esse possano prevedere le probabili decisioni di un giudice in situazioni analoghe» *Carta etica europea*, cit., Punto 79, pp. 27-28. Una conferma di tale considerazione si è avuta all'esito di un esperimento condotto dalla University College of London il quale ha cercato di dimostrare, esaminando la giurisprudenza della Corte europea dei diritti dell'uomo, che un modello di apprendimento automatico potesse predire l'esito di una causa innanzi a un dato tribunale con un elevato grado di accuratezza: l'esito di tale lavoro ha dimostrato una maggiore accuratezza nella ricostruzione dei fatti piuttosto che nella ricostruzione del ragionamento giuridico. *Carta etica europea*, cit., Punto 81, p. 28.

La Carta si dedica poi all'analisi dei possibili utilizzi dell'intelligenza artificiale nell'ambito della giustizia civile[74] e penale.

Con particolare riferimento alla materia penalistica, l'intelligenza artificiale può trovare applicazione relativamente al tema della prevedibilità della condotta criminale[75] ed ai c.d. strumenti di "polizia predittiva", volti a prevenire la commissione di reati[76].

La CEPEJ si concentra nel prosieguo su due dei più noti strumenti di *risk assessment tools* (strumenti di valutazione del rischio) giunti a noi dall'esperienza anglosassone e statunitense.

In Inghilterra l'Università di Cambridge ha realizzato un sistema (dal valore meramente consultivo e sempre sottoposto al controllo stringente delle forze di polizia) denominato HART (*Harm Assessment Risk Tool*) con lo scopo di valutare il rischio di recidiva di un determinato soggetto prendendo in esame una trentina di fattori, alcuni dei quali indipendenti dal reato commesso[77].

Negli Stati Uniti è invece stato realizzato il COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) il quale ha il compito di coadiuvare il giudice nel valutare il rischio di recidiva in fase di determinazione edittale. Tale algoritmo sfrutta le informazioni contenute nel casellario giudiziale dell'imputato e le risposte fornite da quest'ultimo ad un elenco di 137 domande per fornire una risposta in scala che dovrebbe rispecchiare il rischio di recidiva del soggetto.

La ONG ProPublica ha denunciato gli esiti discriminatori dell'algoritmo, dimostrando come quest'ultimo attribuisse un più elevato tasso di rischio di

---

[74] In ambito civilistico tali sistemi possono essere adoperati come strumento computazionale per il calcolo delle tabelle oppure impiegati per le c.d. *online dispute resolution*, tenendo sempre ferme le garanzie fornite dal principio del contraddittorio, del diritto di accesso ad un tribunale, della parità delle armi, imparzialità del decidente e diritto all'assistenza legale. La Carta ricorda in tal sede l'esperimento avvenuto nelle Corti d'appello di Rennes e di Douai, le quali avevano accettato di testare un software di giustizia predittiva il cui obiettivo era quello di «creare uno strumento di ausilio alla decisione in modo da ridurre, se necessario, l'eccessiva variabilità delle decisioni giudiziarie, in nome del principio di uguaglianza dei cittadini di fronte alla legge. L'esito dell'esperimento, oggetto di un dibattito contraddittorio tra le due Corti di appello, il Ministero della giustizia e la società di legal tech che aveva progettato il prodotto, ha sfortunatamente indicato l'assenza di valore aggiunto della versione del software testato per il lavoro di riflessione e il processo decisionale dei magistrati» *Carta etica europea*, cit., pp. 30-31.

[75] L'approccio al succitato tema pone particolari problemi proprio con riguardo all'esperienza italiana la quale sancisce, al secondo comma dell'art. 220 c.p.p., il divieto di effettuare perizie per stabilire l'abitualità o la professionalità al reato, nonché la tendenza a delinquere, la personalità dell'imputato o le sue qualità psichiche.

[76] Tali strumenti informatici vengono utilizzati per individuare i possibili luoghi in cui un dato reato potrebbe essere commesso trasponendo tali indicazioni su una mappa geografica in modo tale da individuare i punti nevralgici da sottoporre a costante osservazione. La Carta Etica parla a tal riguardo di "*predictive criminal mapping*", meccanismo alimentato dai precedenti elementi di localizzazione di certi reati forniti dalle forze di polizia. Tali strumenti possono, da un lato, generare un effetto dissuasivo in grado di prevenire la commissione di un dato crimine, dall'altro però tale sistema comporta il rischio delle c.d. "profezie che si auto-adempiono" in quanto la polizia potrebbe scoprire un maggior numero di reati in quanto massicciamente concentrata nella zona considerata "a rischio" dall'algoritmo. Si dovrebbe in tale ambito evitare altresì che il giudizio umano venga annullato per ossequiare ciecamente lo strumento dotato di IA, ricordando sempre che quest'ultima deve limitarsi alla funzione di strumento volto a coadiuvare l'attività umana e non a sostituirla. *Carta etica europea*, cit., Punti 120-121, p. 35.

[77] *Carta etica europea*, cit., Punti 125 ss., p. 36.

recidiva ai membri della popolazione afro-americana piuttosto che ad altri gruppi etnici[78].

I profili critici del COMPAS non si sono però arrestati qui.

Essendo stato sviluppato da una società privata[79] (ed in ragione del rispetto della proprietà intellettuale), i processi di funzionamento di questo algoritmo difettano di trasparenza[80].

Tali strumenti predittivi sono in grado di incidere sulla libertà personale dell'imputato: per tale ragione è di fondamentale importanza che essi siano costantemente monitorati dagli attori giudiziari che se ne servono.

I sostenitori dell'utilizzo di tali tecnologie adducono, come elementi a favore la loro efficienza e neutralità, il fatto che esse siano basate su metodi oggettivi e trasparenti. Di contro l'inserimento di dati relativi ai precedenti penali di un soggetto ed al suo contesto sociale di provenienza rischia di stigmatizzare un dato gruppo e di non tener conto delle specificità di ciascun essere umano[81].

Il reale problema di fondo – che rende la neutralità dell'algoritmo concretamente insostenibile – è rappresentato dal fatto che tali strumenti sono realizzati dall'uomo e, conseguentemente, rischiano di risentire dei pregiudizi di quest'ultimo e di alimentare diseguaglianze non giustificate e già presenti nel sistema.

Al fine di ridurre i possibili esiti pregiudizievoli che potrebbero derivare dall'uso di tali strumenti ed al fine di garantire la parità delle armi e la presunzione di innocenza di cui all'articolo 6 CEDU, ogni parte interessata dovrebbe essere posta in condizione di contestare la validità scientifica dell'esito fornito dall'algoritmo[82], principio che nel nostro ordinamento parrebbe già garantito dal disposto dell'art. 15 comma 1 lett. h del Regolamento generale sulla protezione dei dati[83].

---

[78] *Carta etica europea,* cit., Punti 128 ss., p. 36. Per un approfondimento della succitata inchiesta https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[79] Queste società ricevono i dati necessari ad alimentare il software proprio dalle autorità statali, senza però che a queste ultime possa essere imputata alcuna responsabilità nei confronti dei consociati, ponendo conseguentemente un serio problema di democrazia. *Carta etica europea,* cit., Punto 131, p. 36.

[80] La dottrina riflette in termini chiarificatori sul principio della trasparenza (ossia il quarto dei cinque principi individuati in apertura dalla Carta in esame), chiarendo che l'interesse proprio della giustizia alla comprensibilità e verificabilità dei procedimenti computazionali deve prevalere sugli interessi privati alla protezione della proprietà intellettuale, così S. QUATTROCOLO, *Intelligenza artificiale e giustizia,* cit., p. 7.

[81] La CEPEJ fornisce un utile esempio a riguardo: «un giudice potrebbe disporre il rilascio su cauzione di un'autrice di reato a rischio di recidiva sulla base di una gerarchia di valori, per esempio attribuendo maggiore importanza al suo ruolo di madre e di protettrice dei suoi figli, mentre l'algoritmo sarebbe in grado di determinare il rischio di recidiva con maggiore precisione, ma non sarebbe in grado di operare una simile gerarchia di priorità» *Carta etica europea,* cit., Punto 136, p. 37.

[82] «Tutte le persone hanno diritto a non essere sottoposte a una decisione che incide significativamente sulla loro persona, adottata unicamente sulla base di un trattamento automatico di dati, senza che si tenga preliminarmente conto del loro punto di vista» *Carta etica europea,* cit., Punto 138, pp. 37-38.

[83] «L'interessato ha il diritto di ottenere dal titolare del trattamento la conferma che sia o meno in corso un trattamento di dati personali che lo riguardano e in tal caso, di ottenere l'accesso ai dati personali e alle seguenti informazioni: (…) h) l'esistenza di un processo decisionale automatizzato, compresa la profilazione di cui all'articolo 22, paragrafi 1 e 4, e, almeno in tali casi, informazioni significative sulla logica utilizzata, nonché l'importanza e le conseguenze previste di

In tale ambito si porrebbe un ulteriore e non secondario problema: il giudice che dovesse decidere di determinarsi in senso contrario rispetto a quanto stabilito dall'algoritmo si troverebbe ad assumersi una maggiore responsabilità e, magari, nella posizione di dover motivare anche in ordine alle ragioni per le quali riterrebbe di disattendere le indicazioni del software. Ciò potrebbe comportare il rischio che il giudice appiattisca la propria decisione sugli esiti dell'algoritmo per non farsi carico di un ulteriore onere motivazionale, così però riducendo le garanzie dei soggetti coinvolti nel processo[84].

La Carta in esame riprende poi il nodale argomento della protezione dei dati personali, ritenendo di fondamentale importanza l'applicazione del (già citato) principio di precauzione nelle politiche di valutazione del rischio. Le misure volte a prevenire i potenziali rischi di un uso improprio di siffatti dati potrebbero trovare applicazione già in fase di progettazione o, financo, per impostazione predefinita[85].

In tal sede la Carta richiama inoltre il principio di lealtà, il cui rispetto impone che i dati personali vengano trattati per scopi legittimi e con modalità che il titolare di tali dati non possa ritenere inappropriate, dovendosi sempre evitare distorsioni discriminatorie o conseguenze negative per le libertà fondamentali dei soggetti coinvolti[86].

La CEPEJ opina poi nel senso di un totale abbandono dell'espressione "giustizia predittiva" in ragione della sua eccessiva ambiguità chiarendo, da un lato, come tali strumenti si fondino su un'analisi statistica della giurisprudenza (senza tentare di riprodurre il ragionamento giuridico) e rimarcando, dall'altro, la necessità che l'uso e la progettazione di siffatti strumenti venga rapportato ad un solido quadro etico[87].

In prospettiva futuribile la CEPEJ sottolinea l'importanza di un approccio multidisciplinare, che porti i realizzatori di tali strumenti e gli operatori del diritto a lavorare in sinergia sin dall'inizio, anche al fine di delineare una vera e propria "cyberetica" volta a garantire la trasparenza e la correttezza degli algoritmi che un domani potrebbero coadiuvare il giudice nell'adozione delle proprie decisioni.

La Commissione si dice consapevole della difficoltà di delineare un quadro normativo concernente l'intelligenza artificiale in ragione della transnazionalità del proprio ambito applicativo, tuttavia non perde occasione per ribadire l'essenzialità di un corpo di norme in materia volto a spianare la strada agli

---

tale trattamento per l'interessato» Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016.

[84] La Commissione autrice della Carta in esame evidenzia ancora il rischio che i risultati proposti dal software di giustizia predittiva assurgano al rango di vera e propria norma, considerando di contro la possibilità per gli operatori del diritto di aggirare tale ostacolo disattendendo tali risultati. *Carta etica europea*, cit., Punti 149-150, p. 39.

[85] *Carta etica europea*, cit., Punti 141 ss., p. 38. Sul concetto di *privacy by design* e *privacy by default* v. Cap. II, Sez. I, Par. 6.

[86] «Quando si utilizza l'intelligenza artificiale, i diritti degli interessati rivestono particolare importanza e il controllo che spetta a ciascuno di noi sulle proprie informazioni personali implica che debba essere possibile esercitare i seguenti diritti: il diritto degli interessati di non essere sottoposti a decisioni automatizzate che incidano significativamente su di essi, senza che si tenga conto del loro punto di vista, il diritto di ottenere informazioni sul ragionamento alla base del trattamento dei dati effettuato dagli algoritmi, il diritto di opporsi a tale trattamento, e il diritto a un ricorso giurisdizionale» *Carta etica europea*, cit., Punto 145, p. 38.

[87] *Carta etica europea*, cit., p. 38.

operatori del settore e a garantire i principi di trasparenza, correttezza e neutralità di tali strumenti, i quali devono essere sottoposti ad un costante monitoraggio da parte di esperti indipendenti per cercare di prevenire derive discriminatorie[88].

La Carta evidenzia, in conclusione[89], che un compiuto sviluppo della cyberetica richiede una specifica formazione dei progettisti degli algoritmi ed un coinvolgimento delle scienze umane interdisciplinari «affinché l'intelligenza artificiale diventi un vettore di sviluppo positivo del genere umano»[90].

Tale strumento di *soft law* costituisce significativo punto di partenza per l'introduzione dei sistemi di IA nell'ambito giudiziario e porta con sé il pregio di tentare (mirabilmente) di conciliare due scienze apparentemente incompatibili: le scienze giuridiche e le c.d. *hard science*[91].

5. Il Piano coordinato sull'intelligenza artificiale.

Tale testo merita di essere fatto oggetto di una (seppur succinta) analisi in quanto richiamato non solo nella Comunicazione "L'intelligenza artificiale per l'Europa", ma anche dai successivi contributi dell'Unione in materia. Esso costituisce uno dei passaggi fondamentali degli interventi europei nel settore e merita di essere ricordato in quanto propone una strategia operativa per agire con cognizione di causa nel settore dell'intelligenza artificiale, anche per quanto riguarda l'ambito giuridico.

---

[88] *Carta etica europea,* cit., Punti 153-157-159, pp. 40-41. «A tale proposito si potrebbe certificare la qualità dei sistemi migliori con l'assegnazione di un marchio o di una certificazione. In particolare deve essere garantita la totale trasparenza e la perfetta correttezza delle modalità di trattamento dell'informazione sia nei confronti dei professionisti che nei confronti dei cittadini, per impedire il ripetersi di errori come quelli commessi dall'algoritmo COMPAS. I professionisti della giustizia debbono partecipare attivamente in modo da essere in grado di valutare correttamente i rischi e l'impatto di tali applicazioni sui sistemi giudiziari» *Carta etica europea,* cit., Punto 160, p. 41.

[89] Nella Appendice II, che non si riporta nel corpo del testo per ragioni di sintesi, la CEPEJ suddivide i potenziali campi di applicazione dell'intelligenza artificiale nei sistemi giudiziari come segue: annovera tra gli utilizzi che devono essere incoraggiati la valorizzazione del patrimonio giurisprudenziale, l'accesso al diritto e la creazione di nuovi strumenti strategici; tra i possibili utilizzi che richiedono considerevoli precauzioni metodologiche fa rientrare l'aiuto nella redazione di tabelle relative ad alcune controversie di carattere civile, il supporto a misure di risoluzione alternativa delle controversie in materia civile, la risoluzione della controversia online e l'utilizzo di algoritmi nelle indagini penali al fine di individuare i luoghi in cui sono commessi reati; tra gli utilizzi da esaminare al termine di supplementari studi scientifici annovera la profilazione dei magistrati e l'anticipazione delle decisioni dei tribunali; da ultimo fa rientrare tra gli utilizzi da esaminare con le più estreme riserve l'utilizzo di algoritmi in materia penale al fine di profilare le persone e la creazione di una norma basata sull'insieme delle decisioni.

[90] *Carta etica europea,* cit., Punto 162, p. 41.

[91] A proposito dell'"incontro" tra scienze dure e scienze sociali G. UBERTIS, *Intelligenza artificiale, giustizia penale, controllo umano significativo*, in *Sistema Penale*, 11.11.2020, p. 7, il quale richiama a sua volta S. QUATTROCOLO, *Intelligenza artificiale e giustizia,* cit., p. 4. Sono già state avviate le sperimentazioni per la creazione di un "meccanismo di certificazione" dei sistemi di intelligenza artificiale destinati ad essere impiegati nelle aule di giustizia al fine di accertarne la conformità ai principi stabiliti dalla su esaminata Carta Etica, P. SEVERINO, *Intelligenza artificiale e diritto penale,* in U. RUFFOLO, *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020, p. 532. Si dedica all'analisi della Carta etica anche G. CORASANITI, *Intelligenza artificiale e diritto: il nuovo ruolo del giurista*, in U. RUFFOLO, *Intelligenza artificiale. Il diritto, i diritti, l'etica*, cit., pp. 404-405, il quale, a proposito della creazione di un *certification mechanism* evidenzia come venga «inoltre incoraggiato un sistema di *certificazione uniforme*, oggetto di *revisione continua*», corsivo nostro.

L'esaminando Piano coordinato chiarisce, in apertura, che l'Unione Europea ha come obiettivo, da un lato, la realizzazione di un'intelligenza artificiale affidabile e ispirata ai valori della Carta dei diritti fondamentali e, dall'altro, diventare leader mondiale nel settore, sviluppando sistemi dotati di IA secondo un approccio etico, sicuro e, soprattutto, antropocentrico.

Tale Piano incoraggia gli Stati membri a realizzare strategie e progetti nel settore dell'intelligenza artificiale – ciascuno sulla scorta delle proprie caratteristiche nazionali – da condividere con gli altri Stati e con la Commissione. Il "gruppo degli Stati membri sulla digitalizzazione dell'industria europea e sull'IA" avrà il compito di guidare le discussioni sul tema tra Commissione e Alte Parti le quali, a loro volta, dovranno individuare parametri di riferimento per effettuare efficaci valutazioni d'impatto in materia di intelligenza artificiale[92].

L'Unione europea si rende conto che, per diventare un punto di riferimento nel settore in esame, è necessario che la Commissione e gli Stati membri indirizzino i propri investimenti in una direzione condivisa, cooperando nella realizzazione di una comune agenda di ricerca e servendosi del supporto di associazioni del settore come l'EurAI[93] e della collaborazione tra l'ambito industriale e quello accademico[94].

Nel riconoscere l'importanza di consolidare il settore scientifico, la Commissione evidenzia la necessità di investire nell'ambito delle possibili applicazioni dell'intelligenza artificiale a fini di sicurezza «da un lato per la prevenzione dell'utilizzo illecito delle tecnologie di IA da parte di soggetti malintenzionati per attività criminali o terroristiche, e dall'altro per la diffusione di strumenti e soluzioni di IA a sostegno delle autorità di contrasto per una migliore prevenzione, individuazione e investigazione di attività criminali e terrorismo»[95].

Punti cardine della c.d. IA "made in Europe" saranno *l'etica fin dalla progettazione* – dunque basata sul rispetto dei principi etici e giuridici sanciti dal GDPR – e la *sicurezza fin dalla progettazione*, al fine di tener conto, *ab origine*, della protezione delle potenziali vittime nonché della cybersicurezza[96].

La Commissione si assume il compito di potenziare le strutture di prova del settore[97] e i già citati poli d'innovazione digitale[98], garantendone un'omogenea

---

[92] *Allegato della Comunicazione della Commissione al Parlamento europeo, al Consiglio europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni. Piano coordinato sull'intelligenza artificiale*, COM(2018) 795, Bruxelles, 7.12.2018, pp. 5-6.

[93] https://www.eurai.org/.

[94] Considerando che l'intelligenza artificiale è già oggetto di diversi partenariati pubblico-privati, specie nel campo della robotica e dei big data, l'Europa mira altresì ad individuare la futura generazione di IA per potersi concentrare su di essa mettendo a disposizione investimenti adeguati per start-up e imprese, anche servendosi di strumenti già esistenti quali il Fondo europeo per gli investimenti strategici e dell'ausilio del Consiglio europeo per l'innovazione (CEI). *Piano coordinato sull'intelligenza artificiale,* cit., p. 7.

[95] *Piano coordinato sull'intelligenza artificiale,* cit., p. 8.

[96] *Piano coordinato sull'intelligenza artificiale,* cit., pp. 8-9.

[97] «Esempi di strutture di prova di questo tipo comprendono quelle utilizzate per le prove transfrontaliere di guida autonoma e connessa, siti di prova per le spedizioni autonome e la creazione di spazi di dati. La Commissione e gli Stati membri identificheranno l'esigenza di istituire nuove strutture di prova su larga scala per le tecnologie di IA più recenti in settori chiave quali mobilità, assistenza sanitaria, attività produttive, agroalimentare o sicurezza. Queste strutture di prova possono includere spazi di sperimentazione normativa (ovvero aree dove la

distribuzione a livello geografico al fine di realizzare una sinergia dei centri di eccellenza in materia di intelligenza artificiale, idonea a rendere l'Europa competitiva sul mercato globale.

Oltre a porsi l'obiettivo di individuare e finanziare i centri di eccellenza nazionali in materia di IA, la Commissione si propone – nell'ambito del programma "Europa digitale" – di istituire siti di prova e sperimentazione nel settore dell'intelligenza artificiale di livello mondiale, così accelerando lo sviluppo dell'IA medesima.

Nel prosieguo la Commissione si dice consapevole del fatto che lo sviluppo degli strumenti dotati di intelligenza artificiale sia strettamente connesso alla disponibilità di considerevoli quantità di dati (sicuri, solidi e di qualità), pertanto propone di agevolare la creazione di uno "spazio comune europeo dei dati"[99] e di promuovere l'accessibilità, l'interoperabilità e il riutilizzo di siffatti dati nei settori di pertinenza dell'IA.

Per realizzare tale scopo agli Stati spetta il compito di individuare, di concerto con la Commissione, set di dati pubblici utili all'addestramento degli strumenti dotati di intelligenza artificiale, nonché di sviluppare il c.d. "cloud europeo per la scienza aperta" e di puntare su forme di tecnologia blockchain per garantire un accesso sicuro ai suddetti dati, specie nei settori della sanità[100], delle geo-informazioni[101], del linguaggio[102] e dell'industria[103].

---

regolamentazione è limitata o favorevole all'esecuzione di prove su nuovi prodotti e servizi) in aree selezionate dove la legge consente alle autorità di regolamentazione sufficiente libertà di azione, attenuando determinati requisiti giuridici e normativi per la durata dello spazio di sperimentazione» *Piano coordinato sull'intelligenza artificiale,* cit., p. 9.

[98] «I poli dell'innovazione digitale possono contribuire a garantire che tutte le imprese, piccole o grandi, ad alta tecnologia o no, e il settore pubblico, possano approfittare delle opportunità digitali. Essendo il loro nucleo costituito da università tecniche o istituti di ricerca, i poli dell'innovazione digitale fungono da sportello unico dove le imprese e il settore pubblico possono avere accesso alla tecnologia, alle prove e al supporto tecnico, alla consulenza finanziaria, alle informazioni di mercato e alle opportunità di creare reti. Più in dettaglio, nel campo dell'IA, i poli dell'innovazione digitale possono aiutare le PMI e le pubbliche amministrazioni a individuare i set di dati necessari, sviluppare algoritmi e addestrare l'IA, e possono collegarsi alle strutture di calcolo basandosi sulla piattaforma di "IA on demand". Possono contribuire a formare il personale delle PMI nell'utilizzo di soluzioni di IA e fornire consulenza circa il sostegno finanziario esistente. Costituiscono un collegamento sia ai centri di eccellenza nella ricerca, sia alle strutture di prova disponibili» *Piano coordinato sull'intelligenza artificiale,* cit., p. 10.

[99] «Un'area digitale senza ostacoli, di dimensioni tali da permettere lo sviluppo di nuovi prodotti e servizi basati sui dati. In particolare, i dati generati e detenuti dal settore pubblico sono spesso di qualità molto elevata e costituiscono una risorsa importante per gli innovatori e le imprese in Europa» *Piano coordinato sull'intelligenza artificiale,* cit., p. 15.

[100] Il settore sanitario è probabilmente uno dei più adatti a sfruttare le potenzialità dell'IA consentendo la circolazione delle informazioni riguardanti pazienti, cartelle cliniche, risultati diagnostici e studi clinici. La Commissione si propone di realizzare un collegamento tra archivi genomici ed un registro delle malattie rare, nonché una banca dati comune di immagini medicali anonimizzate e fornite spontaneamente dai pazienti. *Piano coordinato sull'intelligenza artificiale*, cit., pp. 16-17.

[101] La Commissione riporta l'esempio del programma europeo Copernicus, grazie al quale oggi siamo in grado di acquisire un ingente quantitativo di dati e informazioni risultanti dall'osservazione satellitare della Terra secondo una politica di dati aperti nota anche come *Data and Information Access Services* (DIAS). «Sulla base di ciò, la Commissione propone di sviluppare e adottare capacità di IA utilizzando i dati e le infrastrutture di Copernicus per promuovere servizi basati sulla geolocalizzazione per il clima, l'agricoltura, la qualità dell'aria, le emissioni, l'ambiente marino, la gestione delle risorse idriche, la sicurezza e il monitoraggio delle

Il Piano coordinato in esame si concentra, nel prosieguo, sulla necessità di strutturare un quadro normativo[104] etico, basato sul rispetto dei diritti e delle libertà fondamentali e, pertanto, in grado di generare fiducia da parte dei cittadini nei confronti dell'IA. In particolare l'Europa mira a fare della componente etica il *quid pluris* del suo impegno nel campo dell'intelligenza artificiale a livello mondiale. Per far ciò la Commissione aveva già anticipato nella Comunicazione "L'intelligenza artificiale per l'Europa"[105] l'attribuzione dell'incarico ad un gruppo di esperti nel campo dell'intelligenza artificiale per la creazione di quelli che oggi conosciamo come gli "Orientamenti etici per un'IA affidabile" (che saranno fatti oggetto di trattazione più avanti) fondati sul principio dell'"etica fin dalla progettazione".

Per far ciò il Piano coordinato in esame incoraggia, altresì, gli scambi di buone pratiche e il lavoro sinergico di competenze strategiche tecniche[106] e non propriamente tecniche, quali diritto e scienze umane[107].

La Commissione si propone l'obiettivo di dar vita ad «un quadro di sicurezza e responsabilità adeguato che garantisca un livello elevato di sicurezza e

migrazioni, nonché la scienza dei cittadini» *Piano coordinato sull'intelligenza artificiale,* cit., p. 17.

[102] Ci si riferisce ai servizi di traduzione automatica e di elaborazione del linguaggio naturale che sfruttano l'intelligenza artificiale e per la quale la Commissione si propone di effettuare nuovi investimenti per integrare risorse linguistiche meno rappresentate in rete. Ad esempio «Nel 2019 la Commissione intende offrire eTranslation, il servizio di traduzione automatica che sfrutta l'IA, sviluppato nell'ambito del meccanismo per collegare l'Europa, alle pubbliche amministrazioni degli Stati membri. Le proposte della Commissione per i programmi Orizzonte Europa ed Europa digitale prevedono investimenti orientati allo sviluppo ulteriore dei servizi e degli strumenti di elaborazione del linguaggio naturale per favorire il multilinguismo nel settore pubblico» *Piano coordinato sull'intelligenza artificiale,* cit., p. 22.

[103] L'Unione si propone di realizzare piattaforme di dati industriali per realizzare una condivisione di dati sicura e la creazione di uno spazio comune europeo dei dati al fine di renderli maggiormente accessibili. *Piano coordinato sull'intelligenza artificiale,* cit., p. 18.

[104] «È importante che la legislazione offra un quadro adeguato per l'innovazione guidata dall'IA e per la diffusione delle soluzioni di IA, affrontando nel contempo i possibili rischi che emergono dall'utilizzo e dall'interazione con la tecnologia, compresi i problemi di cibersicurezza. Ciò significa garantire la "cibersicurezza" in termini di prevenzione degli abusi (ad esempio l'hacking o la manipolazione di algoritmi di IA o dei dati elaborati da un algoritmo di IA) e l'adozione di meccanismi che garantiscano la sicurezza dei consumatori e un processo di ricorso efficace che tuteli le vittime in caso di danni e faciliti le indagini qualora il sistema di IA venga compromesso» *Piano coordinato sull'intelligenza artificiale,* cit., p. 19.

[105] Cfr. Cap. I, Sez. I, Par. 2.

[106] Tali competenze sono definite con l'acronimo STEM (Science, Technology, Engineering, Math - scienze, tecnologia, ingegneria, matematica) o TIC (tecnologie dell'informazione e della comunicazione). *Piano coordinato sull'intelligenza artificiale,* cit., p. 12.

[107] L'Unione Europea si rende conto che il settore dell'intelligenza artificiale necessita di specifiche competenze e che sia il mondo del lavoro sia quello dell'istruzione specialistica non sono attualmente in grado di soddisfarle adeguatamente. Per tale ragione ci si propone di istituire programmi di qualificazione professionale per sopperire a tali carenze organiche – anche attraverso la creazione dei c.d. MOOC (*Massive Open Online Courses*) – e di studiare azioni volte a trattenere (o attirare) in Europa i talentuosi studiosi del settore, concentrandosi anche su competenze non propriamente tecniche al fine di realizzare un'IA antropocentrica fondata sull'etica. Un'agevolazione di tale processo potrebbe realizzarsi mediante il reciproco riconoscimento delle certificazioni (quali diplomi d'istruzione superiore o di periodi di studio all'estero) e con lo sfruttamento delle potenzialità della Carta blu, ossia un permesso di lavoro volto ad attirare cittadini di paesi terzi particolarmente qualificati consentendogli di lavorare e vivere nell'UE *Piano coordinato sull'intelligenza artificiale,* cit., pp. 12-13.

meccanismi efficaci di ricorso per le vittime in caso di danni»[108], entrando così a pieno titolo nel settore di competenza dei giuristi, ai quali spetterà il compito di collaborare allo sviluppo della materia *de qua* per garantire la realizzazione di un quadro normativo completo, garantista ed in grado di generare fiducia nell'IA.

A tal proposito la Commissione si interroga se la legislazione attualmente esistente sia in grado di fronteggiare compiutamente le problematiche poste dall'intelligenza artificiale (quali la responsabilità del processo decisionale dello strumento dotato di IA o le questioni sulla relativa proprietà intellettuale) e se sia opportuno creare spazi di sperimentazione normativa in grado di favorire la ricerca scientifica per specifiche applicazioni dell'intelligenza artificiale[109].

Sul finire del Piano oggetto di disamina la Commissione si propone di migliorare i servizi pubblici (quali sanità, trasporti, sicurezza e giustizia) per il tramite dell'adozione di strumenti dotati di intelligenza artificiale, esortando allo scambio di buone pratiche e all'individuazione di ambiti per l'appalto di soluzioni intelligenti volte a rendere la pubblica amministrazione maggiormente efficiente.

In chiusura la Commissione propone un approccio di cooperazione internazionale in materia di intelligenza artificiale, puntando ad uno sviluppo sinergico di obiettivi comuni e standard internazionali in modo da facilitare la diffusione dell'IA e da ingenerare fiducia nei suoi confronti. La Commissione evidenzia, infine, l'importanza che l'Unione faccia fronte comune, promuovendo orientamenti etici ed un proficuo dialogo con Paesi terzi e privati interessati disposti a condividere i valori europei[110].

6. La Risoluzione del Parlamento europeo su una politica industriale europea globale in materia di robotica e intelligenza artificiale.

La nostra carrellata sugli interventi dell'Unione Europea nel settore dell'intelligenza artificiale prosegue con l'emanazione, da parte del Parlamento europeo, della Risoluzione concernente la politica industriale europea in materia di robotica ed intelligenza artificiale, datata 12 febbraio 2019[111].

In linea con quanto sancito dagli interventi europei precedentemente esaminati, il succitato Parlamento evidenzia la necessità di sviluppare un approccio europeo in materia dell'IA, fondato sul rispetto della Carta dei diritti fondamentali, nonché sulla protezione dei dati, della vita privata, della sicurezza, della dignità, dell'autonomia e dell'autodeterminazione degli individui[112].

---

[108] *Piano coordinato sull'intelligenza artificiale,* cit., p. 19. Il tal sede la Commissione propone altresì l'utilizzo dei c.d. patti per l'innovazione, i quali «possono fungere da strumenti interni alla legislazione esistente per valutare gli ostacoli normativi in termini di sviluppo e diffusione dell'IA. I patti per l'innovazione sono accordi di cooperazione volontari tra l'UE, gli innovatori e le autorità nazionali, regionali e locali. L'obiettivo di un patto per l'innovazione è quello di raggiungere un livello di comprensione approfondito del modo in cui una norma o un regolamento UE funzionano in pratica. Il patto farà emergere eventuali ostacoli posti all'innovazione da una norma o un regolamento, consentendo eventuali azioni ulteriori» *Piano coordinato sull'intelligenza artificiale,* cit., pp. 19-20.

[109] *Piano coordinato sull'intelligenza artificiale,* cit., pp. 20-21.

[110] *Piano coordinato sull'intelligenza artificiale,* cit., pp. 22-23.

[111] Si concentra su tale Risoluzione anche F. CAROCCIA, *Soggettività giuridica dei robot?,* cit., p. 216.

[112] *Risoluzione del Parlamento europeo del 12 febbraio 2019 su una politica industriale europea globale in materia di robotica e intelligenza artificiale,* C 449/37, Considerando L e M.

In tal sede viene altresì richiamata la necessità di garantire un flusso di dati (necessario per alimentare i sistemi di intelligenza artificiale) libero e sicuro, volto ad evitare il riprodursi delle discriminazioni endemicamente presenti nella società[113].

Il Parlamento chiarisce sin da subito che «la tendenza all'automazione esige che i soggetti coinvolti nello sviluppo e nella commercializzazione di applicazioni di intelligenza artificiale integrino gli aspetti relativi alla sicurezza e all'etica fin dal principio, riconoscendo pertanto che devono essere pronti ad accettare la *responsabilità giuridica* della qualità della tecnologia da loro prodotta»[114] così portando l'attenzione su un tema particolarmente delicato nella materia in esame (e che invero costituisce il *leitmotiv* del presente lavoro), ossia l'individuazione del soggetto responsabile per i danni causati dal sistema dotato di intelligenza artificiale.

In questa fase iniziale il Parlamento richiama la necessità di adottare un approccio antropocentrico nello sviluppo dei sistemi intelligenti, senza però dimenticare le problematiche di natura etica e giuridica connesse ad una sempre maggiore automazione dei sistemi di IA, nonché il principio di responsabilità secondo il quale è l'uomo a servirsi della macchina[115].

Fornendo maggiori spunti di riflessione alla materia penalistica che qui ci occupa, il Parlamento evidenzia come un uso doloso o negligente dell'intelligenza artificiale potrebbe costituire una minaccia per la sicurezza pubblica, per la democrazia e per i diritti fondamentali, specie se indirizzato a realizzare campagne di disinformazione volte a minare l'autodeterminazione del singolo[116].

Il succitato Parlamento rivolge poi una specifica indicazione alla Commissione, invitandola a «proporre un quadro che *penalizzi* le pratiche di manipolazione della percezione quando i contenuti personalizzati o i "news feed" conducono a sentimenti negativi e alla distorsione della realtà con possibili conseguenze negative»[117]. Già da una preliminare lettura di siffatto invito appare però difficile immaginare come poter strutturare una normativa di tal fatta, a meno di non voler dar vita ad una fattispecie talmente generica e vaga da porsi in contrasto con i capisaldi penalistici della tassatività[118] e dell'offensività[119].

---

[113] *Politica industriale europea*, cit., Considerando P, T e U.

[114] *Politica industriale europea*, cit., Considerando S, corsivo nostro.

[115] *Politica industriale europea*, cit., Considerando V, AJ e AK. Il Parlamento si concentra nel prosieguo (ai Punti 1 e 7) sulla futura struttura di una società fondata sull'intelligenza artificiale, sottolineando come questa nuova rivoluzione tecnologica porterà ad un incremento della produttività, ad una rimodulazione di svariati impieghi, nonché alla sostituzione degli esseri umani nello svolgimento di lavori usuranti o pericolosi e alla necessità di predisporre una vera e propria alfabetizzazione digitale.

[116] *Politica industriale europea*, cit., Punto 9.

[117] *Politica industriale europea*, cit., Punto 10, corsivo nostro.

[118] G. MARINUCCI, E. DOLCINI, G.L. GATTA, *Manuale di diritto penale. Parte generale*, Milano, 2022, pp. 89 ss.; G. DE VERO, *Corso di diritto penale. Parte generale*, Torino, 2020, pp. 99 ss.; F. MANTOVANI, *Diritto Penale. Parte Generale*, Padova, 2020, pp. 66 ss.; G. FIANDACA, E. MUSCO, *Diritto Penale*, cit., pp. 88 ss.; L. RISICATO, *Gli elementi normativi della fattispecie. Profili generali e problemi applicativi*, Milano, 2004; F. PALAZZO, *Il principio di determinatezza nel diritto penale: la fattispecie*, Padova, 1979; M RONCO, *Il principio di tipicità della fattispecie penale nell'ordinamento vigente,* Torino, 1979.

[119] G. MARINUCCI, E. DOLCINI, G.L. GATTA, *Manuale di diritto penale*, cit., pp. 10 ss.; F. MANTOVANI, *Diritto Penale*, cit., pp. 197; G. FIANDACA, E. MUSCO, *Diritto Penale*, cit., p. 164; G. DE VERO, *Corso di diritto penale.*, cit., pp. 114 ss.; G. MANIACI, Harm principle *e* offence

Nel rammentare il ruolo cardine dell'Europa nel settore dell'intelligenza artificiale[120] il Parlamento sottolinea come la ricerca in tale ambito debba svolgersi non solo dal punto di vista tecnologico ma anche da quello sociale ed etico, dedicando maggior attenzione alle questioni concernenti la responsabilità (seppur solo civile) per i danni cagionati da sistemi dotati di IA ed al rispetto del principio di precauzione e dei diritti fondamentali[121].

Viene in tal sede altresì sottolineata la necessità di predisporre un adeguato meccanismo di valutazione e gestione dei rischi – proprio in considerazione dell'ineliminabile presenza di una componente di incertezza nello sviluppo dei sistemi di intelligenza artificiale[122] – accogliendo con favore l'istituzione di spazi di sperimentazione normativa volti a rendere tali tecnologie sicure fin da subito[123].

Il Parlamento europeo rimarca la necessità di servirsi di dati veritieri e qualitativamente elevati al fine di realizzare un proficuo sviluppo dell'IA, ciò anche allo scopo di incrementare la competitività sul mercato delle aziende europee che se ne occupano, di creare standard di sicurezza e affidabilità elevati e di evitare previsioni inadeguate o discriminatorie[124].

---

principle *secondo un'etica liberale*, in *Criminalia*, 30.1.2019; M. LAVACCHINI, *La legittimazione dell'intervento penale tra principio di offensività e principio del danno (*harm principle*)*, in *disCrimen*, 2.8.2019; M. DONINI, *Il principio di offensività. Dalla penalistica italiana ai programmi europei*, in *Dir. pen. cont. - Riv. Trim.*, 4/2013, pp. 4 ss.; V. MANES, *I recenti tracciati della giurisprudenza costituzionale in materia di offensività e ragionevolezza*, in *Dir. pen. cont.* - Riv. Trim. 1/2012, pp. 99 ss.; ID., *Il principio di offensività nel diritto penale. Canone di politica criminale, criterio ermeneutico, parametro di ragionevolezza*, Torino, 2005; F. PALAZZO, *Offensività e ragionevolezza nel controllo di costituzionalità sul contenuto delle leggi penali*, in *Riv. it. dir. proc. pen*., 1998, pp. 350 ss.

[120] «Considerando che, su scala globale, circa un quarto di tutti i robot industriali e metà di tutti i robot dedicati ai servizi professionali sono prodotti da società europee e che pertanto l'UE dispone già di importanti risorse su cui dovrebbe basare la sua politica industriale europea» *Politica industriale europea globale in materia di robotica e intelligenza artificiale*, cit., Considerando E.

[121] *Politica industriale europea globale in materia di robotica e intelligenza artificiale*, cit., Punti 18 e 19. In particolare quest'ultimo punto specifica che «tutti i soggetti coinvolti nello sviluppo, nell'attuazione, nella diffusione e nell'utilizzo dell'intelligenza artificiale dovrebbero tenere in considerazione e rispettare la dignità umana, nonché l'autodeterminazione e il benessere, sia fisico che psicologico, dell'individuo e della società in generale, anticipare le potenziali conseguenze sulla sicurezza e prendere le dovute precauzioni in proporzione al livello di protezione, compresa la tempestiva divulgazione di fattori che potrebbero mettere a rischio i cittadini o l'ambiente».

[122] *Politica industriale europea*, cit., Punto 24. Ai Punti 25 ss. il Parlamento evidenzia la necessità di alimentare gli investimenti nel settore (eccezion fatta per il finanziamento dei sistemi d'arma autonomi e per le attività di ricerca concernenti la coscienza dell'intelligenza artificiale), anche aprendosi al dialogo con Paesi terzi, al fine di rendere l'Europa competitiva nel settore. Il Parlamento raccomanda inoltre il rispetto della normativa in materia di proprietà intellettuale al fine di garantire che le ricerche e le scoperte effettuate in Europa rimangano all'interno dell'Unione.

[123] Il Parlamento evidenzia, inoltre, l'importanza di condurre un'adeguata campagna informativa per sensibilizzare la comunità sui potenziali vantaggi di tali tecnologie, garantendone la trasparenza e la sicurezza, anche al fine di incrementare l'innovazione nel settore. *Politica industriale europea*, cit., Punti 30, 31 e 32.

[124] *Politica industriale europea*, cit., Punti 38, 39 e 42. Ai Punti 47, 48 e 51 il Parlamento si dice consapevole della ritrosia di molti cittadini a condividere i propri dati, pertanto accoglie con favore i regolamenti che si occupano della tutela e della libera circolazione di questi ultimi, evidenziando altresì l'importanza dello sviluppo di sistemi di *cloud computing* necessari per l'elaborazione dei megadati.

Spostandosi dal tema dei dati a quello della politica industriale, il succitato Parlamento chiarisce l'importanza di sviluppare un quadro normativo flessibile – in grado di adeguarsi al progresso delle nuove tecnologie dotate di intelligenza artificiale – nonché di coinvolgere in siffatta evoluzione il settore pubblico, rappresentando quest'ultimo una inesauribile fonte di dati[125].

Nel prosieguo della presente Risoluzione viene evidenziato come l'intelligenza artificiale trovi uno dei suoi più proficui campi d'elezione nell'ambito sanitario (basti pensare, a titolo esemplificativo, alla chirurgia robotica, alla gestione di cartelle cliniche o alla medicina di precisione)[126]. Essa può incidere, in particolar modo, nei settori della ricerca, della diagnosi e della prevenzione, garantendo un'assistenza personalizzata (anche domiciliare)[127] e più efficiente[128]. In tal sede il Parlamento europeo invita la Commissione a «chiarire il quadro relativo alla ripartizione della responsabilità civile tra l'utilizzatore (medico/professionista), il produttore della soluzione tecnologica e la struttura sanitaria che propone il trattamento», così ponendo l'accento sull'importanza del tema della responsabilità giuridica per i danni cagionati dall'IA ed evidenziando, altresì, la necessità di evitare che gli operatori sanitari – i quali si servono degli strumenti di intelligenza artificiale – assecondino acriticamente le soluzioni diagnostiche proposte dal sistema intelligente al fine di sottrarsi a qualsivoglia conseguenza dal punto di vista giuridico[129].

Per quanto il Parlamento circoscriva il proprio invito esclusivamente alla ripartizione delle responsabilità dal punto di vista civilistico, è possibile ritenere che una suddivisione delle responsabilità in tale settore possa giovare anche in ambito penale, fornendo spunti di riflessione e di confronto per poter strutturare il relativo riparto anche nella materia di nostro interesse.

Il Parlamento europeo evidenzia nel prosieguo le potenzialità applicative dell'intelligenza artificiale nel settore energetico[130] (sfruttando una manutenzione predittiva e maggiormente efficiente), in quello agricolo[131], in quello della cybersicurezza[132] ed in quello dei trasporti[133]. Con particolare riferimento a tale ultimo ambito viene accolta con favore la prospettiva dell'introduzione di veicoli autonomi in grado di ridurre gli errori umani e, conseguentemente, il verificarsi di incidenti stradali. Anche in tale settore ritorna il problema della responsabilità giuridica, affrontato forse un po' troppo sbrigativamente dal Parlamento europeo

---

[125] *Politica industriale europea*, cit., Punti 54, 63 e 67.
[126] *Politica industriale europea*, cit., Considerando AF e AH.
[127] *Politica industriale europea*, cit., Punto 82.
[128] *Politica industriale europea*, cit., Punto 72.
[129] *Politica industriale europea*, cit., Punto 77.
[130] *Politica industriale europea*, cit., Punti 84 ss.
[131] Gli strumenti di intelligenza artificiale potrebbero essere adoperati per individuare e affrontare i problemi connessi alla sicurezza della filiera alimentare, migliorando la gestione delle risorse e dei fattori produttivi, riducendo gli sprechi e prevedendo condizioni atmosferiche avverse. *Politica industriale europea*, cit., Punti 93 ss.
[132] *Politica industriale europea* cit., Punti 100 ss. Il Parlamento europeo pone l'accento sull'importanza della sicurezza informatica nel settore dell'intelligenza artificiale, evidenziando che l'Unione dovrebbe acquisire una propria indipendenza tecnologica sul punto, non priva di un costante controllo umano: «sebbene l'intelligenza artificiale consentirà di migliorare la rilevazione delle minacce, è assolutamente necessaria un'interpretazione umana di dette minacce per comprendere se siano reali oppure no» Punto 106.
[133] *Politica industriale europea*, cit., Punti 89 ss.

nei seguenti termini: «l'ampia diffusione di veicoli autonomi in futuro pone rischi per la riservatezza dei dati e i guasti tecnici e sposterà la responsabilità dal conducente al fabbricante, il che comporterà la necessità per le compagnie assicurative di modificare il modo in cui è incorporato il rischio nella sottoscrizione delle loro polizze»[134].

Nel prosieguo il Parlamento rivolge un ulteriore invito alla Commissione, chiedendo di sostenere un quadro normativo che incentivi lo sviluppo dell'intelligenza artificiale, anche riesaminando la normativa attualmente in vigore al fine di porla in continuità con i valori fondanti dell'Unione[135].

Viene evidenziata nel prosieguo l'importanza di lavorare sugli aspetti etici dell'IA e di creare un unico insieme di regole da applicare all'interno del mercato dell'UE, agevolando il reciproco riconoscimento transfrontaliero dei beni intelligenti[136]. In tal sede viene inoltre chiarito che il responsabile di qualsivoglia processo decisionale in astratto attribuibile all'IA deve essere sempre individuato nel soggetto umano, ritenendo opportuno riflettere sull'opportunità di un costante controllo da parte di un professionista competente[137].

Il Parlamento richiama l'importanza di garantire la sicurezza e la riservatezza dei dati utilizzati per le interazioni tra umani e sistemi intelligenti, nonché di rispettare il diritto alla vita privata, il trattamento equo, il giusto processo, la libertà di espressione e di informazione[138]. Proprio con riferimento al diritto all'informazione si evidenzia la necessità di stabilire procedure chiare per garantire che i soggetti coinvolti prestino un consenso (realmente) informato al trattamento dei loro dati personali e che le eventuali richieste di cancellazione di siffatti dati da parte degli interessati possano sempre essere assecondate[139].

Nel prosieguo della Risoluzione in esame il Parlamento si concentra sui possibili profili di responsabilità degli operatori dell'IA, evidenziando come questi ultimi dovrebbero essere considerati, per l'appunto, responsabili dei possibili impatti dei sistemi intelligenti sulla società, non senza evidenziare che tali prodotti *intelligenti* dovrebbero essere sottoposti a controlli di sicurezza ed al rispetto delle norme poste a tutela dei consumatori che affrontino, tra le altre cose, il problema del rischio degli incidenti causati dall'IA[140].

Viene richiamata l'attenzione sulla necessità che le applicazioni dell'intelligenza artificiale rispettino i principi etici[141] e la pertinente legislazione, promuovendo la collaborazione tra il settore pubblico, privato e accademico, nonché lo scambio di buone pratiche e la creazione di un sistema volto a garantire

[134] *Politica industriale europea*, cit., Punto 91.
[135] Il Parlamento favorisce altresì l'adozione di strumenti di democrazia diretta quali piattaforme informative che consentano ai cittadini di poter fornire un utile contributo in materia. *Politica industriale europea*, cit., Punto 114.
[136] *Politica industriale europea*, cit., Punti 118 e 120.
[137] *Politica industriale europea*, cit., Punto 123.
[138] *Politica industriale europea*, cit., Punti 125, 126 e 128.
[139] *Politica industriale europea*, cit., Punto 129.
[140] *Politica industriale europea*, cit., Punti 133 e 135.
[141] Il Parlamento nel prosieguo chiarisce la necessità di «mettere a punto norme etiche volte ad assicurare uno sviluppo dell'intelligenza artificiale incentrato sull'uomo, la responsabilità e la trasparenza dei sistemi decisionali algoritmici, chiare norme in merito alla responsabilità ed equità» *Politica industriale europea*, cit., Punto 143.

che i dati dei soggetti interessati siano utilizzati entro le finalità previste e per le quali hanno prestato il consenso[142].

Nella prospettiva di dar vita ad un'intelligenza artificiale che possa definirsi etica (obiettivo principe dell'Europa nel settore) appare di fondamentale importanza basare gli orientamenti etici più volte richiamati sui principi di beneficenza, autonomia, giustizia[143], trasparenza e spiegabilità[144], istituendo relativi organi di controllo.

Il Parlamento evidenzia inoltre la necessità che le persone sappiano (ed abbiano conseguentemente il diritto di opporsi o ricorrere) quando una decisione che li riguarda viene assunta da un sistema di IA, ben potendo da ciò derivare un serio rischio per i diritti e le libertà dei soggetti coinvolti in un dato processo decisionale[145].

Si riconosce in tal sede la circostanza che gli algoritmi in esame possano risultare opachi: per tale ragione si richiama la necessità di una maggiore trasparenza e responsabilità nel trattamento dei dati personali e del processo decisionale dello strumento dotato di IA, sempre consentendo ai soggetti coinvolti di ottenere l'intervento umano in sostituzione della decisione automatizzata[146].

Trasparenza, spiegabilità e responsabilità sono concetti da dover rispettare durante l'intero ciclo di vita dell'IA, nonché convergenti verso la necessità che il meccanismo di funzionamento del sistema intelligente sia comprensibile all'uomo al fine di poter generare in esso fiducia, in particolar modo attraverso lo sviluppo del requisito della "responsabilità". In tal sede si chiarisce inoltre l'importanza di coinvolgere competenze differenziate al fine di scongiurare le distorsioni cui gli sviluppatori informatici potrebbero – anche involontariamente – dar vita, ad esempio servendosi di dati di scarsa qualità o utilizzando in modo improprio dati qualitativamente elevati[147].

In conclusione, il Parlamento valuta positivamente le iniziative nazionali adottate dai vari Stati membri, ritenendo opportuno che questi ultimi collaborino con la Commissione per realizzare un sistema di norme coerente, idoneo a garantire uniformemente il medesimo livello di sicurezza, nonché il rispetto dei principi etici summenzionati. Un quadro europeo armonizzato contribuirebbe a

---

[142] *Politica industriale europea*, cit., Punti 138, 140 e 141.

[143] In tal sede il Parlamento richiama altresì i più volte citati «principi sanciti all'articolo 2 del trattato sull'Unione europea e nella Carta dei diritti fondamentali dell'Unione europea — quali la dignità umana, l'uguaglianza, la giustizia e l'equità, la non discriminazione, il consenso informato, la vita privata e familiare e la protezione dei dati, così come sugli altri principi e valori alla base del diritto dell'Unione come la non stigmatizzazione, la trasparenza, l'autonomia, la responsabilità individuale e sociale» *Politica industriale europea*, cit., Punto 147.

[144] *Politica industriale europea*, cit., Punto 149. Il concetto della spiegabilità viene ripreso al Punto 151 nel quale il Parlamento europeo «sottolinea la difficoltà e la complessità nel prevedere i comportamenti futuri di molti sistemi complessi di intelligenza artificiale e i comportamenti emergenti dei sistemi di intelligenza artificiale che interagiscono fra loro; invita la Commissione a valutare se esista la necessità di regolamenti specifici concernenti il *processo decisionale* basato sull'intelligenza artificiale», corsivo nostro.

[145] *Politica industriale europea*, cit., Punto 153. Al Punto 154 il Parlamento raccomanda che i sistemi dotati di intelligenza artificiale programmati per assumere decisioni siano sottoposti a una preventiva analisi d'impatto algoritmica (AIA), a meno che non si tratti di decisioni che non incidono sulla vita di alcuno.

[146] *Politica industriale europea*, cit., Punti 157 e 158.

[147] *Politica industriale europea*, cit., Punti 159 ss.

generare fiducia nell'IA e a sostenere il mercato unico, ferma restando la necessità di cooperare con partner internazionali per affrontare sfide comuni ed ottenere benefici condivisi[148].

Il documento appena esaminato si caratterizza, più di altri, per i frequenti riferimenti al tema della responsabilità, fornendo spunti interessanti ma senza mai affrontare esplicitamente il tema della responsabilità penale, nonostante in esso siano disseminati non pochi riferimenti ad essa (basti riportare, a titolo esemplificativo, il riferimento agli incidenti stradali causati da auto a guida autonoma). Sembra pertanto che il Parlamento abbia perso l'occasione di fare chiarezza sul tema del riparto di responsabilità penale, lasciando privi di approfondimento i (non pochi) input penalistici in esso presenti.

7. Gli Orientamenti etici per un'IA affidabile.

Il "Gruppo indipendente di esperti ad alto livello sull'intelligenza artificiale", istituito dalla Commissione europea nel giugno 2018 ha reso pubblici, in data 8 aprile 2019, i più volte citati "Orientamenti etici per un'IA affidabile".

Nel passare in rassegna i più significativi interventi dell'Unione Europea in materia di intelligenza artificiale non si può far a meno di spendere qualche battuta sulle Linee guida in parola. Esse, pur non concentrandosi su aspetti propriamente giuridici, hanno il merito di tracciare una linea programmatica di intervento rivolta a tutti gli operatori del settore e di porre le basi per il futuro sviluppo in termini etici dell'intelligenza artificiale. Esso si propone (e ha effettivamente le potenzialità per farlo) di diventare solida base della materia cui far ricorso per risolvere ogni potenziale dubbio interpretativo, anche (perché no?) in ambito giuridico.

Il presente documento, già in fase introduttiva, evidenzia l'importanza di creare sistemi di intelligenza artificiale antropocentrici (dunque al servizio dell'uomo) ed affidabili, ossia capaci di generare fiducia nella comunità, adottando un approccio etico in grado di rendere l'Europa leader nel settore.

Il succitato gruppo di esperti individua le tre componenti necessarie (ma non sufficienti) per un'IA affidabile, raccomandando che queste ultime siano sempre rispettate durante l'intero ciclo di vita dell'IA:

1. *Legalità*: i sistemi di intelligenza artificiale devono rispettare tutte le norme giuridicamente vincolanti a livello nazionale, europeo ed internazionale. Il documento in esame non si propone l'obiettivo di creare diritti o imporre obblighi, limitandosi in tal sede a raccomandare l'osservanza delle norme generalmente applicabili (quali il diritto primario e derivato dell'UE) e delle norme di settore (ad esempio il regolamento sui dispositivi medici nel settore sanitario);

2. *Eticità*: i meccanismi dotati di intelligenza artificiale devono essere compatibili con principi e valori etici;

---

[148] *Politica industriale europea*, cit., Punti 183 ss. In particolare, al Punto 187, il Parlamento propone la creazione di un'"agenzia europea di regolamentazione per l'intelligenza artificiale e il processo decisionale algoritmico" cui poter affidare diversi compiti, tra cui effettuare una valutazione di rischio in grado di classificare e differenziare i vari tipi di algoritmo ed i relativi campi di applicazione, verificare la violazione dei diritti umani da parte di siffatti sistemi intelligenti e, soprattutto per quanto di nostro interesse, «potenziare l'efficacia del meccanismo responsabilità da illecito quale mezzo per disciplinare la responsabilità dei sistemi algoritmici, fornendo un punto di contatto per i cittadini che non hanno familiarità con i procedimenti giudiziari».

3. *Robustezza*: i sistemi di IA potrebbero causare danni involontari, per questa ragione è importante garantirne la robustezza da un punto di vista tecnico e sociale[149].

Il Capitolo I del documento in esame delinea, alla luce dei tre requisiti succitati, la basi per realizzare un'IA affidabile, etica e fondata sui diritti fondamentali, segnatamente su:

- *Dignità umana*: intesa come valore irriducibile ed irrinunciabile dell'uomo;
- *Libertà individuale*: da interpretare come il controllo della propria libertà di espressione, della propria vita privata e della propria riservatezza;
- *Democrazia, giustizia e Stato di diritto*: i meccanismi dotati di IA devono rispettare le scelte dei consociati e la pluralità dei valori incarnati da una determinata società, nonché i suoi processi democratici;
- *Uguaglianza, non discriminazione e solidarietà*: è importante che i sistemi di IA non generino risultati distorti e non diano vita a discriminazioni ingiustificate, garantendo sempre una particolare tutela nei confronti dei soggetti più vulnerabili;
- *Diritti dei cittadini*: i sistemi dotati di IA possono migliorare i beni e i servizi offerti ai cittadini, tuttavia essi potrebbero altresì comportare effetti negativi su questi ultimi, richiedendo pertanto la loro salvaguardia[150].

Il gruppo di esperti individua poi, alla luce dei diritti fondamentali suelencati, quattro principi etici che costituiscono veri e propri imperativi per gli operatori dell'IA, i quali saranno vincolati ad aderirvi al fine di garantire l'affidabilità dei sistemi intelligenti[151]:

1. *Rispetto dell'autonomia umana*: gli individui che interagiscono con i sistemi di IA – i quali sono progettati per aumentare e integrare le potenzialità umane – devono sempre mantenere una piena autodeterminazione. Nell'ottica di una programmazione antropocentrica[152] l'uomo deve poter conservare sempre ampie opportunità di scelta;

2. *Prevenzione dei danni*: i sistemi di intelligenza artificiale non possono influenzare negativamente gli esseri umani, né causare loro danni o aggravare situazioni già patologiche. Proprio in tale prospettiva si raccomanda che tali meccanismi siano "robusti", per garantire la loro sicurezza e, soprattutto, che non vengano esposti ad usi dolosi;

3. *Equità*: da un punto di vista sostanziale si mira a garantire che lo sviluppo, la diffusione e l'uso dei sistemi intelligenti siano equi, ossia volti ad evitare che

---

[149] *Orientamenti etici per un'IA affidabile*, Bruxelles, 8.4.2019, pp. 7-8.

[150] *Orientamenti etici per un'IA affidabile,* cit., p. 12.

[151] *Orientamenti etici per un'IA affidabile,* cit., pp. 13-14.

[152] Secondo la definizione di "IA antropocentrica" fornita nel Glossario contenuto alla fine del documento in esame «L'approccio antropocentrico all'IA è volto a garantire che i valori umani rivestano un ruolo centrale nelle modalità di sviluppo, distribuzione, utilizzo e monitoraggio dei sistemi di IA, garantendo il rispetto dei diritti fondamentali, tra cui quelli sanciti nei trattati dell'Unione europea e nella Carta dei diritti fondamentali dell'Unione europea, accomunati dal riferimento a un fondamento condiviso radicato nel rispetto della dignità umana, nei quali l'essere umano gode di uno status morale unico e inalienabile. Ciò implica anche il rispetto dell'ambiente naturale e di altri esseri viventi che fanno parte dell'ecosistema umano e un approccio sostenibile che consenta alle generazioni future di prosperare» *Orientamenti etici per un'IA affidabile*, cit., p. 46.

gli individui subiscano distorsioni e discriminazioni. Da un punto di vista procedurale, invece, il rispetto dell'equità comporta la possibilità di contestare le decisioni elaborate dai sistemi intelligenti, garantendo l'identificabilità del reale responsabile della decisione e la comprensibilità del percorso decisionale seguito dalla macchina.

4. *Esplicabilità*: garantire la trasparenza dei sistemi di IA e la possibilità di rendere le loro decisioni comprensibili ai soggetti interessati appare di fondamentale importanza, da un lato, per alimentare la fiducia in questi sistemi e, dall'altro, per rendere le loro determinazioni impugnabili. È bene però notare che le forme più sofisticate di IA spesso non consentono di ottenere siffatta trasparenza a causa del c.d. *black box effect*[153].

Il Gruppo di esperti chiarisce che, per quanto sarebbe auspicabile una coesistenza armoniosa tra i principi summenzionati, non è da escludere che tra di essi possano crearsi tensioni che richiedano una specifica risoluzione caso per caso. Proprio portando l'esempio della polizia predittiva si ipotizza il potenziale contrasto tra prevenzione dei danni (che deriverebbe dall'evitare la consumazione di determinati crimini) ed autonomia umana (connessa alla potenziale violazione del diritto alla riservatezza degli individui). Il Gruppo si dice consapevole della possibilità che gli operatori dell'IA non trovino la soluzione alle summenzionate tensioni servendosi dei principi enunciati – essendo possibile che non si raggiungano compromessi accettabili, specie ove vengano coinvolti diritti non bilanciabili, quali il rispetto della dignità umana –, chiarendo tuttavia che questi ultimi costituiscono prescrizioni astratte utili a indicare la via per raggiungere possibili soluzioni[154].

---

[153] Un'efficace definizione di tale fenomeno è contenuta nel documento redatto dal medesimo gruppo di esperti che ha dato vita agli Orientamenti Etici in esame e mediante il quale si cerca di fornire una definizione compiuta di cosa debba intendersi per "Artificial Intelligence" e di come possano definirsi i meccanismi di funzionamento di quest'ultima. Il *black box effect* può essere così spiegato: «Alcune tecniche di apprendimento automatico, sebbene di grande successo dal punto di vista della precisione, sono molto opachi in ordine alla comprensione di come assumono le decisioni. La nozione di *black-box AI* si riferisce a tali scenari, dove non è possibile risalire al motivo di determinate decisioni. La spiegabilità è una proprietà di quei sistemi di IA che invece possono fornire una forma di spiegazione per le proprie azioni» *A definition of AI: main capabilities and disciplines. Definition developed for the purpose of the AI HLEG's deliverables*, Bruxelles, 8.4.2019, p. 5. Questo stesso documento integrativo fornisce una definizione aggiornata di Intelligenza Artificiale rispetto a quella fornita dalla Comunicazione della Commissione intitolata "Intelligenza Artificiale per l'Europa" e contenuta nella nota 37 del presente capitolo: «I sistemi di intelligenza artificiale (IA) sono sistemi software (ed eventualmente hardware) progettati dall'uomo che, dato un obiettivo complesso, agiscono nella dimensione fisica o digitale percependo il proprio ambiente attraverso l'acquisizione di dati, interpretando i dati strutturati o non strutturati raccolti, ragionando sulla conoscenza o elaborando le informazioni derivate da questi dati e decidendo le migliori azioni da intraprendere per raggiungere l'obiettivo dato. I sistemi di IA possono usare regole simboliche o apprendere un modello numerico, e possono anche adattare il loro comportamento analizzando gli effetti che le loro azioni precedenti hanno avuto sull'ambiente. Come disciplina scientifica, l'IA comprende diversi approcci e diverse tecniche, come l'apprendimento automatico (di cui l'apprendimento profondo e l'apprendimento per rinforzo sono esempi specifici), il ragionamento meccanico (che include la pianificazione, la programmazione, la rappresentazione delle conoscenze e il ragionamento, la ricerca e l'ottimizzazione) e la robotica (che comprende il controllo, la percezione, i sensori e gli attuatori e l'integrazione di tutte le altre tecniche nei sistemi ciberfisici)» *A definition of AI: main capabilities and disciplines*, cit., p. 6.

[154] *Orientamenti etici per un'IA affidabile*, cit., p. 15.

Il capitolo II del documento in esame entra nello specifico, individuando sette requisiti cui sviluppatori, distributori ed utilizzatori dei sistemi intelligenti dovranno adeguarsi in ciascun contesto di pertinenza[155]:

1. *Intervento e sorveglianza umani*: i sistemi di IA devono rispettare l'autonomia umana e i diritti fondamentali, senza incidere negativamente su questi ultimi. Prima di procedere allo sviluppo di un sistema intelligente occorre effettuare una valutazione d'impatto sui suddetti diritti e sulle connesse libertà, valutando i potenziali rischi e riducendoli al minimo. Gli operatori dell'IA devono poter assumere decisioni informate in merito al funzionamento dei sistemi intelligenti in modo da potersi compiutamente relazionare con essi e da non sottostare acriticamente alle decisioni assunte dall'IA stessa senza poterle contestare. La sorveglianza umana – esercitata a vari livelli – serve inoltre ad evitare che il sistema intelligente provochi effetti negativi, secondo un approccio di intervento, supervisione o controllo umano[156], da ciò derivando che «a parità di condizioni, minore è la sorveglianza che un essere umano può esercitare su un sistema di IA, maggiore è la necessità di prove esaurienti e di una governance rigorosa»[157];

2. *Robustezza tecnica e sicurezza*: tale requisito è collegato al principio di prevenzione dei danni. Si richiede infatti che i sistemi di IA si comportino in maniera affidabile e nel rispetto della salute fisica e psichica degli individui e che siano sviluppati nell'ottica della prevenzione di potenziali rischi. I sistemi intelligenti possono essere esposti ad attacchi hacker illeciti volti a colpire i dati, il modello o la struttura dei sistemi medesimi: per tale ragione è necessario che gli strumenti dotati di IA vengano protetti da eventuali abusi e che siano studiate misure per prevenirli o per ridurne le conseguenze nocive. A tal fine potrebbe essere utile dotare tali sistemi di appositi piani d'emergenza e misure di sicurezza per evitare di causare danni ai loro utilizzatori. È inoltre necessario che il sistema intelligente – specie ove idoneo ad incidere sulla vita degli individui coinvolti – sia preciso, in grado cioè di produrre un output corretto, nonché in grado di correggere errori involontari o di individuare la percentuale di probabilità che questi si verifichino, fornendo risultati riproducibili ed affidabili;

3. *Riservatezza e governance dei dati*: anche tale requisito è connesso al principio di prevenzione dei danni che potrebbero derivare dal mancato rispetto del diritto alla riservatezza, nonché da un uso di dati non integri o non qualitativamente elevati. Riservatezza e protezione di questi ultimi – da garantire durante l'intero ciclo di vita dell'IA – sono elementi fondamentali per generare

---

[155] *Orientamenti etici per un'IA affidabile*, cit., pp. 18 ss.

[156] Tale passaggio appare di particolare rilevanza. Il Gruppo di esperti, infatti, distingue «L'approccio HITL (human-in-the-loop) prevede la possibilità di intervento umano in ogni ciclo decisionale del sistema, che in molti casi non è né possibile né auspicabile. L'approccio HOTL (human-on-the-loop) prevede l'intervento umano durante il ciclo di progettazione del sistema e il monitoraggio del funzionamento del sistema. L'approccio HIC (human-in-command) prevede il controllo dell'attività del sistema di IA nel suo complesso (compresi i suoi effetti generali a livello economico, sociale, giuridico ed etico) e la capacità di decidere quando e come utilizzare il sistema in qualsiasi particolare situazione. Si potrebbe anche decidere di non utilizzare un sistema di IA in una data situazione, di stabilire livelli di discrezionalità umana durante l'uso del sistema, o di garantire la capacità di ignorare una decisione presa da un sistema» *Orientamenti etici per un'IA affidabile*, cit., p. 18.

[157] *Orientamenti etici per un'IA affidabile*, cit., p. 18.

fiducia nelle procedure di raccolta dei dati medesimi, dovendo utilizzare questi ultimi limitatamente alle finalità per le quali sono stati raccolti. I dati utilizzali per alimentare l'IA potrebbero essere distorti o imprecisi, per questo la loro qualità deve essere controllata prima di adoperarli, stabilendo altresì protocolli volti a regolare l'accesso ai suddetti dati;

4. *Trasparenza*: tale requisito è connesso al principio di esplicabilità e si riferisce non soltanto alla trasparenza dei dati, ma anche a quella concernente il sistema e i modelli di funzionamento dell'IA. È importante garantire la tracciabilità dei processi che conducono alla decisione del sistema intelligente al fine di garantirne la verificabilità e migliorarne la trasparenza. È inoltre necessario garantire la spiegabilità degli sviluppi tecnici, affinché gli individui che interagiscono con l'IA possano comprenderne i "ragionamenti", specie ove le conclusioni da essa raggiunte siano idonee ad incidere sulla vita dei soggetti coinvolti. Il Gruppo prende in considerazione l'evenienza di dover mediare tra l'esigenza di aumentare la spiegabilità del sistema a scapito della precisione o quella di aumentare la suddetta precisione sacrificando la spiegabilità. Da ultimo, agli individui va garantito il diritto di essere resi edotti del fatto che stanno interagendo con un tipo di intelligenza artificiale e non umana[158], consentendogli di opzionare eventualmente l'interazione con un soggetto in carne ed ossa;

5. *Diversità, non discriminazione ed equità*: tale requisito richiama il principio di equità e richiede la parità di trattamento ed accesso per il tramite di metodologie di progettazione inclusive che valorizzino le diversità durante il ciclo di vita dell'IA. I dati utilizzati per addestrare i sistemi intelligenti possono essere distorti, incompleti o mal governati, da ciò potrebbe derivare un pregiudizio che, invece, dovrebbe essere evitato sin dalla fase di ricognizione dei dati medesimi mediante adeguati processi di sorveglianza. È opportuno servirsi di una progettazione universale che garantisca un accesso equo all'IA, coinvolgendo attivamente tutti i soggetti interessati durante il ciclo di vita del sistema;

6. *Benessere sociale e ambientale*: tale requisito è collegato al principio di prevenzione dei danni e a quello di equità in quanto mira a tutelare la società e l'ambiente, promuovendo uno sviluppo sostenibile durante l'intera catena di approvvigionamento dell'IA (sviluppo, distribuzione ed utilizzo del sistema). È importante valutare l'impatto sociale di questi sistemi, considerandone potenzialità e rischi;

7. *Accountability*: il requisito della "responsabilità", di particolare interesse per il penalista, appare strettamente connesso al principio di equità. Al fine di rispettare il suddetto requisito è importante garantire la verificabilità dei dati e dei processi di programmazione e progettazione dei sistemi intelligenti, garantendo la possibilità di contestare le risultanze da essi fornite. Nell'ottica di ridurre i potenziali effetti negativi dell'IA e di generare fiducia in quest'ultima occorre, da un lato, tutelare coloro che segnalano le irregolarità cui i sistemi intelligenti possono dar luogo e, dall'altro, garantire la possibilità di ricorso in caso di esiti

[158] Si comincia a parlare in tal senso del «diritto ad essere resi consapevoli della natura, umana o artificiale, del nostro interlocutore (…) la non conoscenza del carattere umano o artificiale del nostro interlocutore, perlomeno nello svolgimento di alcune attività, potrebbe assurgere ad una violazione della nostra dignità in quanto esseri umani» C. CASONATO, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Diritto pubblico comparato ed europeo*, 2019, pp. 125-126.

sfavorevoli. Anche in tal sede viene richiamata la possibilità che si vengano a creare tensioni tra i requisiti presi in esame, tuttavia qui si effettua un passaggio ulteriore: il Gruppo prende in considerazione la possibilità che non si riesca a raggiungere una soluzione di compromesso e che, pertanto, possa doversi percorrere la via di non far procedere l'IA secondo impostazioni contrarie all'etica o ai diritti fondamentali. Si chiarisce inoltre che ogni decisione di compromesso debba essere motivata e si attribuisce uno specifico ruolo al decisore il quale «deve assumersi la responsabilità delle modalità di attuazione del compromesso più adeguato e deve riesaminare costantemente l'adeguatezza della decisione che ne deriva per garantire che possano essere apportate le necessarie modifiche al sistema ove opportuno»[159].

Nel prosieguo della trattazione il Gruppo di esperti propone l'adozione di metodi tecnici[160] per dar vita ad un'IA affidabile, proponendo *in primis* la realizzazione di una lista bianca, contenente un elenco di regole che il sistema deve sempre osservare, ed una lista nera di condotte che il sistema non può mai porre in essere. Si prendono poi in considerazione i c.d. "sistemi non deterministici", ossia quei sistemi dotati di una capacità di apprendimento che li rende in grado di adattarsi all'ambiente circostante, così rendendo il proprio comportamento non del tutto prevedibile[161].

Viene raccomandata in tal sede l'adozione di metodi volti a garantire i valori fondamentali sin dalla progettazione ("*by design*"), secondo quanto già avvenuto con i meccanismi di *privacy by design* e *security by design*. Si esorta l'implementazione di meccanismi di arresto *fail safe* (volti a consentire un riavvio a seguito di un attacco o, comunque, di un arresto forzato), del campo di ricerca dell'*Explainable IA* (XAI) volto a comprendere i meccanismi di funzionamento dei sistemi intelligenti e dei metodi di prova dell'IA[162].

Subito dopo il Gruppo di esperti propone di adottare anche dei metodi non tecnici[163], ad esempio predisponendo codici di condotta che i vari operatori del settore potrebbero sottoscrivere, strutturando norme di progettazione che possano aiutare a controllare e garantire la qualità dei prodotti dotati di IA, magari ideando anche dei sistemi di accreditamento per confermare la sussistenza dei requisiti di spiegabilità, robustezza e sicurezza, certificabile da parte di organizzazioni

---

[159] *Orientamenti etici per un'IA affidabile*, cit., p. 23.

[160] *Orientamenti etici per un'IA affidabile*, cit., pp. 24 ss.

[161] Con riferimento a tali sistemi si evidenzia come spesso essi siano «valutati secondo la prospettiva teorica del ciclo "percezione/pianificazione/azione". Per adattare questa architettura affinché garantisca un'IA affidabile è necessario integrare i requisiti in tutte e tre le fasi del ciclo: i) nella fase di "percezione" si dovrebbe sviluppare il sistema in modo tale che riconosca tutti gli elementi ambientali indispensabili a garantire l'aderenza ai requisiti; ii) nella fase di "pianificazione" il sistema dovrebbe considerare solo i piani che aderiscono ai requisiti; iii) nella fase di "azione", le azioni del sistema dovrebbero essere limitate ai comportamenti che realizzano i requisiti» *Orientamenti etici per un'IA affidabile*, cit., p. 24.

[162] Con particolare riferimento a questo aspetto si valuta l'opportunità di adottare la pratica del c.d. *red teaming* ossia, secondo la definizione fornita dal Glossario contenuto nelle pagine finali degli Orientamenti etici in esame, la «pratica in cui un "red team" o un gruppo indipendente sfida un'organizzazione a migliorare la propria efficacia assumendo il ruolo o punto di vista di avversario. Tale pratica è impiegata soprattutto nell'ambito della sicurezza per contribuire a individuare e risolvere le potenziali vulnerabilità» *Orientamenti etici per un'IA affidabile*, cit., p. 46.

[163] *Orientamenti etici per un'IA affidabile*, cit., pp. 26 ss.

competenti attraverso la c.d. etichetta "IA affidabile". Viene però in tal sede chiarito che «la certificazione non può tuttavia mai sostituire la *responsabilità* e dovrebbe essere quindi integrata da quadri di accountability, tra cui clausole di esclusione della responsabilità nonché meccanismi correttivi e di riesame»[164].

Volendo fare un passo ulteriore, a nostro avviso non sembra neanche che l'operatore dell'IA possa andare esente da responsabilità per il sol fatto di essersi omologato alle regole stabilite dai codici e dalle norme di settore.

Traducendo la considerazione appena svolta in termini penalistici, il rispetto di una regola precauzionale di fonte scritta (integrante un'ipotesi di colpa specifica) contenuta nel codice di condotta o nella normativa di settore, non escluderebbe la responsabilità del soggetto coinvolto nel caso in cui venga fatto un uso negligente (dunque in violazione di una regola di fonte sociale, anche nota come colpa generica) dello strumento dotato di intelligenza artificiale[165].

Tornando all'esame dei metodi non tecnici necessari per garantire un'IA affidabile, si evidenzia che le organizzazioni di settore dovrebbero garantire il requisito della responsabilità in ognuna delle fasi di vita del sistema intelligente (sviluppo, distribuzione ed utilizzo), istituendo canali informativi volti alla condivisione di pratiche, dilemmi e problemi etici, chiarendo come tali meccanismi possano solo limitarsi a coadiuvare la supervisione giuridica, non potendosi sostituire ad essa[166].

Il Capitolo III del documento in esame fornisce le c.d. "liste di controllo per la valutazione dell'affidabilità dell'IA"[167], fornendo una specificazione dei

---

[164] *Orientamenti etici per un'IA affidabile*, cit., p. 26, corsivo nostro.

[165] Chiara sul punto autorevole dottrina penalistica secondo cui «occorre di volta in volta verificare se le norme "scritte" esauriscano la misura di diligenza richiesta all'agente nelle situazioni considerate: solo in questo caso l'osservanza di dette norme esclude la responsabilità penale. In caso contrario, ove residui cioè uno spazio di esigenze preventive non coperte dalla disposizione scritta, il giudizio di colpa può tornare a basarsi sulla inosservanza di una "generica" misura precauzionale» G. FIANDACA, E. MUSCO, *Diritto Penale*, cit. p. 582. Altra parte della dottrina ha analizzato il rapporto tra colpa generica e colpa specifica sotto un'altra visuale prospettica: «Si pone il quesito se l'inosservanza di una regola cautelare codificata *a contenuto rigido* sia di per sé sola sufficiente a fondare la colpa. La risposta è nel senso che l'inosservanza dà vita a colpa, a meno che siano presenti circostanze concrete tali da rendere il rispetto della norma stessa fonte di un aumento del rischio della realizzazione di un fatto che integra un reato colposo. In questa evenienza l'inosservanza della regola cautelare codificata è irrilevante, perché la vera regola di diligenza da osservare non è quella prescritta dalla regola cautelare codificata, bensì quella che l'agente modello avrebbe rispettato nelle circostanze concrete per evitare che quel maggior rischio si traducesse in un evento lesivo» G. MARINUCCI, E. DOLCINI, G.L. GATTA, *Manuale di diritto penale*, cit., p. 431.

[166] Il Gruppo di esperti incoraggia la formazione dei diversi portatori di interessi nel settore in modo che essi possano partecipare compiutamente allo sviluppo dell'IA, con ciò riferendosi non soltanto ai tecnici, bensì anche a giuristi, studiosi di etica ed organi rappresentativi di consumatori e lavoratori, anche al fine di incoraggiare un dialogo multidisciplinare e di includere competenze differenti nei gruppi responsabili della progettazione dei sistemi intelligenti. *Orientamenti etici per un'IA affidabile*, cit., p. 27.

[167] Tali liste di controllo verranno adottate in forma di progetto pilota – per strutturare un quadro normativo applicabile orizzontalmente a tutte le tipologie di IA – garantendo la rappresentatività di imprese ed organizzazioni interessate (processo qualitativo) nonché la partecipazione di tutti i soggetti interessati (processo quantitativo), *Orientamenti etici per un'IA affidabile*, cit., p. 28. Data la natura del progetto, in chiusura del Capitolo III, il Gruppo di esperti invita «tutti i portatori di interessi a sperimentare nella pratica la presente lista di controllo e a fornire un riscontro sulla sua attuabilità, la sua completezza, la sua pertinenza per l'applicazione o il settore specifici dell'IA,

requisiti contenuti nel Capitolo II. Tali liste sono rivolte a sviluppatori e distributori, per quanto – un po' riprendendo le considerazioni svolte pocanzi – l'osservanza di tali liste non esime dalla responsabilità per i potenziali danni causati dal sistema intelligente[168].

Queste liste di controllo (per quanto non esaustive) servono a valutare se un dato sistema di IA possa considerarsi affidabile e ad orientare gli operatori del settore, incoraggiandoli a riflettere su quali passaggi delle suddette liste potrebbero essere integrati per garantire una maggiore affidabilità dell'IA. In alcuni punti le succitate liste possono sovrapporsi con le disposizioni normative attualmente vigenti, in particolare per quanto concerne la legislazione sulla protezione dei dati personali, i quali devono essere trattati in modo etico e nel rispetto della normativa già esistente sul punto.

Nell'ultima sezione del documento in esame vengono proposti alcuni settori in cui l'IA potrebbe trovare terreno fertile per crescere e portare benefici per la società intera, come ad esempio il settore ambientale – per realizzare una migliore gestione del fabbisogno energetico e dei consumi –, il settore dei trasporti[169] e quello sanitario[170]. Al contempo vengono però individuati specifici ambiti in cui l'uso di forme di intelligenza artificiale desta non poche preoccupazioni, basti pensare ai sistemi di IA in grado di identificare gli individui mediante il riconoscimento facciale senza il loro consenso[171], la possibilità che i soggetti

---

come pure sulla sua sovrapposizione o complementarietà con i processi di conformità o di valutazione esistenti. In base a tali riscontri, all'inizio del 2020 verrà proposta alla Commissione una versione riveduta della lista di controllo» *Orientamenti etici per un'IA affidabile*, cit., p. 39.

[168] «Il fatto di essersi attenuti a tale lista di controllo non è una prova di conformità giuridica, né tale lista va intesa come un orientamento per garantire la conformità alle leggi vigenti. Data la specificità applicativa dei sistemi di IA, la lista di controllo dovrà essere adattata ai casi d'uso e ai contesti specifici in cui operano i sistemi» *Orientamenti etici per un'IA affidabile*, cit., p. 28. Analogo ragionamento viene ripreso più avanti ove il Gruppo di esperti ribadisce che «La conformità alla lista di controllo non prova, tuttavia, la conformità giuridica, né tale lista va intesa come un orientamento per garantire la conformità alle leggi vigenti. Lo scopo della lista di controllo è piuttosto quello di porre una serie di domande specifiche a coloro che intendono garantire che il loro approccio allo sviluppo o alla distribuzione dell'IA è orientato verso un'IA affidabile, e tenta di garantire tale affidabilità» *Orientamenti etici per un'IA affidabile*, cit., p. 30.

[169] «Nel settore del trasporto pubblico i sistemi di IA per i trasporti intelligenti possono essere utilizzati per ridurre al minimo le code, ottimizzare i percorsi, consentire alle persone con deficit visivi di essere più indipendenti e ottimizzare i motori ad alta efficienza energetica, contribuendo così agli sforzi di decarbonizzazione e alla riduzione dell'impatto ambientale, a favore di una società più verde. Attualmente nel mondo ogni 23 secondi un essere umano muore in un incidente d'auto. I sistemi di IA potrebbero contribuire a ridurre significativamente il numero di vittime, ad esempio migliorando i tempi di reazione e il rispetto delle regole» *Orientamenti etici per un'IA affidabile*, cit., pp. 39-40.

[170] I sistemi intelligenti possono coadiuvare il personale sanitario nella loro attività, strutturando terapie personalizzate, prevenendo patologie mortali, svolgendo analisi accurate, servendosi della robotica per assistere gli anziani, nonché per supervisionare i parametri vitali dei pazienti e poter intervenire tempestivamente. *Orientamenti etici per un'IA affidabile*, cit., p. 40.

[171] «L'identificazione automatica tuttavia desta enormi preoccupazioni di natura sia giuridica che etica, in quanto può avere effetti non previsti sotto molti aspetti a livello psicologico e socioculturale. Per salvaguardare l'autonomia dei cittadini europei è necessario ricorrere alle tecniche di controllo tramite l'IA in modo proporzionato. Definire chiaramente se, quando e come l'IA può essere utilizzata per l'identificazione automatica degli individui e differenziare tra l'identificazione di un individuo e la sua tracciatura e localizzazione, e tra sorveglianza mirata e sorveglianza di massa, sarà fondamentale per ottenere un'IA affidabile. L'applicazione di tali tecnologie deve essere chiaramente motivata dal diritto vigente. Se la base giuridica di tale attività

interagiscano con un sistema intelligente nell'erronea convinzione di dialogare con un umano[172], nonché l'uso di sistemi d'arma autonomi letali[173].

Gli Orientamenti etici enucleati nel presente paragrafo hanno costituito oggetto di una specifica Comunicazione della Commissione, volta a creare fiducia nell'intelligenza artificiale antropocentrica e con la quale ci si propone di avviare una fase pilota per testare le suesposte Linee Guida per uno sviluppo affidabile dell'IA e per raccogliere feedback da parte dei soggetti interessati, specie per quanto concerne la guida autonoma e le banche dati medicali, sfruttando il supporto dell'Alleanza europea per l'IA e della piattaforma di IA "on demand" AI4EU[174].

è rappresentata dal "consenso", devono essere sviluppati mezzi pratici che permettano di dare un consenso eloquente e verificato ad essere identificati automaticamente da un sistema di IA o da tecnologie equivalenti. Ciò vale anche per l'utilizzo di dati personali "anonimi" che possono essere ripersonalizzati» *Orientamenti etici per un'IA affidabile*, cit., p. 41.

[172] Si dovrà prestare particolare attenzione alle problematiche etiche che potranno porsi con lo sviluppo dei c.d. androidi. Non si tratta più di prospettivi futuristiche ma di tangibile realtà. Nell'agosto del 2021 Elon Musk ha annunciato l'arrivo nel 2022 del prototipo del Tesla Bot, un robot umanoide sfruttabile per i più diversificati utilizzi, cfr. https://www.tesla.com/it_IT/AI.

[173] «Attualmente, un numero imprecisato di paesi e industrie si dedica alla ricerca e allo sviluppo di sistemi d'arma autonomi letali, come missili capaci di selezionare gli obiettivi o macchine ad apprendimento automatico con abilità cognitive che consentono di decidere chi, quando e dove combattere senza l'intervento umano. Questa situazione pone interrogativi etici fondamentali, per esempio la possibilità di una corsa incontrollabile agli armamenti a un livello storicamente senza precedenti e la creazione di contesti militari in cui il controllo umano è quasi del tutto assente e i rischi di malfunzionamento non sono presi in considerazione. Il Parlamento europeo ha chiesto l'elaborazione urgente di una posizione comune giuridicamente vincolante che affronti questioni etiche e giuridiche fondamentali relative al controllo e alla supervisione da parte dell'uomo, all'accountability e all'attuazione del diritto internazionale in materia di diritti umani, del diritto internazionale umanitario e delle strategie militari. Ricordando che l'Unione europea si prefigge di promuovere la pace come sancito dall'articolo 3 del trattato sull'Unione europea, sosteniamo la risoluzione del Parlamento europeo del 12 settembre 2018 e tutti gli sforzi correlati in materia di sistemi d'arma autonomi letali» *Orientamenti etici per un'IA affidabile*, cit., p. 42.

[174] *Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni. Creare fiducia nell'intelligenza artificiale antropocentrica*, Bruxelles, 8.4.2019, pp. 7-8. La Commissione evidenzia in tal sede come la fiducia costituisca un elemento fondamentale realizzare un'intelligenza artificiale antropocentrica. L'Unione Europea è stata edificata su principi cardine quali il rispetto della dignità umana, della libertà, della democrazia, dell'uguaglianza, dello Stato di diritto, dei diritti umani, del pluralismo, della non discriminazione, della tolleranza, della giustizia e della solidarietà. Tutti questi principi integrano il quadro normativo che andrà a costituire lo standard comune per realizzare un'IA antropocentrica. La Comunicazione in parola prende altresì in considerazione i sistemi di IA più sofisticati i quali sono in grado di assumere decisioni e portarle ad esecuzione senza l'intervento dell'uomo, ma queste decisioni possono essere prese sulla base di dati incompleti o errati, dando così luogo ad output condizionati. Tali effetti collaterali dovrebbero essere evitati rispettando le leggi ed i principi etici ad essi sottesi, anche al fine di generare fiducia nell'uso di questi sistemi e di aumentare la competitività europea sul mercato mondiale grazie ad un'intelligenza artificiale "made in Europe" affidabile e basata sulle esigenze dell'uomo. La Commissione si propone di cooperare con partner internazionali che aderiscono ai suddetti valori e di partecipare a tavoli tecnici ed iniziative sul tema dell'intelligenza artificiale, contribuendo a creare reti di centri di eccellenza per la ricerca sull'IA, poli d'innovazione digitale e modelli per la condivisione dei dati. *Creare fiducia nell'intelligenza artificiale antropocentrica*, cit., pp. 1-2 e 9-10.

8. Il Libro Bianco sull'intelligenza artificiale.

Il presente Libro Bianco merita di essere segnalato nel corpo di questa nostra analisi in quanto, per un verso, richiama (seppur indirettamente) problemi penalistici di non secondario rilievo nel settore e, per altro verso, fornisce input importanti che saranno ripresi ed approfonditi dalla Proposta di Regolamento del 21.4.2021 che sarà fatta oggetto di analisi sul finire di questo primo capitolo.

Il 19 febbraio del 2020 la Commissione europea ha pubblicato il Libro Bianco sull'intelligenza artificiale con lo scopo di stabilire una strategia operativa – rispettosa dei valori europei ed ispirata ad uno sviluppo affidabile – per raggiungere due specifici obiettivi: incoraggiare l'adozione dell'intelligenza artificiale e, conseguentemente, fronteggiare i potenziali rischi ad essa connessi[175].

Nella fase iniziale il presente Libro Bianco individua una serie di azioni volte a realizzare un "*ecosistema di eccellenza*" per la creazione del quale appare necessario incentivare la collaborazione tra gli Stati membri[176] e a livello internazionale[177], concentrare gli sforzi della comunità scientifica nella ricerca[178], sviluppare le competenze necessarie per poter lavorare nell'IA[179], sostenere le piccole e medie imprese[180], incentivare partenariati col settore privato[181], promuovere l'adozione dell'IA nel settore pubblico[182] e, da ultimo, garantire l'accesso ai dati e alle infrastrutture di calcolo[183].

---

[175] *Libro Bianco sull'intelligenza artificiale. Un approccio europeo all'eccellenza e alla fiducia*, Bruxelles, 19.2.2020, p. 1. In fase introduttiva la Commissione qualifica l'intelligenza artificiale come una moltitudine di tecnologie, evidenziando le potenzialità dell'Europa nel settore e chiarendo come essa sia in grado di realizzare prodotti di alta qualità e di dar vita ad un quadro normativo fondato sui diritti fondamentali, tra cui la tutela della dignità umana e della privacy.

[176] La Commissione si propone di coinvolgere gli Stati in una revisione del piano coordinato sull'intelligenza artificiale (cfr. Cap. I, Sez. I, Par. 4) agevolando gli investimenti pubblici e privati nel settore. *Libro Bianco sull'intelligenza artificiale,* cit., p. 6.

[177] *Libro Bianco sull'intelligenza artificiale,* cit., pp. 5 ss., lett. a) e h). La Commissione evidenzia l'importanza di creare alleanze internazionali fondate sui valori europei, sul rispetto dei diritti fondamentali e su un'intelligenza artificiale etica. L'OCSE (l'Organizzazione per la Cooperazione e lo Sviluppo Economico) ha individuato cinque principi etici per l'intelligenza artificiale: crescita inclusiva, sviluppo sostenibile e benessere; valori incentrati sull'uomo e correttezza; trasparenza e spiegabilità; robustezza, sicurezza e protezione; responsabilità. Per un approfondimento di tali ultimi principi https://oecd.ai/en/ai-principles.

[178] La Commissione prende atto della natura frammentaria degli attuali centri di competenza nel settore, evidenziando la necessità di coordinare questi sforzi facendoli confluire in un unico centro. *Libro Bianco sull'intelligenza artificiale,* cit., p. 6, lett. b).

[179] La Commissione sostiene la realizzazione delle competenze necessarie per migliorare l'intelligenza artificiale e per guidare coloro che si troveranno a dover lavorare con essa, coinvolgendo le parti sociali al fine di realizzare un approccio antropocentrico. *Libro Bianco sull'intelligenza artificiale,* cit., p. 7, lett. c).

[180] La Commissione sottolinea la necessità di consolidare i poli d'innovazione digitale e la piattaforma di IA on demand, garantendo alle piccole e medie imprese un ampio accesso all'intelligenza artificiale. *Libro Bianco sull'intelligenza artificiale,* cit., p. 8, lett. d).

[181] La Commissione incoraggia il coinvolgimento del settore privato al fine di realizzare partenariati pubblico-privati e di collaborare con altri partenariati nella di strutture di prova e poli d'innovazione. *Libro Bianco sull'intelligenza artificiale,* cit., p. 8, lett. e).

[182] Trasporti, ospedali e pubbliche amministrazioni inizieranno a utilizzare servizi che sfruttano l'intelligenza artificiale, per questo la Commissione incoraggia un ampio dialogo nel settore. *Libro Bianco sull'intelligenza artificiale,* cit., p. 9, lett. f).

[183] La Commissione, alla luce della complementarità degli ambiti di intervento dell'IA, incoraggia lo sviluppo dell'accesso ai dati, necessari ad alimentare i sistemi intelligenti, promuovendone una gestione conforme ai principi di reperibilità, accessibilità, interoperabilità e riutilizzabilità. Il Libro

Nel prosieguo il Libro Bianco affronta il tema della creazione di un "*ecosistema di fiducia*", chiarendo come inevitabilmente l'avvento dell'intelligenza artificiale sia foriero di considerevoli potenzialità in termini di tutela dei diritti dei consociati ma anche di possibili rischi, come ad esempio lo sfruttamento dell'IA per scopi dolosi. La Commissione evidenzia l'importanza di un solido quadro normativo, in grado di generare fiducia e di garantire il rispetto dei valori europei, specialmente nei casi in cui l'IA venisse applicata in settori delicati quali gli uffici giudiziari o delle forze dell'ordine.

Gli sviluppatori dei sistemi intelligenti devono rispettare la normativa europea – per quanto ciò possa essere reso più complesso dalle caratteristiche specifiche dell'IA, quali l'opacità –, specie per quanto concerne la tutela dei consumatori ai quali deve essere garantito un livello di sicurezza adeguato. La Commissione evidenzia la necessità di verificare se la vigente legislazione sia idonea a fronteggiare i dubbi posti dai sistemi intelligenti – specie alla luce della loro rapida evoluzione – o se sia piuttosto necessario effettuare eventuali modifiche. Si constata l'attuale assenza di un quadro europeo comune ed il conseguente rischio di realizzare così una disgregazione del mercato interno, conseguentemente si incentiva la realizzazione di un assetto condiviso volto ad agevolare lo sviluppo dell'IA e a consolidare la base industriale europea nel settore[184].

I sistemi intelligenti possono causare danni tanto materiali quanto immateriali: per questa ragione è necessario dar vita ad un quadro normativo in grado di minimizzare i potenziali rischi – derivabili da difetti di progettazione o dall'uso di dati distorti[185] – e di affrontare le problematiche connesse alla sicurezza ed alla responsabilità. Con particolare riferimento al rischio di creare distorsioni, la Commissione evidenzia che le discriminazioni tipiche dell'agire umano possono riflettersi nel sistema intelligente il quale potrebbe, a sua volta, realizzare discriminazioni pregiudizievoli[186].

Il Libro Bianco si concentra nel prosieguo sui rischi per la sicurezza degli utilizzatori dell'IA[187], evidenziando come l'assenza di un quadro giuridico chiaro

---

Bianco fa riferimento ai c.d. principi FAIR (*Findable*, *Accessible*, *Interoperable* e *Reusable*). *Libro Bianco sull'intelligenza artificiale,* cit., p. 9, lett. g).

[184] *Libro Bianco sull'intelligenza artificiale,* cit., pp. 10-11.

[185] Con riferimento all'ambito penalistico, segnatamente con riferimento materia della recidiva, il Libro Bianco chiarisce che «alcuni algoritmi dell'IA, se usati per prevedere il rischio di recidiva di atti delittuosi, possono riflettere distorsioni legate alla razza e al genere, prevedendo probabilità di rischio di recidiva diverse per le donne rispetto agli uomini, oppure per i cittadini di un determinato paese rispetto agli stranieri» *Libro Bianco sull'intelligenza artificiale,* cit., p. 13.

[186] Nel prosieguo la Commissione chiarisce efficacemente come «Le caratteristiche specifiche di molte tecnologie di IA, tra cui l'opacità (effetto "scatola nera"), la complessità, l'imprevedibilità e un comportamento parzialmente autonomo, possono rendere difficile verificare il rispetto delle normative dell'UE in vigore volte a proteggere i diritti fondamentali e possono ostacolarne l'applicazione effettiva. Le autorità preposte all'applicazione della legge e le persone interessate potrebbero non disporre dei mezzi per verificare come sia stata presa una determinata decisione con il coinvolgimento di sistemi di IA e, di conseguenza, se sia stata rispettata la normativa pertinente. Le persone fisiche e giuridiche possono incontrare difficoltà nell'accesso effettivo alla giustizia in situazioni in cui tali decisioni possono avere ripercussioni negative su di loro» *Libro Bianco sull'intelligenza artificiale,* cit., p. 13.

[187] Il Libro Bianco in esame riporta l'esempio di un'auto a guida autonoma che, per un difetto strutturale, non riconosca un oggetto lungo la carreggiata e finisca con il provocare un incidente

a riguardo sia potenzialmente foriero di incertezza, rischiando così di danneggiare la competitività delle imprese europee. Non solo: dalla mancanza di un preciso quadro giuridico deriverebbe anche che, ove i suddetti rischi per la sicurezza dovessero concretizzarsi, risulterebbe arduo per le persone offese coinvolte ottenere il risarcimento dei danni subiti. Inoltre la difficoltà di ricostruire il procedimento decisionale seguito dal sistema intelligente renderebbe arduo per le vittime rinvenire gli elementi probatori necessari per agire in giudizio e, conseguentemente, ottenere il relativo risarcimento[188].

Viene ripreso in tal sede, ed approfondito, l'interrogativo concernente la possibilità di applicare efficacemente la normativa europea per fronteggiare i rischi tipici dell'intelligenza artificiale oppure se non sia più opportuno adeguare il quadro normativo vigente per affrontare i seguenti, specifici, rischi[189]:

- *Effettiva applicazione e rispetto della normativa nazionale e dell'UE in vigore*: da un lato le caratteristiche tipiche dell'intelligenza artificiale rischiano di porsi in contrasto con la normativa europea e, dall'altro lato, l'opacità dei sistemi intelligenti renderebbe arduo comprendere quando si sia effettivamente realizzata la violazione di una norma, a chi sia possibile attribuire la responsabilità del danno e come muoversi per ottenerne il risarcimento. Potrebbe in tal senso essere utile rivedere la normativa europea proprio per quanto concerne l'aspetto della responsabilità;
- *Limiti dell'ambito di applicazione della legislazione dell'UE vigente*: ci si chiede, in particolare, quale dovrebbe essere la disciplina dei software indipendenti e se questi possano essere soggetti alla normativa europea in materia di sicurezza dei prodotti;
- *Funzionalità mutevole dei sistemi di IA*: il funzionamento dei sistemi intelligenti può subire modificazioni nel corso del tempo grazie agli aggiornamenti software ed al relativo apprendimento automatico, dando così vita a rischi non preventivati, non preventivabili e, soprattutto, non regolamentati dalla normativa vigente;
- *Incertezza in merito all'attribuzione delle responsabilità tra i diversi operatori economici lungo la catena di approvvigionamento*: la normativa europea attualmente vigente attribuisce la responsabilità del prodotto e dei suoi componenti al fabbricante, tuttavia è possibile che l'intelligenza artificiale venga integrata all'interno di un prodotto in una fase successiva rispetto all'immissione sul mercato e, soprattutto, da un soggetto diverso. Come dovrebbe ripartirsi la responsabilità tra i vari membri della catena di approvvigionamento? Il presente

---

stradale. *Libro Bianco sull'intelligenza artificiale,* cit., p. 13. Non si tratta, invero, di mere supposizioni ma di fatti realmente verificabili, come approfondito al Cap. III, Par. 2.1.

[188] Attualmente, «A norma della direttiva sulla responsabilità per danno da prodotti difettosi, il produttore è responsabile dei danni causati da un prodotto difettoso. Tuttavia, nel caso di sistemi basati sull'IA, come quelli delle auto a guida autonoma, può rivelarsi difficile provare che il prodotto è difettoso e dimostrare il danno cagionato e il nesso di causalità tra difetto e danno. In aggiunta non è chiaro come e in che misura si applichi la direttiva sulla responsabilità per danno da prodotti difettosi nel caso di alcuni tipi di difetti, ad esempio per quelli risultanti da carenze della cibersicurezza del prodotto» *Libro Bianco sull'intelligenza artificiale,* cit., p. 14.

[189] *Libro Bianco sull'intelligenza artificiale,* cit., pp. 15 ss.

inciso sembra richiamare il complesso meccanismo della successione di posizioni di garanzia[190];

- *Evoluzione del concetto di sicurezza*: esistono rischi attualmente non affrontati dalla normativa europea che possono presentarsi sia in fase di immissione del prodotto nel mercato sia in una fase successiva derivante da un aggiornamento del sistema. Sarebbe pertanto utile valutare le potenziali minacce del settore anche grazie all'aiuto dell'Agenzia dell'Unione europea per la cibersicurezza (ENISA).

Un approccio comune europeo eviterebbe il proliferare di normative nazionali non armonizzate e garantirebbe una disciplina organica, ferma restando la possibilità che si renda necessaria l'introduzione di una nuova legislazione che disciplini specificatamente la materia dell'intelligenza artificiale, così adeguando compiutamente il quadro normativo europeo attualmente vigente.

Dopo aver richiamato le definizioni di "Intelligenza Artificiale" fornite dalla Comunicazione "L'Intelligenza artificiale per l'Europa" e dagli "Orientamenti Etici per un'IA affidabile"[191], la Commissione evidenzia che ogni nuovo strumento giuridico dovrà far proprio una definizione di "IA" che sia al contempo flessibile e chiara[192].

Si richiama altresì la necessità di adottare un approccio normativo "basato sul rischio"[193], definendo il differente grado di incognita delle varie applicazioni di IA secondo criteri chiari e predeterminati, anche al fine di garantire un intervento normativo proporzionato e ferma restando la soggezione di tali sistemi intelligenti al diritto europeo. La Commissione considera ad alto rischio i sistemi intelligenti che siano, da un lato, utilizzati in settori in cui possono verificarsi seri rischi (come, ad esempio, quello sanitario o energetico) e, dall'altro, che siano impiegati con modalità tali da poter generare rischi considerevoli[194].

---

[190] In ordine alla suddivisione di compiti e doveri ed alla relativa ripartizione delle responsabilità v. F. SGUBBI, *Responsabilità penale per omesso impedimento dell'evento*, Padova, 1975, p. 198 e p. 239. Più in generale sul trasferimento della posizione di garanzia G. FIANDACA, *Il reato commissivo mediante omissione*, Milano, 1979, p. 188. Sui doveri di sorveglianza del soggetto delegato e in ordine alla relativa *culpa in eligendo* ID., *op. ult. cit.*, pp. 202 ss. Sull'efficacia liberatoria del trasferimento della posizione di garanzia, G. GRASSO, *Il reato omissivo improprio. La struttura obiettiva della fattispecie*, Milano, 1983, p. 319 e p. 326. Sull'assunzione "di fatto" della posizione di garanzia G. FIANDACA, *op. ult. cit.*, p. 86; G. FIANDACA, E. MUSCO, *Diritto Penale*, cit., p. 653. Per ulteriori riflessioni sul tema delle posizioni di garanzia nel settore dell'intelligenza artificiale cfr. Cap. II, Sez. III, Par. 17.7.

[191] Riportate rispettivamente alle note n. 37 e 153 del presente capitolo.

[192] Da un punto di vista squisitamente tecnico il Libro Bianco chiarisce come i principali elementi che compongono l'IA sono da individuare nei "dati" e negli "algoritmi". Questi ultimi «vengono addestrati per dedurre determinati modelli partendo da un set di dati al fine di stabilire le azioni necessarie al conseguimento di un determinato obiettivo». La Commissione chiarisce poi che, nonostante sia ben possibile che i sistemi intelligenti agiscano autonomamente, in realtà il loro comportamento è predeterminato dagli sviluppatori che programmano l'IA per raggiungere un dato obiettivo. *Libro Bianco sull'intelligenza artificiale,* cit., pp. 18-19.

[193] Approccio che verrà ripreso nel testo esaminato al Cap. I, Sez. I, Par. 9.1.

[194] La Commissione chiarisce inoltre che, al fine di valutare se ci si trovi in presenza di un sistema di IA ad alto rischio occorra considerare gli interessi coinvolti e i potenziali rischi per il settore di riferimento ma anche, in prospettiva, le possibili conseguenze per i soggetti interessati: «Ad esempio, usi delle applicazioni di IA che producono effetti giuridici, o effetti altrettanto rilevanti, sui diritti di una persona o di una società; usi che presentano il rischio di lesioni, morte o danni materiali o immateriali significativi; usi che producono effetti non ragionevolmente evitabili dalle persone fisiche o giuridiche». Nel prosieguo la Commissione prende anche in considerazione la

In prospettiva *de iure condendo* il quadro normativo concernente l'IA potrebbe imporre una serie di prescrizioni giuridiche obbligatorie riguardanti i seguenti aspetti[195]:

- *Dati di addestramento*: come abbiamo ormai avuto modo di imparare, i dati costituiscono il motore dell'IA, per questa ragione è di fondamentale importanza che i set di dati utilizzati per alimentare questi sistemi siano rispettosi dei valori europei. Si potrebbe a tal fine richiedere la conformità di siffatti dati a standard prestabiliti, la garanzia che il sistema intelligente non realizzi discriminazioni grazie all'uso di dati che garantiscano un'ampia rappresentatività nonché, chiaramente, la protezione dei dati personali e della privacy;
- *Tenuta dei dati e dei registri*: l'opacità dei sistemi intelligenti rende arduo comprendere se siano state rispettate le norme vigenti, per questa ragione appare opportuno conservare registri che documentino la programmazione dell'IA e i dati utilizzati per alimentarla, anche al fine di ricostruire e verificare le decisioni problematiche da questa eventualmente assunte. Tali registri – la cui conservazione sarebbe limitata ad un ragionevole periodo di tempo – dovrebbero essere accessibili dietro richiesta ove funzionali a fornire elementi probatori presso le autorità competenti, come ad esempio in sede processuale;
- *Obblighi di informazione*: la trasparenza costituisce elemento fondamentale per creare fiducia in un'intelligenza artificiale responsabile e maggiormente idonea a garantire il risarcimento di eventuali danni. Il rafforzamento della trasparenza potrebbe passare attraverso informazioni chiare sulle capacità del sistema intelligente e sulle prestazioni che è dato attendersi da quest'ultimo, nonché sul fatto che non si stia interagendo con un umano bensì con una forma di intelligenza artificiale;
- *Robustezza e precisione*: i sistemi intelligenti devono essere tecnicamente precisi e realizzati secondo una preliminare ed accurata valutazione dei possibili rischi, essendo importante che questi meccanismi funzionino in modo prevedibile ed affidabile, adottando eventualmente misure idonee a ridurre il pericolo di causazione dei danni;
- *Sorveglianza umana*: garantire l'autonomia umana costituisce presupposto necessario per assicurare lo sviluppo di un'IA antropocentrica ed etica. Il livello di controllo umano può chiaramente variare a seconda dei potenziali effetti sui soggetti coinvolti. Alcune accortezze volte a garantire tale sorveglianza potrebbero essere la convalida umana del risultato fornito dall'IA o la possibilità dell'uomo di intervenire tempestivamente per disattivare il sistema intelligente;
- *Prescrizioni specifiche per l'identificazione biometrica remota*: la raccolta di dati biometrici funzionale all'identificazione dei consociati rischia di porsi in contrasto con i diritti fondamentali. Ai sensi del GDPR tale identificazione può avvenire solo in un esiguo numero di casi di interesse pubblico e purché ciò avvenga in modo proporzionato e giustificato, nonché riservando adeguate garanzie al soggetto identificato.

Il Libro Bianco in esame, nel prosieguo, mette in luce l'opportunità che il quadro giuridico che si delineerà in materia affronti il tema della ripartizione degli

---

possibilità che esistano casi in cui l'IA sia considerata di per sé ad alto rischio, indipendentemente dai due requisiti succitati, ad esempio in ragione dell'importanza che quest'ultima potrebbe rivestire per i consociati. *Libro Bianco sull'intelligenza artificiale,* cit., p. 20.

[195] *Libro Bianco sull'intelligenza artificiale,* cit., pp. 20 ss.

obblighi – in base ai potenziali e specifici rischi – in capo a ciascuna delle figure professionali coinvolte nel ciclo della vita del sistema intelligente. Ciò consentirebbe di garantire l'attribuzione della responsabilità per eventuali danni derivati agli utilizzatori del prodotto nonché una maggiore agevolazione per questi ultimi nel richiedere il relativo risarcimento del danno[196]. La Commissione raccomanda inoltre che tale quadro giuridico trovi un'omogenea applicazione dal punto di vista geografico, essendo uniformemente applicabili su tutto il territorio dell'Unione Europea.

In tal sede la Commissione evidenzia altresì l'importanza di prevenire applicazioni dell'IA potenzialmente nocive per i consociati mediante valutazioni d'impatto volte a garantire il rispetto delle indicazioni legislative del settore. Ove le procedure necessarie per realizzare la valutazione d'impatto[197] venissero normativizzate potrebbero integrare a tutti gli effetti regole cautelari la cui inosservanza comporterebbe la responsabilità colposa dell'incaricato della catena operativa dell'IA che aveva il compito di osservarle. Spetterebbe dunque agli organi inquirenti dimostrare in giudizio che, ove le suddette regole fossero state osservate, l'evento lesivo non si sarebbe verificato. Si riproporrebbe in tal sede anche il problema, preliminarmente accennato concernente la difficoltà di individuare il soggetto effettivamente responsabile del malfunzionamento dell'IA, stante la complessità delle catene di montaggio e funzionamento che possono connotare i sistemi intelligenti[198].

La realizzazione di una valutazione d'impatto di questi sofisticati sistemi – e, conseguentemente, l'individuazione del soggetto cui imputare la responsabilità del danno da essi cagionato, sempre che esista un reale responsabile – può diventare particolarmente complessa con riferimento a quegli strumenti dotati di intelligenza artificiale capaci di apprendere dall'esperienza e che, dunque, si evolvono sfuggendo al controllo umano.

La strutturazione di un sistema di verifiche preliminari (le quali non escluderebbero anche un controllo *ex post* del rispetto della normativa di settore) consentirebbe anche di poter individuare i sistemi intelligenti che non si pongano in linea con le regole prestabilite e di intervenire per correggere gli errori e "riaddestrare" il sistema[199].

La Commissione propone nel prosieguo la creazione di un sistema di etichettatura volontaria per le applicazioni dell'IA non "ad alto rischio"[200] nonché

---

[196] A tal riguardo la Commissione chiarisce che attualmente «in base al diritto dell'UE in materia di responsabilità per danno da prodotti difettosi, tale responsabilità è attribuita al produttore, fatta salva la normativa nazionale, che può anche prevedere la possibilità di rivalersi su altri soggetti» *Libro Bianco sull'intelligenza artificiale,* cit., p. 25.

[197] «Le valutazioni della conformità per le applicazioni di IA ad alto rischio dovrebbero far parte dei meccanismi di valutazione della conformità già esistenti per un gran numero di prodotti immessi sul mercato interno dell'UE. Qualora non sia possibile avvalersi di tali meccanismi, potrebbe essere necessario istituire meccanismi analoghi, avvalendosi delle migliori prassi e di eventuali contributi dei portatori di interessi e delle organizzazioni europee di normazione. Tali nuovi meccanismi dovrebbero essere proporzionati e non discriminatori e utilizzare criteri trasparenti e obiettivi in conformità agli obblighi internazionali» *Libro Bianco sull'intelligenza artificiale*, cit., p. 26.

[198] Si tornerà sul tema del c.d. *many hands problem* al Cap. II, Sez. II, Par. 12.

[199] *Libro Bianco sull'intelligenza artificiale*, cit., p. 26.

[200] «Nel quadro di tale sistema, gli operatori economici interessati non soggetti alle prescrizioni obbligatorie potrebbero decidere, su base volontaria, di conformarsi a tali prescrizioni o di

la creazione di una *governance* europea in materia di IA volta ad evitare l'affastellarsi di responsabilità e a realizzare un proficuo scambio di informazioni e buone pratiche[201].

La Commissione reputa inoltre di tale importanza la necessità di garantire un ricorso alla giustizia effettivo per coloro i quali venissero danneggiati dall'IA da dedicare al tema – seppur nel contesto del testo in esame – un documento a sé stante che sarà oggetto di trattazione del paragrafo seguente.

Il Libro Bianco, infine, conclude evidenziando come un'IA etica, antropocentrica e rispettosa dei diritti fondamentali sia in grado di portare con sé considerevoli vantaggi per i consociati, annoverando tra le sfide cui quest'ultima è in grado di contribuire – per quanto di nostra competenza – la lotta alla criminalità (chiaramente nei termini di un uso proporzionato e necessario).

8.1. La Relazione sulle implicazioni dell'intelligenza artificiale, dell'Internet delle cose e della robotica in materia di sicurezza e di responsabilità.

La Relazione della Commissione che ci accingiamo ad esaminare e che accompagna il Libro Bianco verte su due temi particolarmente cari al penalista che intenda avvicinarsi allo studio della materia *de qua*, ossia la sicurezza e la responsabilità[202]. Scopo della Relazione in esame è quello di individuare le possibili lacune e le potenziali implicazioni dell'avvento dell'intelligenza artificiale sulla normativa attualmente vigente.

---

impegnarsi al rispetto di una serie specifica di prescrizioni analoghe, stabilite appositamente ai fini del sistema volontario. Agli operatori economici interessati sarebbe allora assegnato un marchio di qualità per le loro applicazioni di IA. (…) Tale opzione richiederebbe la creazione di un nuovo strumento giuridico che definisca il quadro relativo a un sistema di etichettatura su base volontaria per gli sviluppatori e/o i soggetti che applicano i sistemi di IA non considerati ad alto rischio. Sebbene la partecipazione al sistema di etichettatura sia facoltativo, una volta che gli sviluppatori o i soggetti che applicano l'IA hanno deciso di aderirvi le relative prescrizioni sarebbero vincolanti. La combinazione di attività di applicazione della legge ex ante ed ex post dovrebbe garantire il rispetto di tutte le prescrizioni» *Libro Bianco sull'intelligenza artificiale,* cit., p. 27.

[201] «Considerando che vi sono strutture già esistenti, ad esempio nei settori finanziario, farmaceutico, dell'aviazione, dei dispositivi medici, della tutela dei consumatori o della protezione dei dati, la struttura di governance proposta non dovrebbe duplicare le funzioni esistenti, bensì stabilire stretti legami con altre autorità competenti a livello nazionale e dell'UE nei vari settori, al fine di integrare le competenze esistenti e aiutare le autorità nel monitoraggio e nella sorveglianza delle attività svolte dagli operatori economici in cui intervengono sistemi di IA nonché prodotti e servizi basati sull'IA» *Libro Bianco sull'intelligenza artificiale,* cit., p. 28.

[202] «L'obiettivo generale dei quadri giuridici in materia di sicurezza e di responsabilità è garantire che tutti i prodotti e servizi, compresi quelli che integrano le tecnologie digitali emergenti, funzionino in modo sicuro, affidabile e costante e che vi siano rimedi efficaci in caso di danni. Livelli elevati di sicurezza dei prodotti e dei sistemi che integrano le nuove tecnologie digitali e meccanismi solidi per rimediare ai danni verificatisi (ossia il quadro della responsabilità) contribuiscono a tutelare meglio i consumatori. Creano inoltre fiducia in queste tecnologie, che è un prerequisito per la loro adozione da parte di imprese e utilizzatori. Questa contribuirà, a sua volta, a rafforzare la competitività delle nostre imprese e a realizzare gli obiettivi dell'Unione. Con l'emergere di nuove tecnologie, come l'intelligenza artificiale, l'Internet delle cose e la robotica, acquista particolare importanza un quadro chiaro in materia di sicurezza e di responsabilità, sia per tutelare i consumatori sia per garantire la certezza del diritto per le imprese» *Relazione della Commissione al Parlamento europeo, al Consiglio e al Comitato economico e sociale europeo. Relazione sulle implicazioni dell'intelligenza artificiale, dell'Internet delle cose e della robotica in materia di sicurezza e di responsabilità*, COM(2018) 237, Bruxelles, 19.2.2020, p. 1.

Come abbiamo ormai avuto modo di imparare dallo studio dei documenti precedentemente esaminati, i sistemi intelligenti sono in grado di apprendere dall'esperienza e, pertanto, di migliorare le loro prestazioni. Tale abilità, unita alla connettività ed all'opacità del funzionamento del sistema, rendono arduo prevedere il comportamento dei sistemi di IA maggiormente sviluppati e, conseguentemente, individuare le possibili cause di eventuali danni.

La Relazione evidenzia le molteplici opportunità che l'intelligenza artificiale può portare con sé[203], non potendo però evitare di prendere in considerazione i potenziali effetti pregiudizievoli che possono derivare per una serie di interessi giuridicamente rilevanti, da ciò scaturendo la necessità di vagliare i quadri normativi attualmente vigenti in materia di sicurezza e responsabilità per comprendere se necessitino di specifiche integrazioni.

La Commissione procede poi ad affrontare il primo dei due argomenti fatti oggetto della presente Relazione, ossia quello della *sicurezza*. Passando in rassegna la normativa dell'Unione in materia di sicurezza dei prodotti, la quale «mira a garantire che i prodotti immessi sul mercato dell'Unione soddisfino requisiti elevati in materia di salute, sicurezza e ambiente e che tali prodotti possano circolare liberamente in tutta l'Unione»[204], la Commissione inizia a chiedersi se l'attuale assetto normativo sia in grado di fronteggiare le nuove tecnologie.

Per tentare di fornire una risposta a questo interrogativo la Commissione inizia a passare in rassegna le caratteristiche dell'intelligenza artificiale, in particolare di quelle che possono influenzare il concetto di sicurezza[205]:

- *Connettività*: il concetto di sicurezza merita di essere rivisto alla luce dell'interconnessione esistente tra i dispositivi intelligenti. La perdita della connettività da parte di queste tecnologie può comportare rischi concreti per l'incolumità delle persone che con esse interagiscono[206], con conseguenti ricadute nell'individuazione del responsabile di un potenziale evento lesivo. Nel concetto di sicurezza di tali prodotti rientra la protezione non soltanto dai rischi collegati alla perdita di connettività ma anche dai rischi informatici e da quelli riferibili ad un uso improprio (per quanto prevedibile) del prodotto. In tale ottica una previsione normativa specifica consentirebbe di garantire la certezza del diritto e di realizzare una maggiore tutela per i consumatori;
- *Autonomia*: i sistemi dotati di intelligenza artificiale potrebbero essere in grado di assumere determinazioni non predeterminabili, che pertanto sfuggono al

[203] La Commissione evidenzia le potenzialità dell'uso dell'IA in ambito sanitario, nel settore della cybersicurezza e dei cambiamenti climatici, nonché in quello della circolazione stradale. Con particolare riferimento a quest'ultimo aspetto, richiamando la Relazione della Commissione dal titolo "Salvare vite umane: migliorare la sicurezza dei veicoli nell'UE" considera che «i veicoli connessi e automatizzati potrebbero migliorare la sicurezza stradale, dato che la maggior parte degli incidenti stradali è attualmente causata da errore umano» *Relazione sulle implicazioni dell'intelligenza artificiale*, cit., p. 3.

[204] *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., p. 4.

[205] *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., pp. 6 ss.

[206] «Ad esempio, se un allarme antincendio collegato perde la connettività, potrebbe non avvertire l'utilizzatore in caso di incendio». A sua volta l'interconnettività di questi sistemi può essere utilizzata per scopi illeciti: «La radio del veicolo può presentare lacune nella *security* del software che consentono l'accesso non autorizzato da parte di terzi ai sistemi di controllo interconnessi del veicolo. Se dette lacune fossero sfruttate da terzi per scopi dolosi, potrebbe verificarsi un incidente stradale» *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., p. 6.

controllo umano, così potendo cagionare un danno agli utenti che se ne servono. La valutazione del rischio relativa al prodotto effettuata preliminarmente rispetto all'immissione di quest'ultimo nel mercato potrebbe, pertanto, non essere attendibile: il comportamento originariamente previsto dal progettista potrebbe modificarsi, in tal modo non verrebbero più rispettati gli originari requisiti di sicurezza e si imporrebbe una nuova valutazione del prodotto in grado di apprendere autonomamente[207]. Tale caratteristica rischia di porsi in contrasto con la necessità che l'uomo debba sempre mantenere il controllo sul sistema intelligente: per tale ragione la normativa dell'Unione dovrebbe prevedere specifichi obblighi in materia, specie per quanto concerne la sorveglianza umana, non soltanto durante la fase di progettazione del sistema, ma anche durante il suo intero ciclo vitale[208];

- *Dipendenza dai dati*: i sistemi basati sull'IA si alimentano di dati ed è importante che questi ultimi siano accurati se si vuole garantire la correttezza della decisione assunta dal sistema. Chiaramente la normativa attualmente vigente non si occupa dei rischi per la sicurezza connessi all'inserzione di dati errati, pertanto sarebbe opportuno prevedere specifiche disposizioni volte garantire un'elevata qualità dei dati durante l'utilizzo del sistema;
- *Opacità*: ormai costituisce dato acquisito che i sistemi intelligenti sono in grado di apprendere dall'esperienza, tuttavia è spesso arduo (se non forse impossibile) ricostruire il procedimento con cui ciò avviene (c.d. *black box effect*). La comprensione di tale meccanismo – ed il relativo controllo *ex post* – appare invece di fondamentale importanza, specie nel caso in cui tali sistemi intelligenti vengano usati in settori delicati. Stabilire requisiti di trasparenza, solidità e responsabilità risulta dunque necessario al fine di poter alimentare la fiducia in tali sistemi e di verificare il rispetto della normativa di settore;
- *Complessità dei prodotti e dei sistemi*: nell'effettuare la valutazione dei potenziali rischi del sistema intelligente il produttore deve considerare non soltanto l'utilizzo previsto e prevedibile dello strumento dotato di IA, ma anche il prevedibile uso scorretto che di questo si possa fare. Inoltre tale valutazione del rischio non dovrà riguardare esclusivamente il prodotto in sé, bensì anche il software[209] che in esso viene integrato prima dell'immissione sul mercato. Le

[207] «Una simile valutazione del rischio è già obbligatoria nella normativa in materia di trasporti; ad esempio, nel trasporto ferroviario, la normativa prevede che, in caso di modifica di un veicolo ferroviario dopo che quest'ultimo è stato certificato, il soggetto che introduce la modifica deve seguire una procedura specifica e attenersi a criteri chiari e definiti per determinare se occorra rivolgersi alle autorità» *Relazione sulle implicazioni dell'intelligenza artificiale*, cit., p. 8.

[208] «Nella pertinente normativa dell'Unione si potrebbero introdurre obblighi espliciti, anche a carico dei produttori di robot umanoidi dotati di intelligenza artificiale, di tener conto esplicitamente dei danni immateriali che i loro prodotti potrebbero causare agli utilizzatori, in particolare agli utilizzatori vulnerabili come le persone anziane in contesti di cura» *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., p. 9. Un precedente riferimento ai danni immateriali viene altresì svolto nel corpo della Comunicazione della Commissione dal titolo "Creare fiducia nell'intelligenza artificiale antropocentrica" ove si chiarisce che «I sistemi di IA dovrebbero inoltre contenere meccanismi di sicurezza fin dalla progettazione, per garantire che siano sicuri in modo verificabile in ogni fase, considerando soprattutto la sicurezza fisica e mentale di tutte le persone coinvolte» *Creare fiducia nell'intelligenza artificiale antropocentrica*, cit., p. 5.

[209] «Nella normativa dell'Unione in materia di sicurezza dei prodotti, gli aggiornamenti del software potrebbero essere assimilati a interventi di manutenzione per motivi di sicurezza, purché non modifichino in misura significativa il prodotto già immesso sul mercato e non introducano

tecnologie fondate sull'intelligenza artificiale coinvolgono per il loro funzionamento "catene di valore complesse" eppure, sulla scorta dell'attuale normativa, «la responsabilità della sicurezza del prodotto rimane a carico del produttore che immette il prodotto sul mercato»[210].

Successivamente la Commissione affronta il secondo punto in oggetto alla presente Relazione e che, probabilmente, rappresenta l'argomento di maggior interesse per il penalista, ossia quello della *responsabilità*. La Commissione apre così un'interessante riflessione evidenziando la complementarità tra le disposizioni concernenti la sicurezza dei prodotti e la responsabilità per danno da prodotti difettosi, chiarendo come il loro scopo sia quello di operare congiuntamente al fine di creare un mercato unico in grado di garantire elevati livelli di sicurezza, di minimizzare i rischi per gli utilizzatori e di assicurare il risarcimento degli eventuali danni.

In tal sede la Commissione dedica maggiore attenzione alla materia della responsabilità civile[211], evidenziando altresì che la direttiva sulla responsabilità per danno da prodotti difettosi «introduce un regime di responsabilità oggettiva del produttore per i danni causati dai difetti del prodotto»[212], dunque indipendentemente dalla colpa del produttore medesimo.

La Commissione distingue a questo punto i regimi di responsabilità per colpa – ove grava sulla persona offesa l'onere di dimostrare (al fine di ottenere il risarcimento del danno) la colpa del produttore, la sussistenza del danno ed il nesso di causalità tra questi due elementi – da quelli di responsabilità oggettiva, in

---

nuovi rischi non previsti nella valutazione del rischio iniziale. Tuttavia, se l'aggiornamento del software modifica in misura sostanziale il prodotto in cui viene scaricato, l'intero prodotto potrebbe essere considerato un nuovo prodotto e la conformità alla pertinente normativa in materia di sicurezza dei prodotti deve essere rivalutata al momento della modifica» *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., p. 11.

[210] *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., p. 12. La Commissione specifica nel prosieguo che «La normativa dell'Unione in materia di sicurezza dei prodotti tiene conto della complessità delle catene di valore, imponendo obblighi a diversi operatori economici secondo il principio della "responsabilità condivisa". Sebbene la responsabilità del produttore per la sicurezza del prodotto finale si sia rivelata adeguata per le complesse catene di valore attuali, disposizioni esplicite che impongano specificamente la cooperazione tra gli operatori economici nella catena di approvvigionamento e gli utilizzatori potrebbero creare certezza giuridica in catene di valore forse ancora più complesse. In particolare, ogni partecipante alla catena di valore avente un impatto sulla sicurezza del prodotto (ad esempio i produttori di software) e sugli utilizzatori (ad esempio, se modifica il prodotto) si assumerebbe la propria responsabilità e fornirebbe al partecipante successivo nella catena le informazioni e le misure necessarie» p. 13.

[211] A tal riguardo la Relazione chiarisce che «Le norme in materia di responsabilità civile hanno una duplice funzione nella nostra società: da una parte, garantiscono che quanti hanno subito un danno causato da altri ottengano il risarcimento e, dall'altra, creano un incentivo economico per spingere la parte responsabile a evitare di causare il danno. Le norme in materia di responsabilità devono sempre trovare un equilibrio, tutelando i cittadini dai danni e consentendo allo stesso tempo alle imprese di innovare» *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., p. 13.

[212] *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., p. 13. Riteniamo che in tal sede la Commissione voglia fare riferimento all'inciso della direttiva in questione, la quale sancisce che «solo la responsabilità del produttore, indipendente dalla sua colpa, costituisce un'adeguata soluzione del problema, specifico di un'epoca caratterizzata dal progresso tecnologico, di una giusta attribuzione dei rischi inerenti alla produzione tecnica moderna» *Direttiva del Consiglio relativa al ravvicinamento delle disposizioni legislative, regolamentari ed amministrative degli stati membri in materia di responsabilità per danno da prodotti difettosi* (85/374/CEE), 25.7.1985.

base ai quali la responsabilità del rischio viene attribuita ad un dato soggetto senza che gravi alcun onere sulla persona offesa.

Sempre ad avviso della Commissione, ove si dovesse opzionare un regime di responsabilità colposa, le caratteristiche delle tecnologie dotate di IA renderebbero difficile individuare il comportamento umano causativo del danno e, conseguentemente, riuscire ad articolare una domanda di risarcimento danni. Il tutto alla luce della necessità di garantire alle vittime di incidenti collegati all'uso di sistemi intelligenti un livello di tutela analogo a quello garantito rispetto a prodotti e servizi non integrati da simili tecnologie.

In tal sede la Commissione richiama due concetti, invero, squisitamente penalistici: la responsabilità oggettiva e la responsabilità colposa. Tali due meccanismi di imputazione della responsabilità si connotano per una intrinseca problematicità, specie se rapportati alla materia dell'intelligenza artificiale. La responsabilità oggettiva, da un lato, non dovrebbe trovare cittadinanza nel nostro ordinamento (chiaramente in termini penalistici) ma a tratti sembra essere la sola via percorribile, specie in considerazione dello stretto connubio che si instaura nella materia *de qua* tra la responsabilità civilistica del danno da prodotto difettoso e l'interrogativo penalistico dell'attribuzione della responsabilità in caso di un reato commesso dall'IA. La responsabilità colposa, di contro, porta con sé l'incertezza dell'individuazione delle regole precauzionali dalla cui violazione deriverebbe una responsabilità per colpa.

La Relazione in esame evidenzia poi la necessità di valutare se gli attuali quadri normativi possano adattarsi alle evoluzioni tecnologiche oppure se la loro applicazione genererebbe incertezza dal punto di vista giuridico. Bisogna evitare che i singoli Stati membri diano vita a disposizioni frammentarie in materia di responsabilità, riservando alle imprese del settore un quadro normativo chiaro che consenta loro di ridurre e prevenire i rischi insiti nel corso della catena di funzionamento del sistema.

Attualizzando il problema, la Commissione evidenzia che, applicando la direttiva sulla responsabilità per danno da prodotti difettosi a sistemi software o dotati di intelligenza artificiale, si finirebbe per attribuire la responsabilità di eventuali danni al fabbricante. In realtà, però, l'ambito di applicazione della direttiva in parola andrebbe ulteriormente chiarito e specificato alla luce dell'evoluzione scientifica e della complessità dei prodotti dotati di IA.

Come si è cercato di accennare in precedenza, la complessità di questi sistemi, unita alla moltitudine di professionisti coinvolti nel loro funzionamento, rendono arduo individuare l'origine del danno ed il soggetto cui poter imputare la responsabilità (e, aggiungeremmo, a che titolo imputare siffatta responsabilità). Conseguentemente, a fronte di cotante incertezze, risulterà altrettanto complesso per le vittime strutturare una domanda di risarcimento del danno.

In tal sede la Commissione pone un interessante interrogativo per il penalista: «in che modo la nozione di colpa si applicherebbe ai danni causati dall'intelligenza artificiale?»[213].

La Commissione prende in considerazione l'evenienza di adeguare le normative nazionali in modo da agevolare l'onere probatorio dei soggetti che hanno patito un danno a causa del malfunzionamento di un sistema intelligente, ad

[213] *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., p. 15.

esempio, prevedendo specifici obblighi in materia di sicurezza la cui mancata osservanza comporterebbe la relativa modifica dell'onere probatorio a carico della persona offesa in ordine alla colpa o al nesso causale.

Le problematiche relative alla sicurezza sono tuttavia destinate a mutare alla luce della connettività che caratterizza tali sistemi. Si evidenzia però in tal sede l'importanza di tracciare un quadro giuridico chiaro in materia di cybersicurezza al fine di consentire agli operatori di settore di prevedere ed evitare (nell'ottica di un'imputazione colposa) di incorrere in responsabilità per inosservanza delle relative norme[214].

La Commissione, con una riflessione di particolare interesse, considera importante comprendere se il produttore, tenendo conto dell'uso ragionevole del sistema intelligente, avrebbe potuto prevedere eventuali modifiche del prodotto. A tal fine vengono distinti due sistemi di difesa applicabili in tale campo:

- *Difesa basata sul difetto successivo*: «il produttore non è responsabile se il difetto non esisteva al momento in cui il prodotto è stato messo in circolazione»;
- *Difesa basata sui rischi di sviluppo*: «le migliori conoscenze del momento non consentivano di prevedere il difetto»[215].

La Relazione propone inoltre un ulteriore strumento da poter adoperare nel settore – sempre nell'ottica di un utilizzo ragionevolmente prevedibile – ossia il concorso di colpa della vittima (che, in una prospettiva penalistica *de iure condendo*, potrebbe consentire di ridurre la pena riservata al produttore, principale responsabile del danno) la quale, ad esempio, avrebbe potuto evitare o ridurre il danno patito ove avesse aggiornato il software come richiesto dal sistema stesso. Una circostanza di tal fatta renderebbe maggiormente difficoltoso ottenere il risarcimento del danno da parte della vittima ma tutelerebbe il produttore, ove egli avesse effettivamente rispettato i propri doveri normativamente imposti.

Tornando sul tema dell'autonomia, la Commissione richiama la possibilità che i sistemi dotati di IA siano in grado di agire in modo indipendente o di portare a termine determinati compiti senza la supervisione umana. Alla luce dell'opacità che connota il funzionamento di questi sistemi, richiedere un risarcimento del danno non appare cosa da poco: «non è chiaro in che modo si possa dimostrare la colpa di un'intelligenza artificiale che agisce autonomamente, né in che cosa potrebbe consistere la colpa della persona che utilizza l'intelligenza artificiale»[216]. Alla luce dell'apprendimento automatico di cui sono dotati i più sofisticati sistemi intelligenti sarebbe utile tracciare una linea di demarcazione e stabilire fin dove il produttore sia in grado di prevedere i possibili sviluppi dell'IA e fin quando permanga la responsabilità di quest'ultimo (da quale momento in poi, dunque, il produttore non può più essere considerato responsabile dei danni cagionati

[214] *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., pp. 15-16.
[215] *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., p. 17.
[216] *Relazione sulle implicazioni dell'intelligenza artificiale,* cit., p. 17. La Relazione prosegue sul tema affermando che «Nella normativa dell'Unione in materia di sicurezza dei prodotti e di responsabilità per danno da prodotti difettosi il principio guida è che spetta ai produttori garantire che tutti i prodotti immessi sul mercato siano sicuri, lungo tutto il ciclo di vita e per l'uso del prodotto ragionevolmente prevedibile. Ciò significa che il fabbricante dovrebbe garantire che il prodotto che utilizza l'intelligenza artificiale rispetti determinati parametri di sicurezza. Le caratteristiche dell'intelligenza artificiale non precludono il diritto a pretendere che i prodotti siano sicuri, che si tratti di tosaerba automatici o di robot chirurgici».

dall'IA), chiarendo, una volta per tutte, chi sia il responsabile di eventuali modifiche del prodotto.

Alcuni sistemi di IA potrebbero presentare specifici profili di rischio, specialmente dal punto di vista della responsabilità, ove vengano utilizzati in contesti in cui possano essere posti in pericolo beni giuridici di primaria rilevanza come la vita e l'incolumità fisica. La Commissione evidenzia come una possibile via per affrontare tali nuove problematiche potrebbe essere quella di ricorrere ad un approccio basato sul rischio, spendendo parole positive nei riguardi del regime di responsabilità oggettiva, il quale garantirebbe il risarcimento della vittima al concretizzarsi del rischio, indipendentemente dalla colpa del soggetto sul quale si sceglierebbe di far ricadere la responsabilità oggettiva (*rectius*, responsabilità per posizione).

In conclusione la Relazione, dopo aver passato in rassegna i possibili rischi connessi al rispetto della sicurezza e della responsabilità alla luce delle caratteristiche tipiche dell'intelligenza artificiale, richiama la necessità di colmare eventuali lacune del sistema, garantendo un livello di tutela analogo a quello riservato alle vittime delle c.d. tecnologie tradizionali, anche al fine di alimentare la fiducia nell'IA, fronteggiando le difficoltà connesse alla stesura di una domanda di risarcimento danni e all'adozione di un approccio basato sul rischio o, meglio, sui rischi specifici di ciascun sistema intelligente.

SEZIONE SECONDA
DE IURE CONDENDO

9. Prospettive de iure condendo.

Una volta passati in rassegna i più significativi testi attualmente in vigore nel contesto dell'Unione Europea, si ritiene opportuno vagliare due proposte (rispettivamente della Commissione e del Parlamento europeo) di recentissimo conio.

9.1. La Proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione.

Il 21 aprile 2021 la Commissione Europea ha pubblicato una Proposta di Regolamento volta a stabilire regole armonizzate in materia di intelligenza artificiale, ponendo una serie di obiettivi ambiziosi, quali garantire la sicurezza dei sistemi di IA immessi sul mercato dell'Unione, assicurare la certezza del diritto e il rispetto della normativa attualmente vigente in materia di diritti fondamentali, nonché prevenire la frammentazione del mercato agevolando lo sviluppo di un mercato unico[217].

[217] *Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione*, Bruxelles, 21.4.2021, p. 3. La Commissione individua la base giuridica dell'esaminanda Proposta nell'art. 114 del Trattato sul funzionamento dell'Unione europea (TFUE) il quale stabilisce che «Il Parlamento europeo e il Consiglio, deliberando secondo la procedura legislativa ordinaria e previa consultazione del Comitato economico e sociale, adottano le misure relative al ravvicinamento delle disposizioni legislative, regolamentari ed amministrative

Tali obiettivi non possono essere perseguiti adeguatamente dai singoli Stati membri, rendendosi piuttosto necessario un intervento dell'Unione sul punto. L'affastellarsi di differenti normative in materia, difatti, osterebbe alla libera circolazione dei prodotti e servizi dotati di IA, senza riuscire a garantire un'omogenea protezione dei diritti fondamentali.

La Proposta mira a realizzare un quadro normativo proporzionato, solido (ma al contempo flessibile) ed essenziale, volto a individuare i requisiti minimi per affrontare i rischi connessi ai sistemi intelligenti e ad adeguarsi allo sviluppo tecnologico senza, con questo, ostacolarlo. A tal fine si individuano i requisiti obbligatori comuni che devono essere rispettati dai sistemi di IA e si stabiliscono regole volte a fronteggiare la fase successiva all'immissione sul mercato di detti prodotti mediante meccanismi di controllo *ex post*.

Si tenta, con tale strumento legislativo, di individuare disposizioni armoniche in grado di regolamentare lo sviluppo, la commercializzazione e l'uso dell'IA, fornendo una definizione unitaria di quest'ultima e seguendo un approccio *basato sul rischio*.

La Proposta mira, a sua volta, ad armonizzarsi e integrarsi con gli altri testi normativi attualmente vigenti in Europa e con la normativa di settore in materia di sicurezza dei prodotti, riducendo i rischi di discriminazione algoritmica ed evitando duplicazioni o oneri aggiuntivi. La scelta del Regolamento come atto giuridico più idoneo è dettata dalla necessità di garantire un'applicazione diretta e uniforme delle norme in materia, così evitando fenomeni di frammentazione giuridica[218]. La Commissione ha, dunque, deciso di servirsi di uno strumento legislativo orizzontale che sia proporzionato e basato sul rischio, coadiuvandolo con specifici codici di condotta per i sistemi di intelligenza artificiale considerati non ad alto rischio[219] da adottare su base volontaria.

La Proposta in esame appare consapevole della possibilità che i sistemi di IA, per le loro intrinseche caratteristiche, possano incidere sui diritti fondamentali. L'individuazione dei requisiti necessari per un'IA affidabile e degli obblighi che i membri della catena di valore del sistema devono rispettare, unita ad un serio monitoraggio sul corretto funzionamento di tali sistemi, servirà a promuovere il rispetto di siffatti diritti[220].

---

degli Stati membri che hanno per oggetto l'instaurazione ed il funzionamento del mercato interno» *Proposta di Regolamento,* cit., p. 6.

[218] *Proposta di Regolamento,* cit., pp. 3 ss.

[219] La citata opzione costituiva la n. 3+. Le altre proposte ivi contenute riguardavano: «opzione 1: strumento legislativo dell'UE che istituisce un sistema di etichettatura volontario; opzione 2: approccio settoriale "ad hoc"; opzione 3: strumento legislativo orizzontale dell'UE che segue un approccio proporzionato basato sul rischio; opzione 4: strumento legislativo orizzontale dell'UE che stabilisce requisiti obbligatori per tutti i sistemi di IA, indipendentemente dal rischio che pongono» *Proposta di Regolamento,* cit., p. 10.

[220] Tra di essi si ricordano, a titolo esemplificativo, il diritto alla dignità umana, al rispetto della vita privata e alla non discriminazione. La Commissione sancisce inoltre che «Gli *obblighi di prova ex ante*, di gestione dei rischi e di sorveglianza umana faciliteranno altresì il rispetto di altri diritti fondamentali, riducendo al minimo il rischio di decisioni errate o distorte assistite dall'IA in settori critici quali l'istruzione e la formazione, l'occupazione, servizi importanti, le attività di contrasto e il sistema giudiziario. Nel caso in cui si verifichino comunque violazioni dei diritti fondamentali, un ricorso efficace a favore delle persone lese sarà reso possibile assicurando la trasparenza e la tracciabilità dei sistemi di IA unitamente a rigidi *controlli ex post*» *Proposta di Regolamento,* cit., p. 12, corsivo nostro.

In sede introduttiva la Proposta evidenzia l'importanza di garantire un elevato livello di tutela per interessi pubblici e diritti fondamentali, specie per quanto concerne i sistemi di intelligenza artificiale c.d. "ad alto rischio", chiarendo come questo obiettivo possa essere raggiunto con una legislazione comune, coerente e non discriminatoria che regolamenti questi sistemi. La Commissione prende poi in considerazione la possibilità che tali sistemi vengano utilizzati impropriamente con pratiche dannose o pericolose per i diritti e i valori dell'Unione.

Sotto il profilo strettamente penalistico la Proposta richiama in tal sede «pratiche di manipolazione, sfruttamento e controllo sociale»[221]. Pur non entrando nello specifico, a prima vista, sembra che condotte di tal fatta possano considerarsi particolarmente gravi, al punto da giustificare una sanzione penale al fine di contrastarle, pur essendo di difficile accertamento probatorio. La Commissione evidenzia altresì la necessità di contemperare lo sviluppo della ricerca scientifica con l'indispensabile cautela di non immettere sul mercato prodotti in grado di cagionare danni materiali o immateriali. A tal fine la Proposta di Regolamento individua determinati requisiti obbligatori che i sistemi di IA ad alto rischio devono soddisfare per poter essere immessi sul mercato al fine prevenire potenziali rischi per la sicurezza[222].

Nel cominciare a delineare gli aspetti dei sistemi intelligenti ad alto rischio la Commissione – oltre a menzionare i sistemi di identificazione biometrica remota "in tempo reale" che saranno approfonditi più avanti – chiarisce che possano essere annoverati in tale categoria anche quei sistemi di IA «destinati a essere utilizzati come componenti di sicurezza ai fini della gestione del traffico stradale nonché della fornitura di acqua, gas, riscaldamento ed elettricità, in quanto un loro guasto o malfunzionamento può mettere a rischio la vita e la salute di un grande numero di persone»[223]. Anche qui il richiamo alla materia penalistica, determinato dal rischio per la vita e la salute dei consociati, in particolare se connessa alla materia della circolazione stradale, appare evidente.

Analogo collegamento penalistico emerge nel momento in cui la Commissione evidenzia il rischio che l'uso di sistemi intelligenti non sufficientemente trasparenti o spiegabili – ove impiegati in ambito giudiziario, nel

[221] *Proposta di Regolamento,* cit., Punto 15, p. 23. In particolare, la pratica del "controllo sociale" sembra ricordare le venature dell'art. 603 c.p. – ormai espunto dal nostro codice penale – il quale recitava «Chiunque sottopone una persona al proprio potere, in modo da ridurla in totale stato di soggezione, è punito con la reclusione da cinque a quindici anni». La Corte Costituzionale ha dichiarato l'illegittimità costituzionale di tale norma con sent. 8.6.1981, n. 96 in quanto contrastante con il principio di tassatività della fattispecie e, pertanto, con l'art. 25 comma 2 Cost.

[222] «Ad esempio, i robot sempre più autonomi, sia nel contesto della produzione sia in quello della cura e dell'assistenza alle persone, dovrebbero essere in misura di operare e svolgere le loro funzioni in condizioni di sicurezza in ambienti complessi. Analogamente, nel settore sanitario, in cui la posta in gioco per la vita e la salute è particolarmente elevata, è opportuno che i sistemi diagnostici e i sistemi di sostegno delle decisioni dell'uomo, sempre più sofisticati, siano affidabili e accurati. La portata dell'impatto negativo del sistema di IA sui diritti fondamentali protetti dalla Carta è di particolare rilevanza ai fini della classificazione di un sistema di IA tra quelli ad alto rischio. Tali diritti comprendono il diritto alla dignità umana, il rispetto della vita privata e della vita familiare, la protezione dei dati personali, la libertà di espressione e di informazione, la libertà di riunione e di associazione e la non discriminazione, la protezione dei consumatori, i diritti dei lavoratori, i diritti delle persone con disabilità, il *diritto a un ricorso effettivo* e a un *giudice imparziale*, i *diritti della difesa* e la *presunzione di innocenza* e il diritto a una buona amministrazione» *Proposta di Regolamento,* cit., Punto 28, p. 26, corsivo nostro.

[223] *Proposta di Regolamento,* cit., Punto 34, p. 29.

quale andrebbe garantita l'affidabilità e l'accuratezza – possa ledere il diritto di difesa, il diritto a un ricorso effettivo e il diritto ad esser giudicati da un giudice imparziale, nonché il principio della presunzione di innocenza[224]. I sistemi di IA adoperati in questo delicato settore possono impattare sulle libertà individuali e sulla democrazia stessa, pertanto appare necessario prevenire distorsioni e opacità, senza però rinunciare all'uso di sistemi che, seppur adoperati nell'ambito dell'amministrazione della giustizia, non comportano rischi per i succitati beni (come, ad esempio, le pratiche di anonimizzazione delle decisioni giudiziarie)[225].

La Proposta richiama altresì l'importanza di addestrare il sistema con dati pertinenti e rappresentativi, di fronteggiare l'opacità imponendo un certo grado di trasparenza che consenta agli utilizzatori di questi sistemi di comprenderne gli output, nonché di individuare adeguate misure di sorveglianza umana affidandole a persone dotate della competenza e della formazione necessaria[226].

Ulteriore richiamo alla materia penalistica si ha nel momento in cui la Commissione chiarisce la necessità che questi sistemi siano in grado di resistere e rispondere ad azioni dolose idonee a comprometterne la sicurezza e da cui possono scaturire conseguenze dannose. Ci si riferisce alle condotte di terzi malintenzionati che potrebbero sfruttare le falle di questi sistemi per alterarne l'uso o comprometterne la sicurezza[227].

Sempre da un punto di vista squisitamente penalistico, la Proposta ritiene opportuno individuare una specifica persona, fisica o giuridica, anche nota come "fornitore"[228], che si assuma la responsabilità dell'immissione sul mercato del sistema intelligente ad alto rischio, a prescindere dal fatto che questa persona abbia o meno progettato il sistema[229]. Se da un lato tale inciso richiama alla mente la figura del soggetto che riveste una posizione di garanzia, secondo le comuni regole dell'art. 40 cpv. c.p.[230], d'altro canto sembra richiedere l'individuazione di

[224] «In considerazione della natura delle attività in questione e dei rischi a esse connessi, tra tali sistemi di IA ad alto rischio è opportuno includere, in particolare, i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per valutazioni dei rischi individuali, come poligrafi e strumenti analoghi, oppure per rilevare lo stato emotivo delle persone fisiche, individuare "deep fake", valutare l'affidabilità degli elementi probatori nei procedimenti penali, prevedere il verificarsi o il ripetersi di un reato effettivo o potenziale sulla base della profilazione delle persone fisiche, o valutare i tratti e le caratteristiche della personalità o il comportamento criminale pregresso delle persone fisiche o dei gruppi, nonché ai fini della profilazione nel corso dell'indagine, dell'accertamento e del perseguimento di reati e dell'analisi criminale nei riguardi delle persone fisiche» *Proposta di Regolamento,* cit., Punto 38, p. 30.

[225] *Proposta di Regolamento,* cit., Punto 40, p. 31.

[226] *Proposta di Regolamento,* cit., Punti 44-47-48, pp. 32-33.

[227] *Proposta di Regolamento,* cit., Punti 50-51, p. 33. «Gli attacchi informatici contro i sistemi di IA possono far leva sulle risorse specifiche dell'IA, quali i set di dati di addestramento (ad esempio "avvelenamento dei dati", *data poisoning*) o i modelli addestrati (ad esempio "attacchi antagonisti", *adversarial attacks*), o sfruttare le vulnerabilità delle risorse digitali del sistema di IA dell'infrastruttura TIC sottostante. Al fine di garantire un livello di cibersicurezza adeguato ai rischi, è pertanto opportuno che i fornitori di sistemi di IA ad alto rischio adottino misure adeguate, anche tenendo debitamente conto dell'infrastruttura TIC sottostante».

[228] In capo a quest'ultimo ricadrebbe l'obbligo di istituire un sistema di gestione della qualità, di valutazione di conformità e di monitoraggio, dovendo altresì garantire che il sistema soddisfi i requisiti sanciti dalla Proposta di Regolamento prima di essere immesso sul mercato. *Proposta di Regolamento,* cit., Punti 54-55, p. 34.

[229] *Proposta di Regolamento,* cit., Punto 53, p. 34.

[230] F.C. LA VATTIATA, *Brevi note "a caldo" sulla recente Proposta di Regolamento UE in tema di intelligenza artificiale,* in *Diritto Penale e Uomo,* 30.6.2021, p. 13.

una figura che risponda a prescindere da qualsivoglia nesso, non soltanto eziologico rispetto all'evento, ma anche rispetto al legame con il sistema intelligente, finendo così per delineare una forma di responsabilità oggettiva che non dovrebbe trovare cittadinanza in materia penalistica.

Il fornitore non sembra però essere il solo soggetto in capo al quale ricondurre eventuali responsabilità, in quanto si specifica che gli utenti sono tenuti a servirsi del sistema di IA conformemente alle relative istruzioni per l'uso[231], potendo da ciò dedurre *a contrario* che, in caso di un utilizzo non conforme del sistema intelligente da parte dell'utilizzatore, quest'ultimo potrebbe essere chiamato a rispondere di eventuali danni cagionati dall'IA sottoposta alla sua autorità.

Altre conseguenze penalmente rilevanti potrebbero derivare dalla mancata valutazione di conformità da effettuare prima dell'immissione del prodotto sul mercato, la quale serve a garantire un elevato livello di affidabilità del sistema. La mancata effettuazione di tale valutazione potrebbe integrare gli estremi della mancata osservanza di una regola cautelare, ove tale impegno venisse normativizzato[232]. Tale valutazione dovrebbe essere effettuata dal fornitore – ad eccezione dei sistemi di IA per l'identificazione biometrica remota, per i quali è necessario coinvolgere specifici organismi c.d. notificati[233] – e andrebbe rinnovata ogniqualvolta intervenga una modifica in grado di incidere sulla conformità o sulle finalità del sistema, come potrebbe avvenire, ad esempio, per i sistemi capaci di autoapprendimento[234].

La Commissione menziona gli ormai più volte citati "spazi di sperimentazione normativa", evidenziando la necessità di stabilire regole comuni per la loro attuazione, proponendo il Regolamento in esame come base giuridica. Spetterebbe alle autorità competenti prevedere eventuali sanzioni amministrative pecuniarie per sanzionare comportamenti scorretti dei partecipanti ai detti spazi di sperimentazione[235].

Al fine di assicurare la concreta adozione delle misure sancite dall'esaminando Regolamento potrebbe essere utile per gli Stati prevedere, in caso di violazione di suddette misure, sanzioni effettive, proporzionate e dissuasive. Nell'ottica di un diritto penale inteso come *extrema ratio* del sistema, probabilmente una sanzione penale in tale ambito potrebbe, in generale, non rendersi necessaria. Sarebbe forse più opportuno andare a scandagliare il caso concreto in quanto, magari, una violazione particolarmente grave e con conseguenze lesive per soggetti terzi potrebbe anche rendere l'intervento penalistico giustificato[236].

Prima di entrare nel vivo della Proposta Regolamentare, la Commissione chiarisce come quest'ultima miri a rispettare i principi di proporzionalità e sussidiarietà, i quali caratterizzano la tecnica normativa dell'Unione, chiarendo altresì la necessità che gli Stati stabiliscano un'apposita normativa in materia di

---

[231] *Proposta di Regolamento,* cit., Punto 58, p. 35.
[232] *Proposta di Regolamento,* cit., Punto 62, p. 35.
[233] *Proposta di Regolamento,* cit., Punto 64, p. 36. Secondo la definizione fornita dalla Proposta di Regolamento in esame l'organismo notificato sarebbe «un organismo di valutazione della conformità designato in conformità al presente regolamento e ad altre pertinenti normative di armonizzazione dell'Unione» *Proposta di Regolamento,* cit., art. 3, n. 22, p. 44.
[234] *Proposta di Regolamento,* cit., Punto 66, p. 36.
[235] *Proposta di Regolamento,* cit., Punto 72, pp. 37-38.
[236] *Proposta di Regolamento,* cit., Punto 84, p. 40.

sanzioni – richiamando espressamente le sanzioni ammnistrative pecuniarie – per le violazioni del presente Regolamento[237], senza dunque sbilanciarsi in ordine ad altre tipologie di sanzioni, eventualmente anche penali (?).

Il Titolo I della Proposta individua l'ambito di applicazione[238] del testo medesimo e fornisce una serie di norme definitorie utili per la corretta interpretazione del Regolamento.

Il Titolo II individua le pratiche di intelligenza artificiale vietate, annoverando tra queste i sistemi in grado di influenzare il comportamento delle persone, o di un gruppo di persone, al punto da provocare loro un danno fisico o psicologico e sistemi utilizzati dalle autorità pubbliche al fine di classificare le persone in base alle loro condotte sociali.

La pratica vietata che, in ottica penalistica, suscita maggior interesse è quella concernente i sistemi di identificazione biometrica remota "in tempo reale" in spazi accessibili al pubblico a fini di attività di contrasto. L'uso di tali sistemi viene autorizzato, in caso di necessità, ove funzionale a ricercare vittime di reato, prevenire minacce per la vita o l'incolumità fisica delle persone o individuare gli autori (o i soggetti sospettati) di uno dei reati di cui all'articolo 2, paragrafo 2, della decisione quadro 2002/584/GAI[239].

In ottica garantistica la Commissione specifica che, nell'assumere determinazioni in ordine all'utilizzo di questi sistemi si debba tener conto della natura della situazione, della gravità, della probabilità e dell'entità del danno, nonché delle conseguenze dell'uso del sistema intelligente per i diritti e le libertà delle persone coinvolte. L'uso di questi sistemi deve essere limitato entro predeterminate condizioni temporali, geografiche e personali, nonché subordinato ad una previa autorizzazione da parte dell'autorità giudiziaria rilasciata dietro motivata richiesta, salvo che non sussistano motivazioni d'urgenza tali da rendere

---

[237] *Proposta di Regolamento,* cit., Punti 87-88, p. 41.

[238] Il Regolamento si rivolge «a) ai fornitori che immettono sul mercato o mettono in servizio sistemi di IA nell'Unione, indipendentemente dal fatto che siano stabiliti nell'Unione o in un paese terzo; b) agli utenti dei sistemi di IA situati nell'Unione; c) ai fornitori e agli utenti di sistemi di IA situati in un paese terzo, laddove l'output prodotto dal sistema sia utilizzato nell'Unione» *Proposta di Regolamento,* cit., art. 2, p. 42.

[239] Decisione quadro del Consiglio 2002/584/GAI, del 13 giugno 2002, relativa al mandato d'arresto europeo e alle procedure di consegna tra Stati membri. Segnatamente l'art. 2, par. 2 della direttiva in parola si riferisce ai reati di «partecipazione a un'organizzazione criminale; terrorismo; tratta di esseri umani; sfruttamento sessuale dei bambini e pornografia infantile; traffico illecito di stupefacenti e sostanze psicotrope; traffico illecito di armi, munizioni ed esplosivi; corruzione; frode, compresa la frode che lede gli interessi finanziari delle Comunità europee ai sensi della convenzione del 26 luglio 1995 relativa alla tutela degli interessi finanziari delle Comunità europee; riciclaggio di proventi di reato; falsificazione di monete, compresa la contraffazione dell'euro; criminalità informatica; criminalità ambientale, compreso il traffico illecito di specie animali protette e il traffico illecito di specie e di essenze vegetali protette; favoreggiamento dell'ingresso e del soggiorno illegali; omicidio volontario; lesioni personali gravi; traffico illecito di organi e tessuti umani; rapimento, sequestro e presa di ostaggi; razzismo e xenofobia; furti organizzati o con l'uso di armi; traffico illecito di beni culturali, compresi gli oggetti d'antiquariato e le opere d'arte; truffa; racket e estorsioni; contraffazione e pirateria in materia di prodotti; falsificazione di atti amministrativi e traffico di documenti falsi; falsificazione di mezzi di pagamento; traffico illecito di sostanze ormonali ed altri fattori di crescita; traffico illecito di materie nucleari e radioattive; traffico di veicoli rubati; stupro; incendio volontario; reati che rientrano nella competenza giurisdizionale della Corte penale internazionale; dirottamento di aereo/nave; sabotaggio».

necessario iniziare ad utilizzare il sistema senza autorizzazione, richiedendola comunque durante o dopo l'uso del sistema medesimo.

L'autorizzazione viene rilasciata solo ove l'autorità giudiziaria accerti che sussistano prove oggettive e che l'uso di tali sistemi sia necessario e proporzionato, spettando a ciascuno Stato membro il compito di stabilire regole dettagliate in ordine alla formulazione della richiesta, alla concessione dell'autorizzazione e allo svolgimento della relativa attività di controllo[240].

Il Titolo III è invece dedicato ai sistemi di intelligenza artificiale considerati "*ad alto rischio*". Si reputano tali i sistemi che, da un lato, vengano utilizzati come componenti di sicurezza di un prodotto disciplinato dalla normativa di armonizzazione dell'Unione[241] e soggetto a una valutazione di conformità da parte di terzi ai fini dell'immissione sul mercato[242] e, dall'altro, i sistemi elencati nell'Allegato III della presente Proposta[243].

Per tali particolari forme di IA viene istituito un sistema di gestione dei rischi[244], ossia un procedimento di controllo costantemente aggiornato e svolto durante l'intero ciclo di vita dell'IA che include fasi pregnanti come l'identificazione dei rischi noti e prevedibili associati a ciascun sistema di IA, nonché la stima dei rischi che possono emergere non solo quando il sistema sia usato conformemente alla sua finalità ma anche nell'ipotesi di un uso improprio ragionevolmente prevedibile[245].

La cifra penalistica viene qui richiamata nel riferimento all'uso improprio ragionevolmente prevedibile. Ci si chiede infatti secondo quali criteri si debba procedere all'individuazione della soglia di prevedibilità o meno di un uso illecito, specie nella misura in cui alcun rimprovero potrebbe essere mosso al fornitore di un sistema ad alto rischio nel caso in cui quest'ultimo venga utilizzato (magari cagionando un danno a terzi) secondo modalità imprevedibili.

Ulteriore cenno penalistico si rinviene nel quarto comma dell'art. 9, ove si sancisce che «Nell'eliminare o ridurre i rischi connessi all'uso del sistema di IA ad alto rischio, si tengono debitamente in considerazione le conoscenze tecniche, l'esperienza, l'istruzione e la formazione che ci si può aspettare dall'utente e l'ambiente in cui il sistema è destinato ad essere usato»[246], così richiamando in un certo senso i limiti penalistici al dovere di diligenza (in tal caso riferibili all'utente del sistema) ed il relativo principio di affidamento.

---

[240] *Proposta di Regolamento,* cit., art. 5, pp. 47-48.

[241] L'Allegato II alla presente Proposta riporta l'elenco della c.d. normativa di armonizzazione dell'Unione, ossia un insieme di testi normativi destinati a coadiuvare la Proposta di Regolamento medesima.

[242] *Proposta di Regolamento,* cit., art. 6, comma 1, p. 49.

[243] Le macroaree incluse nell'elenco di cui all'Allegato III sono così sintetizzabili: identificazione e categorizzazione biometrica delle persone fisiche; gestione e funzionamento delle infrastrutture critiche; istruzione e formazione professionale; occupazione, gestione dei lavoratori e accesso al lavoro autonomo; accesso a prestazioni e servizi pubblici e a servizi privati essenziali e fruizione degli stessi; attività di contrasto; gestione della migrazione, dell'asilo e del controllo delle frontiere; amministrazione della giustizia e processi democratici.

[244] «I sistemi di IA ad alto rischio sono sottoposti a prova al fine di individuare le misure di gestione dei rischi più appropriate. Le prove garantiscono che i sistemi di IA ad alto rischio funzionino in modo coerente per la finalità prevista e che siano conformi ai requisiti di cui al presente capo» *Proposta di Regolamento,* cit., art. 9, comma 5, p. 51.

[245] *Proposta di Regolamento,* cit., art. 9, comma 2, p. 50.

[246] *Proposta di Regolamento,* cit., art. 9, comma 4, p. 51.

Gli artt. 10 ss. elencano i requisiti che i sistemi di IA ad alto rischio devono necessariamente rispettare e che sono così sintetizzabili:

- i dati di addestramento di tali sistemi devono essere pertinenti, rappresentativi, completi e privi di errori, sempre tenendo conto del contesto in cui il sistema verrà utilizzato;
- ciascun sistema deve avere la propria documentazione tecnica idonea a dimostrare che quel prodotto rispetti i requisiti necessari per essere immesso sul mercato;
- tali sistemi devono possedere un meccanismo di registrazione automatica degli eventi, in modo tale da garantire la tracciabilità e il monitoraggio del funzionamento dell'IA;
- i sistemi in esame devono essere dotati di comprensibili istruzioni per l'uso, nonché essere progettati in modo da garantire la trasparenza del proprio funzionamento, consentendo agli utenti di poterli utilizzare adeguatamente e di interpretare i vari output correttamente;
- tali sistemi devono poter essere supervisionati efficacemente dalle persone fisiche[247] al fine di prevenire i rischi per i diritti fondamentali insiti nell'utilizzo (proprio o improprio ma ragionevolmente prevedibile) del sistema intelligente;
- questi sistemi devono garantire un elevato livello di accuratezza, robustezza e cibersicurezza, fronteggiando errori, incongruenze ed eventuali attacchi di terzi malintenzionati. A tal fine appare opportuno prevedere misure atte a evitare simili attacchi, nonché adeguati meccanismi per prevedere usi distorti dei sistemi in grado di autoapprendere dall'esperienza.

Il Titolo in esame prosegue individuando specifici obblighi in capo a tutti i soggetti coinvolti nella gestione dei sistemi di IA ad alto rischio, segnatamente i fornitori[248], gli importatori, i distributori[249] e gli utenti[250]. Ruolo particolare è

[247] Tali soggetti devono essere in grado di comprendere e monitorare il funzionamento del sistema, devono evitare di fare eccessivo affidamento sugli output da questo prodotti e devono essere in condizione di interpretarlo correttamente, devono essere in grado di decidere quando non servirsi del sistema, ignorando o ribaltando l'output da quest'ultimo prodotto e devono, da ultimo, essere in condizione di intervenire per arrestare il sistema in caso di necessità. *Proposta di Regolamento*, cit., art. 14, comma 4, p. 55.

[248] A costoro spetta il compito di garantire che i sistemi di IA siano conformi ai requisiti individuati nel Capo II della presente Proposta e che siano sottoposti alla relativa procedura di valutazione di conformità, da effettuarsi prima dell'immissione sul mercato e da rinnovarsi dopo ogni modifica sostanziale del sistema. I fornitori devono altresì adottare adeguate misure correttive, nel caso in cui i sistemi intelligenti non rispettino più i requisiti summenzionati (eventualmente ritirandoli dal mercato), nonché apporre la relativa marcatura CE per indicare la conformità dei sistemi ai requisiti sanciti dalla presente Proposta. Una volta redatta la dichiarazione di conformità il fornitore si assume la responsabilità della conformità del sistema ai requisiti indicati dal Regolamento, dovendo altresì avere cura di aggiornare la suddetta dichiarazione. In capo ai fornitori spetta altresì il compito di istituire un sistema di gestione della qualità il quale include, tra gli altri, una strategia normativa, un sistema di gestione dei rischi ed un quadro atto a individuare le responsabilità dei soggetti coinvolti nella gestione dell'IA. Da ultimo, prima di procedere all'immissione sul mercato, il fornitore deve registrare il sistema nella relativa banca dati UE. *Proposta di Regolamento*, cit., artt. 16 ss., pp. 56 ss. L'art. 24 individua una analoga responsabilità per i fabbricanti, ove vengano in gioco prodotti cui si applica la normativa armonizzata di cui all'Allegato II, Sezione A.

[249] In capo a costoro spetta il compito di garantire che il fornitore abbia eseguito la valutazione di conformità, redatto la relativa documentazione tecnica, nonché che il sistema rechi la marcatura CE e le istruzioni per l'uso. Ove l'importatore o il distributore dovessero ritenere che un sistema di IA non sia conforme al Regolamento costoro non devono immetterlo sul mercato finché il sistema

attribuito al fornitore, tanto da essere considerato come punto di riferimento dall'art. 28 del presente Regolamento ai sensi del quale distributori, importatori e utenti sono considerati come fornitori e sono soggetti ai relativi obblighi ove: a) immettano sul mercato sistemi di IA ad alto rischio con il loro nome o il loro marchio; b) modifichino le finalità previste per i suddetti sistemi già immessi sul mercato; c) apportino modifiche sostanziali ai sistemi in questione. L'articolo in esame specifica al secondo comma che «Qualora si verifichino le circostanze di cui al paragrafo 1, lettera b) o c), il fornitore che ha inizialmente immesso sul mercato o messo in servizio il sistema di IA ad alto rischio non è più considerato un fornitore ai fini del presente regolamento»[251]. Si individua pertanto il momento in cui nessun rimprovero può più essere mosso al fornitore, svincolandolo da ogni responsabilità.

È facile intuire che l'attribuzione di predeterminati obblighi in capo a questi soggetti comporti la loro responsabilità (anche penale?) in caso di inosservanza. La catena degli individui coinvolti durante le fasi di vita del sistema intelligente appare particolarmente nutrita, pertanto individuare il responsabile di un eventuale danno cagionato dal prodotto potrebbe non essere semplice. Tali sintetiche battute potrebbero evocare il meccanismo della successione di posizioni di garanzia (nota in ambito penalistico) e la relativa difficoltà di individuare il momento in cui finisce la responsabilità di un soggetto ed inizia quella di un altro.

Il Titolo IV è dedicato agli obblighi di trasparenza che devono essere rispettati dai sistemi di IA. L'aspetto di maggior interesse dell'unico articolo contenuto in questo Titolo riguarda l'obbligo in capo ai fornitori di garantire che le persone fisiche che interagiscono con i sistemi intelligenti siano edotte del fatto che stanno interagendo con essi, a meno che ciò non sia di palese evidenza. La Commissione specifica che non si rende necessario rispettare tale obbligo nel caso in cui i sistemi coinvolti siano autorizzati dalla legge al fine di indagare, perseguire e prevenire condotte penalmente rilevanti o se si debba garantire l'esercizio del diritto convenzionalmente garantito alla libertà di espressione, alla libertà delle arti e delle scienze.

Desta qualche perplessità l'inciso concernente il perseguimento delle fattispecie di reato. Sembra quasi che la Proposta acconsenta a sacrificare la trasparenza proprio dove essa va maggiormente garantita, ossia nell'accertamento di una fattispecie delittuosa. Per quanto possa effettivamente ritenersi che un simile sacrificio si renda necessario per una causa primaria come quella di ridurre la criminalità, dall'altro sembrano non essere tenute in debita considerazione le garanzie dei soggetti di volta in volta indagati.

Il Titolo V si dedica, invece, alle misure a sostegno dell'innovazione e, segnatamente, agli spazi di sperimentazione normativa. Questi ultimi, come si è già avuto modo di evidenziare nelle pagine precedenti, costituiscono uno spazio volto a facilitare lo sviluppo e la sperimentazione dei sistemi di IA prima

---

non si sia ad esso conformato. Finché il sistema resta sotto la loro responsabilità devono garantire che il prodotto non venga pregiudicato durante le fasi di stoccaggio e trasporto. *Proposta di Regolamento,* cit., artt. 26-27, pp. 60 ss.

[250] Gli utenti di questi sistemi sono tenuti a servirsene in modo conforme alle istruzioni per l'uso che li accompagnano, garantendo l'uso di input adeguati alle finalità del sistema ed informando il fornitore o il distributore circa eventuali rischi del sistema o in ordine ad incidenti gravi o malfunzionamenti, interrompendo l'uso del prodotto. *Proposta di Regolamento,* cit., art. 29, p. 62.

[251] *Proposta di Regolamento,* cit., art. 28, p. 62.

dell'immissione sul mercato. La Proposta evidenzia, ai nostri fini, che coloro i quali partecipano ai suddetti spazi di sperimentazione restano responsabili per i danni ivi arrecati, chiarendo altresì che permane in tali spazi il controllo delle autorità competenti e che, nel caso di individuazione di rischi per la sicurezza e per i diritti fondamentali, questi devono essere prontamente fronteggiati con adeguate misure o, a mali estremi, con la sospensione della sperimentazione.

Sempre in ottica penalistica il presente Titolo chiarisce, altresì, che i dati personali legalmente raccolti, seppur per altre finalità, possono essere adoperati negli spazi di sperimentazione normativa a condizione che vengano utilizzati in sistemi di IA sviluppati per motivi di interesse pubblico, come prevenire o perseguire reati o eseguire sanzioni penali, garantire la sicurezza e la sanità pubblica o migliorare la qualità dell'ambiente.

I Titoli VI e VII si occupano, rispettivamente, di istituire il comitato europeo per l'intelligenza artificiale[252] e la banca dati europea per i sistemi di intelligenza artificiale indipendenti ad alto rischio.

Il Titolo VIII, dedicato al monitoraggio e alla condivisione delle informazioni, attribuisce ai fornitori il compito di attuare sistemi di controllo successivi all'immissione nel mercato dei sistemi intelligenti e di segnalare alle autorità di vigilanza del mercato medesimo ogni incidente o malfunzionamento che possa costituire una minaccia per i diritti fondamentali dell'Unione. Tale segnalazione deve essere preceduta dall'accertamento del nesso causale (secondo il criterio della ragionevole probabilità) intercorrente tra l'operato dell'IA e l'incidente verificatosi.

La conformità del sistema intelligente al presente Regolamento non comporta necessariamente che quest'ultimo sia privo di rischi per la sicurezza o la salute dei consociati. Per tale motivo, ove sussistano residui rischi in tal senso, l'autorità di vigilanza del mercato può chiedere che vengano adottate misure per eliminare tali rischi procedendo, in caso di necessità, anche al ritiro del prodotto[253].

Il Titolo IX incoraggia l'elaborazione di codici di condotta per i sistemi di IA considerati non ad alto rischio al fine di esortare all'applicazione volontaria di questi ultimi.

Il Titolo X si dedica, invece, a due aspetti specifici del presente Regolamento: riservatezza e sanzioni. La Commissione attribuisce alle autorità nazionali deputate all'applicazione del presente Regolamento il compito di rispettare la riservatezza delle informazioni ottenute durante la loro attività, così tutelando i relativi diritti di proprietà intellettuale nonché l'integrità dei procedimenti penali e amministrativi.

La Commissione sancisce inoltre che, entro i limiti stabiliti dal presente Regolamento, spetta agli Stati individuare le sanzioni da applicare in caso di violazione del Regolamento medesimo secondo i criteri (invero penalistici) dell'effettività, dissuasività e proporzionalità. L'art. 71 della Proposta in esame

[252] A tale comitato è attribuito il compito di collaborare con la Commissione al fine di garantire un'efficace cooperazione con le autorità nazionali, un'uniforme applicazione del Regolamento e contribuire a delineare gli orientamenti della Commissione nella materia dell'intelligenza artificiale. *Proposta di Regolamento,* cit., art. 56, p. 76.

[253] *Proposta di Regolamento,* cit., artt. 61-62-67, pp. 79 ss.

utilizza a più riprese la parola "reato"[254], quasi come se mirasse a delineare nuove fattispecie delittuose che, con l'entrata in vigore del presente Regolamento, dovrebbero trovare cittadinanza negli ordinamenti europei.

I Titoli XI e XII concludono la Proposta con specifiche disposizioni di chiusura.

Tale Proposta costituisce un importante passo per l'Unione Europea in quanto cerca di fornire una disciplina uniforme ad un settore in costante evoluzione al fine di evitare il proliferare a macchia di leopardo di normative disarmonizzate. La versione definitiva della Proposta *de qua* potrebbe certamente subire delle modifiche ma, nonostante alcuni nodi interpretativi che si è cercato di evidenziare nel corso della trattazione, l'impianto normativo appare complessivamente soddisfacente e idoneo a costituire una base solida per poter affrontare compiutamente le sfide giuridiche poste dall'intelligenza artificiale.

9.2. La Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale.

Il 13 luglio 2021 è stata pubblicata una proposta di Risoluzione del Parlamento europeo volta a regolamentare il rapporto tra intelligenza artificiale e diritto penale. In apertura il Parlamento si dice consapevole del fatto che l'uso di sistemi dotati di intelligenza artificiale porti con sé non solo concreti benefici ma anche potenziali rischi per i diritti fondamentali (i quali vanno tutelati durante l'intero ciclo di vita dell'IA), nonché del fatto che, anche nel contesto giuridico, esistono sistemi intelligenti in grado di raggiungere livelli di prestazione analoghi a quelli delle persone esperte nel settore[255].

Il Parlamento chiarisce immediatamente che tali sistemi debbano essere al servizio dell'uomo nonché degni di fiducia, dovendo sempre essere possibile per il soggetto umano intervenire per interrompere il funzionamento del meccanismo.

---

[254] «3. Le seguenti violazioni sono soggette a sanzioni amministrative pecuniarie fino a 30.000.000 di EUR o, se l'autore del *reato* è una società, fino al 6% del fatturato mondiale totale annuo dell'esercizio precedente, se superiore: a) inosservanza del divieto delle pratiche di intelligenza artificiale di cui all'articolo 5; b) non conformità del sistema di IA ai requisiti di cui all'articolo 10. 4. La non conformità del sistema di IA ai requisiti o agli obblighi previsti dal presente regolamento, diversi da quelli di cui agli articoli 5 e 10, è soggetta a sanzioni amministrative pecuniarie fino a 20.000.000 di EUR o, se l'autore del *reato* è una società, fino al 4% del fatturato mondiale totale annuo dell'esercizio precedente, se superiore. 5. La fornitura di informazioni inesatte, incomplete o fuorvianti agli organismi notificati e alle autorità nazionali competenti è soggetta a sanzioni amministrative pecuniarie fino a 10.000.000 di EUR o, se l'autore del *reato* è una società, fino al 2% del fatturato mondiale totale annuo dell'esercizio precedente, se superiore» *Proposta di Regolamento*, cit., art. 71, p. 87, corsivo nostro.

[255] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale*, 13.7.2021, Considerando A-B-D, p. 5. «Considerando che alcuni paesi, compresi diversi Stati membri, fanno un maggiore uso delle applicazioni di IA o dei sistemi integrati di IA per le attività di contrasto e nel settore giudiziario rispetto ad altri, in parte a causa della mancanza di regolamentazione e delle differenze normative che consentono o impediscono l'uso dell'IA per talune finalità; che l'uso sempre più frequente dell'IA nel diritto penale si basa, in particolare, sulla promessa che ridurrà determinati tipi di reati e favorirà l'adozione di decisioni più obiettive; che tale promessa non sempre viene mantenuta» *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale*, cit., Considerando C, p. 5.

Essi devono inoltre essere non discriminatori, sicuri, trasparenti, spiegabili e affidabili, nonché in grado di apportare considerevoli vantaggi per i consociati.

Con particolare riferimento al settore giudiziario il Parlamento chiarisce che i sistemi di IA debbano essere trasparenti, posti a tutela dei diritti fondamentali e rispettosi dei principi di necessità e proporzionalità. Essi devono altresì garantire un sistema giustizia equo e privo di discriminazioni[256].

L'uso di questi sistemi nelle attività di contrasto può contribuire alla sicurezza dei cittadini (in termini di prevenzione di specifici reati) ma può altresì comportare un rischio per questi ultimi (basti pensare ad un uso sproporzionato dell'IA che porterebbe a forme di sorveglianza di massa).

Proprio in ragione di questa doppia visuale prospettica (quella dei benefici da un lato e quella dei rischi dall'altro) il Parlamento evidenzia la necessità di stabilire un modello di responsabilità e prevenzione per quanto concerne i possibili effetti nocivi causati dall'IA nell'ambito del diritto penale, dovendosi sempre optare per una forma di responsabilità "umana".

Posto che le conseguenze delle "decisioni" dei sistemi intelligenti adoperati nel settore giudiziario sono in grado di incidere sulla vita e sui diritti dei consociati, il Parlamento chiarisce l'importanza di tenere sempre presente il contemperamento tra una produttiva attività di contrasto[257] e il rispetto dei diritti fondamentali.

Il Parlamento constata che, al momento, nell'ambito dell'Unione i sistemi di IA sono adoperati prevalentemente in ambito civile, mentre in altri Paesi essi sono utilizzati per aspetti più delicati del settore giudiziario, come stabilire la probabilità di recidiva o al fine di corroborare una decisione in ordine all'applicazione delle misure cautelari o delle pene.

Data l'opacità di questi sistemi, la possibilità che essi possano riprodurre errori o discriminazioni, nonché l'evenienza che questi possano essere oggetto di attacchi informatici, rischiano di essere messi sotto sforzo alcuni dei principi

---

[256] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando E-F-G.

[257] «Considerando che l'IA è utilizzata dalle autorità di contrasto in applicazioni quali le tecnologie di riconoscimento facciale, ad esempio per la ricerca in database di sospetti e l'identificazione delle vittime della tratta di esseri umani o di sfruttamento sessuale e abusi nei confronti di minori, riconoscimento automatizzato delle targhe, identificazione di chi parla, identificazione vocale, tecnologie di lettura labiale, analisi di segnali acustici (algoritmi di rilevamento di colpi di arma da fuoco), ricerca autonoma e analisi di database identificati, previsioni (polizia predittiva e analisi della scena del crimine), strumenti di rilevamento dei comportamenti, strumenti avanzati di autopsia virtuale per contribuire a determinare la causa di morte, strumenti autonomi per identificare le frodi finanziarie e il finanziamento del terrorismo, monitoraggio dei social media (estrazione e raccolta di dati per l'estrazione di connessioni) e sistemi di sorveglianza automatica che integrano diverse capacità di rilevamento (come il rilevamento cardiaco e le videocamere termiche); che le applicazioni di cui sopra, unitamente ad altre applicazioni potenziali o future della tecnologia dell'IA nelle attività di contrasto, possono presentare gradi molto diversi di affidabilità e precisione e di impatto sulla protezione dei diritti fondamentali e sulle dinamiche dei sistemi di giustizia penale; che molti di questi strumenti sono utilizzati in paesi non UE ma sarebbero illegali ai sensi dell'acquis dell'Unione in materia di protezione dei dati e della relativa giurisprudenza; che la diffusione di routine degli algoritmi, anche con un tasso limitato di falsi positivi, può generare falsi allarmi in numero decisamente superiore agli allarmi corretti» *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando M, p. 6.

cardine della giustizia penale, come la presunzione di innocenza, il diritto a un ricorso effettivo e il diritto ad un processo equo[258].

Vanno altresì messi in evidenza i possibili rischi connessi all'uso di questi sistemi per la prevenzione dei reati, dato il potenziale contrasto con un diritto penale del fatto, volto a sanzionare un reato dopo che è stato commesso e non diretto a intercettare, quasi in ottica positivistica, la condotta del soggetto considerato (a torto o a ragione) potenzialmente pericoloso[259].

Come abbiamo ormai avuto modo di imparare, l'IA richiede per il proprio funzionamento l'analisi di un'ingente mole di dati, la tutela dei quali dovrebbe essere garantita non solo dalla nutrita normativa dell'Unione in materia di protezione dei dati personali e della vita privata ma anche da un trattamento di siffatti dati svolto per fini legittimi, chiari e predeterminati. Occorre altresì predisporre meccanismi di prevenzione di attacchi dolosi a questi sistemi, garantendo il controllo umano e la sicurezza sin dalla progettazione al fine di prevenire accessi non autorizzati a dati personali e, più in generale, qualsiasi pericolo per l'integrità dei dati medesimi[260].

Gli strumenti di IA utilizzati dalle autorità giudiziarie contribuiscono a prevenire, accertare e perseguire fattispecie delittuose, nonché ad eseguire le relative pene, dunque tali sistemi devono essere considerati ad alto rischio[261] e devono essere sicuri, trasparenti, spiegabili e proporzionati[262].

Il Parlamento esprime preoccupazione in ordine alle possibili distorsioni e discriminazioni intrinseche nei sistemi di IA e determinate dai dati in essi inseriti, chiedendo pertanto l'introduzione di misure di salvaguardia volte ad evitare tali effetti collaterali, specie nel settore giudiziario[263].

---

[258] «Ribadisce che tutte le soluzioni di IA per le attività di contrasto e il settore giudiziario devono inoltre rispettare appieno i principi di dignità umana, non discriminazione, libertà di movimento, presunzione di innocenza e diritto di difesa, compreso il diritto di non rispondere, libertà di espressione e informazione, libertà di riunione e associazione, uguaglianza dinanzi alla legge, principio dell'eguaglianza delle armi e diritto a un ricorso effettivo e a un processo equo, conformemente alla Carta e alla Convenzione europea dei diritti dell'uomo; sottolinea che l'utilizzo dell'IA deve essere proibito se incompatibile con i diritti fondamentali» *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 2, p. 8.

[259] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando Q, p. 7.

[260] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 1-11, p. 8-10.

[261] Secondo quanto stabilito dalla Proposta di Regolamento esaminata nel paragrafo precedente.

[262] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 4, p. 8. «Riconosce il contributo positivo di determinati tipi di applicazioni di IA al lavoro delle autorità di contrasto e giudiziarie in tutta l'Unione; sottolinea, ad esempio, il miglioramento nella gestione della giurisprudenza ottenuto con gli strumenti che garantiscono ulteriori opzioni di ricerca; ritiene che vi sia una serie di altri utilizzi potenziali per l'IA da parte delle autorità di contrasto e giudiziarie che potrebbe essere esaminata, tenendo in considerazione i cinque principi della Carta etica sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambienti connessi, adottata dalla CEPEJ e prestando particolare attenzione agli "utilizzi da esaminare con le più estreme riserve" identificati dalla CEPEJ» *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 5, pp. 8-9.

[263] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale*, cit., Considerando 8-9, p. 9. «Sottolinea che gli insiemi di dati e i sistemi algoritmici utilizzati per condurre classificazioni, valutazioni e previsioni nelle diverse fasi del trattamento dei dati per lo sviluppo dell'IA e delle relative tecnologie potrebbero anche risultare in un trattamento

Il Parlamento mette altresì in evidenza l'asimmetria di potere esistente tra coloro che si servono dei sistemi intelligenti e coloro che, invece, li "subiscono". Tale asimmetria può influire sul diritto di difesa dei soggetti imputati o indagati i quali avrebbero considerevoli difficoltà a confutare in sede giudiziaria l'output prodotto dall'IA[264]. Proseguendo su questo solco il Parlamento chiarisce, inoltre, che i sistemi intelligenti impiegati in ambito giudiziario non possono né attribuire diritti né imporre obblighi agli individui.

Probabilmente la sfida più importante da affrontare in questo settore è quella che concerne la realizzazione di un quadro giuridico idoneo a disciplinare l'imputabilità e la responsabilità per i possibili danni prodotti da tali sistemi. In tal sede il Parlamento evidenzia non solo l'importanza di applicare il principio di precauzione nella materia *de qua* ma anche di far sempre ricadere ogni forma di responsabilità su una persona fisica o giuridica. Non secondario il riferimento alla trasparenza che deve essere garantita dalle aziende che producono sistemi di IA adoperati in sede giudiziaria[265].

Il Parlamento chiarisce inoltre l'importanza di realizzare un quadro giuridico chiaro e completo che individui condizioni, modalità e conseguenze dell'uso di questi sistemi adoperati nelle attività di contrasto, al fine di garantire la trasparenza della giustizia penale e i diritti delle parti coinvolte nel processo. In tale ottica è importante che gli individui siano sempre resi edotti del fatto di essere o meno sottoposti all'uso di un sistema intelligente da parte delle autorità giudiziarie e di contrasto. I componenti di tali autorità non devono riporre cieca fiducia negli esiti apparentemente oggettivi dei sistemi di IA in quanto essi possono fornire risultati errati o discriminatori. È infatti importante che i soggetti che si servono di tali strumenti abbiano le competenze per mettere in dubbio output non condivisibili o, addirittura, per respingerli[266].

---

differenziale e in una discriminazione sia diretta che indiretta di gruppi di persone, in particolare poiché i dati utilizzati per formare gli algoritmi di polizia predittiva rispecchiano le attuali priorità di sorveglianza e, di conseguenza, potrebbero finire con il riprodurre e amplificare le discriminazioni esistenti; sottolinea, pertanto, che le tecnologie di IA, in particolare se diffuse per l'uso nelle attività di contrasto e nel settore giudiziario, richiedono una ricerca e un contributo interdisciplinari, anche da ambiti quali gli studi scientifici e tecnologici, gli studi critici sulla razza, gli studi sulla disabilità e altre discipline in sintonia con il contesto sociale, comprese le modalità di costruzione delle differenze, il lavoro di classificazione e le sue conseguenze» *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 22, p. 13.

[264] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 10, p. 10. Non si tratta di fantascienza bensì di realtà ormai concreta, basti pensare al noto caso Loomis, su cui S. CARRER, *Se l'amicus curiae è l'algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin*, in *Giurisprudenza Penale Web*, 24.4.2019.

[265] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 12-13, p. 10. «Chiede la spiegabilità, la trasparenza, la tracciabilità e la verifica degli algoritmi quali elementi necessari della vigilanza al fine di garantire che lo sviluppo, la diffusione e l'utilizzo di sistemi di IA per il settore giudiziario e delle attività di contrasto rispettino i diritti fondamentali e godano della fiducia dei cittadini, nonché al fine di garantire che i risultati generati dagli algoritmi di IA possano essere resi intelligibili per gli utenti e coloro che sono soggetti a tali sistemi, e che vi sia trasparenza riguardo ai dati di base e alle modalità con cui il sistema è giunto a una certa conclusione» *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 17, p. 11.

[266] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 14-15, pp. 10-11.

Ogni decisione idonea a produrre effetti giuridici su un soggetto deve sempre essere presa da un individuo umano al quale poter imputare la relativa responsabilità[267]. In ragione di ciò le autorità che si servono di strumenti di IA in ambito giudiziario devono considerare il potenziale impatto sulle varie fasi del procedimento penale e garantire sempre, non soltanto il relativo intervento umano, ma anche che le decisioni giudiziarie definitive siano assunte sempre da giudici in carne e ossa.

Il Parlamento incoraggia le autorità coinvolte a individuare le aree di maggior interesse per l'applicazione dei sistemi intelligenti e a definirne funzioni e limiti, prevedendo altresì l'esecuzione di una preliminare valutazione d'impatto su questi sistemi (segnatamente con riferimento ai possibili rischi per i diritti fondamentali) da effettuare prima che essi vengano posti in commercio o prima di un loro controllo periodico[268].

Il Parlamento sembra diffidare degli strumenti di polizia predittiva, avvertendo che se da un lato le autorità di contrasto possono effettuare correlazioni tra i dati storici al fine di prevedere i comportamenti dei consociati, di contro evidenzia l'impossibilità di stabilire nessi causali o fare previsioni affidabili al punto da costituire la sola base giuridica che determina il relativo intervento penale. In tal sede si chiede altresì di vietare l'uso di sistemi di riconoscimento automatici in luoghi pubblici, nonché la limitazione dell'utilizzo dei sistemi di riconoscimento facciale solo per fini giustificati e nel rispetto dei principi di necessità e proporzionalità[269]. Il Parlamento mette in guardia anche dai possibili rischi per i diritti fondamentali derivanti dall'uso di dati biometrici per l'identificazione a distanza, stante anche la dubbia validità scientifica di simili tecnologie. In tal senso invita la Commissione a vietare il trattamento dei dati biometrici per finalità che potrebbero comportare una sorveglianza di massa in spazi pubblici.

Il Parlamento chiude richiamando la necessità di realizzare un quadro giuridico ed etico in materia di intelligenza artificiale che miri a garantire la trasparenza di quest'ultima nonché il rispetto dei diritti umani[270].

---

[267] «Ai sensi del diritto dell'UE, una persona ha il diritto di non essere sottoposta a una decisione che produca effetti giuridici o abbia effetti significativi nei suoi confronti fondata esclusivamente su un trattamento automatizzato dei dati» *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 16, p. 11.

[268] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 18-20, p. 12.

[269] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 24-25-26, p. 14. «Esprime profonda preoccupazione per l'utilizzo di database privati di riconoscimento facciale da parte delle autorità di contrasto e dei servizi di intelligence, come Clearview AI, una banca dati di oltre tre miliardi di immagini raccolte illegalmente dai social network e da altre fonti Internet, comprese immagini di cittadini dell'Unione; invita gli Stati membri a obbligare le autorità di contrasto a indicare se stanno utilizzando la tecnologia Clearview AI o tecnologie equivalenti di altri fornitori; rammenta il parere del comitato europeo per la protezione dei dati secondo cui l'utilizzo di un servizio quale Clearview AI da parte delle autorità di contrasto nell'Unione europea non sarebbe probabilmente coerente con il regime di protezione dei dati dell'UE; chiede un divieto sull'utilizzo di database privati di riconoscimento facciale per le attività di contrasto» *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 28, p. 15.

[270] *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Considerando 30-31-33-34, pp. 15-16.

Nelle motivazioni che hanno condotto il Parlamento a proporre tale Risoluzione si rinvengono assunti ormai acquisiti – come la necessità di creare fiducia in un'intelligenza artificiale antropocentrica e basata sui diritti fondamentali – e propositi "nuovi", come migliorare il lavoro delle autorità di giustizia penale e di contrasto, perseguire più efficacemente determinati reati, analizzare le scene del crimine e stabilire la probabilità di recidiva[271].

Il Parlamento chiarisce l'importanza di rispettare, durante l'intero ciclo di vita dell'IA, i principi di trasparenza e tracciabilità, considerandoli requisiti necessari per poter considerare legittimo e affidabile un sistema intelligente, specie ove esso vada utilizzato nel contesto giudiziario.

10. Conclusioni.

Dall'appena svolta rassegna possiamo trarre alcuni punti fermi, stante la ricorrenza di una serie di concetti idonei a connotare l'approccio europeo al tema dell'intelligenza artificiale.

L'Unione europea sceglie di intervenire in materia servendosi, perlopiù, di strumenti di *soft law* (quali orientamenti o linee guida, su temi che toccano tanto il diritto quanto l'etica), in modo da consentire una più rapida elaborazione e una più agile revisione degli stessi, lasciando aperto il dialogo sui possibili scenari che l'IA è in grado di aprire[272].

L'Unione mira a diventare leader nel settore, facendo del rispetto dei diritti fondamentali e di un approccio etico la sua cifra stilistica. Fa questo fornendo una definizione compiuta di cosa si debba intendere per "intelligenza artificiale", raccomandando il rispetto della relativa proprietà intellettuale e impegnandosi a creare un mercato unico nel settore.

L'obiettivo comune è quello di dar vita ad un quadro giuridico armonico, che garantisca la certezza del diritto ed eviti pericolose frammentazioni normative tra Stati membri. Questi ultimi vengono esortati a lavorare con approccio sinergico, mettendo in atto un proficuo scambio di buone pratiche nonché un reciproco riconoscimento transfrontaliero dei beni intelligenti.

L'Unione intende inoltre realizzare un approccio normativo in grado di garantire la trasparenza e la tracciabilità dei sistemi intelligenti, alimentando la fiducia dei consociati nei loro riguardi.

Abbiamo avuto modo di imparare dallo studio di questi testi che gli strumenti dotati di intelligenza artificiale necessitano, per il loro funzionamento, di un'ingente mole di dati (sicuri, accurati, integri, e completi) e che, per tale ragione, la protezione di questi ultimi si connota come punto nodale della materia *de qua*.

La qualità dei dati inseriti come input nel sistema intelligente appare di fondamentale importanza al fine di garantire un output accurato e, soprattutto, non

---

[271] «Nonostante i benefici che essa apporta, l'IA comporta nel contempo una serie di rischi potenziali, quali processi decisionali opachi, vari tipi di discriminazione, intrusione nella vita privata, rischi per la protezione dei dati personali, per la dignità umana e la libertà di espressione e informazione. Tali rischi potenziali sono ancora più gravi nel settore delle attività di contrasto e della giustizia penale, in quanto possono incidere sulla presunzione di innocenza, sui diritti fondamentali per la libertà e la sicurezza dell'individuo e su un ricorso effettivo e un processo equo» *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale*, cit., p. 19.

[272] A. TURANO, *Robotica e roboetica*, cit., p. 159.

discriminatorio. Non dimentichiamoci infatti che l'inserimento dei suddetti dati dipende pur sempre dalla supervisione umana, essendo pertanto possibile che la macchina riproduca i preconcetti e le distorsioni dell'uomo. Per tentare di scongiurare tale rischio potrebbe essere proficuo adottare un approccio multidisciplinare, che coinvolga dunque nello sviluppo dell'IA non soltanto esperti tecnici ma anche studiosi delle scienze sociali.

Il succitato controllo umano richiama l'importanza di assicurare un'intelligenza artificiale antropocentrica, volta cioè a coadiuvare l'uomo – mai a sostituirlo – residuando in capo a quest'ultimo la possibilità di contestare sempre l'output della macchina (anche per quanto concerne le forme di intelligenza artificiale maggiormente sofisticate le quali, essendo in grado di apprendere dall'esperienza, possono sfuggire al controllo umano e non essere del tutto prevedibili).

I soggetti che si relazionano con questi sistemi intelligenti hanno il diritto di ottenere informazioni complete in ordine al loro funzionamento, dovendo essere edotti, ad esempio, del fatto di star interagendo con una forma di intelligenza artificiale e non umana, ma anche del loro diritto a non essere sottoposti ad una decisione del tutto automatizzata.

L'Unione Europea, evidenziando le opportunità e i rischi connessi all'uso dei sistemi intelligenti, ci ha mostrato che questi ultimi possono trovare applicazione nel settore delle auto a guida autonoma, nel settore sanitario, in quello industriale, in quello dei droni a pilotaggio remoto e, soprattutto, nel settore della giustizia e delle autorità di contrasto.

Possiamo dire di vivere in un'epoca in cui stiamo assistendo ai primi passi dell'Unione Europea nel fornire una base giuridica a un fenomeno nuovo e proteiforme come l'intelligenza artificiale. In tale ottica è giusto cominciare a interrogarsi su quale possa essere, in questo ambito, lo spazio di operatività del diritto penale.

I riferimenti, più o meno diretti, non mancano. Basti a tal proposito ricordare il più volte citato principio di precauzione, il quale dovrà riuscire a trovare un proprio spazio di operatività senza ostacolare la ricerca scientifica.

L'Unione individua non pochi settori di rilevanza penale in cui l'intelligenza artificiale può venire in rilievo. Ci riferiamo, ad esempio, agli strumenti di polizia predittiva, ai *risk assessment tools* e ai sistemi di IA per l'identificazione biometrica remota.

Nel corso di questi interventi viene presa in considerazione la possibilità di un uso doloso dei sistemi intelligenti e l'importanza di fare il possibile per prevenirli al fine di garantire la sicurezza dei suddetti sistemi.

L'argomento che viene più spesso affrontato dai documenti su esaminati e che rappresenta l'aspetto di maggior interesse per il penalista è, indubbiamente, quello concernente l'attribuzione della responsabilità.

Come si comporteranno (*rectius*, come si potrebbero comportare) i classici modelli di imputazione della responsabilità quando a realizzare una fattispecie delittuosa sarà stata la macchina? Come si dovrà graduare la responsabilità dell'uomo che controlla il sistema quando quest'ultimo sia dotato di una certa percentuale di autonomia che, pertanto, sfugge al controllo umano? Si può dire che coloro che controllano la macchina rivestano una posizione di garanzia ex art. 40 cpv. c.p.? Sarà possibile (e se sì, in che termini) fare ricorso ai classici schemi

della responsabilità dolosa o colposa? Bisognerà rifuggire da semplicistiche forme di responsabilità oggettiva o si renderà, piuttosto, necessario prenderle in seria considerazione?

Questi sono solo alcuni degli interrogativi che, agli occhi del penalista, sorgono spontanei dalla disamina degli interventi normativi europei e ai quali si tenterà, senza alcuna pretesta di esaustività, di fornire una risposta nelle pagine seguenti.

CAPITOLO II
LE ANTINOMIE TRA INTELLIGENZA ARTIFICIALE E CATEGORIE PENALISTICHE

SOMMARIO: 1. Perché occuparci del fenomeno? Questioni di politica del diritto – 2. Qualche considerazione di politica criminale – SEZIONE PRIMA. L'INTELLIGENZA ARTIFICIALE – 3. Non conosciamo completamente una scienza se non conosciamo la sua storia – 4. Di cosa stiamo parlando? Per una definizione di "intelligenza artificiale" – 5. Tipologie e caratteristiche dell'IA: il *machine learning* – 5.1 L'imprevedibilità dell'IA: il *black box effect* – 6. Il ruolo dei dati – 7. Le differenze con la robotica – 8. Per un'IA etica e conforme al diritto – 9. Considerazioni conclusive – SEZIONE SECONDA. QUESTIONI DI RESPONSABILITÀ PENALE – 10. Quale ruolo per il diritto penale? – 11. Alcuni spunti penalistici dal diritto civile – 12. Evitare due eccessi: dalla deresponsabilizzazione alla responsabilità oggettiva – 13. Il principio di precauzione – 14. La responsabilità per danno da prodotto – SEZIONE TERZA. INTELLIGENZA ARTIFICIALE: STRUMENTO, AUTORE E VITTIMA DEL REATO – 15. Premessa – 16. Intelligenza artificiale come "strumento" di reato – 16.1. L'imputazione della responsabilità per reato commesso "mediante" l'IA – 16.2. Un problema di "autoria mediata"? – 16.3. Dall'automazione all'autonomia – 17. Intelligenza artificiale come "autore" di reato – 17.1 Nuove soluzioni per nuovi problemi? – 17.2 Problemi di responsabilità: la (presunta) crisi del modello vicario – 17.2.1. Segue: il reato diverso da quello voluto e l'interruzione del nesso causale – 17.3 Personalità giuridica elettronica: considerazioni di carattere generale – 17.3.1 L'assimilazione alle persone giuridiche: una falsa pista – 17.4 La diretta responsabilizzazione penale dell'IA (?) – 17.4.1 Una via difficilmente percorribile – 17.5 Il pensiero di Gabriel Hallevy – 17.5.1 Le obiezioni alla teoria di Hallevy – 17.6 Come declinare l'elemento soggettivo: a) il dolo – 17.6.1 (Segue) b) la colpa – 17.7 Due brevi intermezzi: a) la prospettabilità di una posizione di garanzia – 17.8 (Segue) b) il rischio consentito – 17.9 *De iure condito* vs. *De iure condendo* – 18. Conclusioni (dalle quali siamo ancora lontani) – 19. Intelligenza artificiale come "vittima" di reato.

1. Perché occuparci del fenomeno? Questioni di politica del diritto.

Ci troviamo alle pendici di una nuova rivoluzione industriale[1] e sociale[2]. Parte della dottrina non ha esitato a definirla come la "quarta rivoluzione industriale"[3], altri ancora hanno preferito parlare di un "mutamento di paradigma"[4]. Ciò che è indubbio è che si tratti di una vera e propria rivoluzione,

[1] C. PIERGALLINI, *Intelligenza artificiale: da "mezzo" ad "autore" del reato?*, in *Riv. it. dir. proc. pen*., 2020, p. 1748; A. AMIDEI, *Robotica intelligente e responsabilità: profili e prospettive evolutive del quadro normativo europeo*, in U. RUFFOLO, *Intelligenza artificiale e responsabilità*, Milano, 2017, p. 77; G. CAPILLI, *I criteri di interpretazione della responsabilità*, in G. ALPA, *Diritto e intelligenza artificiale*, cit., p. 459.

[2] A. TURANO, *Robotica e roboetica*, cit., p. 149.

[3] K. SCHWAB, *La quarta rivoluzione industriale*, Milano, 2016; L. FLORIDI, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Milano, 2017. Un richiamo alla quarta rivoluzione industriale è rinvenibile anche in M. DI FLORIO, *Il diritto penale che verrà. Brevi considerazioni sul possibile impiego dell'IA per prevenire il rischio di disastri colposi*, in *Archivio Penale*, 2021, p. 8; M. GABBRIELLI, *Dalla logica al deep learning: una breve riflessione sull'intelligenza artificiale*, in U. RUFFOLO, *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, Torino, 2020, p. 28; C. TREVISI, *La regolamentazione in materia di intelligenza artificiale, robot, automazione: a che punto siamo*, in *MediaLaws*, 25.6.2018, p. 5.

[4] V. MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in U. RUFFOLO, *Intelligenza artificiale. Il diritto, i diritti, l'etica*, cit., p. 547; A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 7; A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal*, 2019, p. 78.

comunque la si voglia chiamare[5], e che spetti alle istituzioni e ai giuristi (trovandoci ancora agli albori di questa trasformazione tecnologica) veicolarla e svilupparla nella corretta direzione: a favore dell'uomo e non contro di esso[6].

Vogliamo rimarcare in tal sede questo concetto di "rivoluzione" ed evidenziare che non si tratta affatto di un termine "troppo forte": l'intelligenza artificiale è destinata a investire una moltitudine di aspetti della vita e degli interessi della collettività, a rendere possibili nuovi comportamenti illeciti, ma anche a far sorgere nuovi interessi meritevoli di protezione giuridica, anche penalistica[7].

Si tratta di tecnologie dirompenti[8], in grado di dar vita ad una "frattura antropologica"[9], a una nuova idea di umanità[10]. Comprendere come dovrà destreggiarsi il diritto, anche penale, tra le pieghe di questa proteiforme tecnologia non sarà semplice. Si avrà spesso la sensazione di navigare senza bussola[11], ma non per questo i giuristi potranno esimersi dal delicato compito di (quantomeno provare a) regolamentare la materia dell'intelligenza artificiale, proprio in ragione dell'impatto che essa è destinata ad avere sulle nostre vite[12].

Occorre preliminarmente constatare che individuare un unico approccio corretto per affrontare una materia in costante divenire come questa è compito tutt'altro che semplice. Nella consapevolezza che probabilmente un *modus operandi* universalmente corretto non può esistere, tenteremo in tal sede di essere quanto più schematici possibile.

Potrebbe essere utile, dunque, prendere la mosse da una duplice valutazione prospettica, in termini di politica del diritto e di politica criminale.

Preliminarmente occorre chiederci perché il diritto dovrebbe occuparsi dell'intelligenza artificiale. A ben vedere quello dell'IA può essere a tutti gli effetti definibile come un "fenomeno sociale" e, come ogni fenomeno sociale di

---

[5] Parlano di "rivoluzione digitale" S. CHIARLONI, *Riflessioni minime su Intelligenza Artificiale e servizi giuridici*, in *Giur. it., Speciale 170 anni*, 2019, p. 9 e A. GULLO, *Nuove frontiere tecnologiche e sistema penale: alcune note introduttive*, in *Dir. pen. cont. - Riv. Trim.* 2/2019, p. VI.

[6] R. CINGOLANI, D. ANDRESCIANI, *Robots, macchine intelligenti e sistemi autonomi: analisi della situazione e delle prospettive*, in G. ALPA, *Diritto e intelligenza artificiale,* cit., p. 45.

[7] L. PICOTTI, *Diritto penale e tecnologie informatiche: una visione d'insieme*, in A. CADOPPI, S. CANESTRARI, A. MANNA, M. PAPA, *Trattato di diritto penale.* C*ybercrime*, Torino, 2019, p. 37.

[8] J. MANYIKA, M. CHUI, J. BUGHIN, R. DOBBS, P. BISSON, A. MARRS, *Disruptive technologies: Advances that will transform life, business, and the global economy*, in *McKinsey Global Institute Report*, 1.5.2013; M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 28.

[9] V. MANES, *L'oracolo algoritmico,* cit., p. 547.

[10] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 7.

[11] Questa suggestiva immagine è suggerita da P. SEVERINO, *Intelligenza artificiale,* cit., p. 531.

[12] Sembra che ormai l'IA regoli la nostra vita quotidiana, anche a nostra insaputa, così G. ITALIANO, *Intelligenza Artificiale: passato, presente, futuro*, in F. PIZZETTI, *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino, 2018, p. 217; G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 1: «l'intelligenza artificiale interviene (o può e potrà intervenire) sempre più in quasi ogni nostra attività, anche senza che noi ne siamo coscienti (si pensi a quando un frigorifero "intelligente" muta automaticamente di temperatura per mantenerla invariata, qualora la presenza di persone modifichi quella del locale in cui l'elettrodomestico si trova) similmente a quanto accade per il diritto, che – tra l'altro conformemente al brocardo ubi ius ibi societas, ubi societas ibi ius – pervade la nostra vita pure quando riteniamo di escluderlo dall'ambito di essa».

una certa portata, esso *deve* trovare una regolamentazione giuridica[13]. Inoltre occorre osservare che alcune forme di intelligenza artificiale sono in grado di porre in essere condotte che, ove fossero realizzate da un uomo, sarebbero regolate da disposizioni legislative[14]. Sembra pertanto che questo primo interrogativo possa trovare una risposta affermativa.

Tali assunti si palesano tanto più veritieri se si considera che il binomio tra innovazione tecnologica e diritto[15], anche penale[16], non costituisce certo una novità. Il diritto viene "sfidato"[17], da queste nuove tecnologie e, d'altronde, questo non dovrebbe sorprenderci più di tanto. Viviamo nella società del rischio[18], ossia in una società connotata da un numero potenzialmente indeterminato di rischi per tutti i consociati. Si tratta, a ben vedere, di una prospettiva che si candida ad essere nuovamente chiamata in causa dall'incedere di queste nuove tecnologie. La materia dell'intelligenza artificiale, infatti, porta con sé una moltitudine di benefici ma anche considerevoli rischi[19], il che condurrà tanto i legislatori quanto i giuristi a bilanciare i possibili vantaggi da essa derivabili con il rispetto dei diritti dei consociati[20].

Diritto e tecnologia trovano nell'intelligenza artificiale una forma di "*reductio ad unum*": non si tratta soltanto di adattare i principi dell'ordinamento al progresso e all'evoluzione scientifica, bensì di fare un uso mirato dell'innovazione, finalizzato a raggiungere una maggiore certezza del diritto, intesa non più come ricerca di sintesi dell'avanzamento tecnologico bensì come *fine* ultimo dell'intelligenza artificiale[21].

Nonostante la «magnetica e inquietante personalità»[22] dell'IA, sarà opportuno rifuggire da atteggiamenti catastrofici e prediligere un approccio propositivo alla

---

[13] L'intelligenza artificiale «non poteva non chiamare in causa il diritto, la sua ricerca, spesso faticosa, di dare un senso alla realtà, allo scorrere delle vicende umane» A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 4.

[14] H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law*, in *Artificial Intelligence and Law*, 25, 2017, p. 342. Basti pensare, a titolo meramente esemplificativo, alle auto a guida autonoma: appare fuor di dubbio che esse pongano in essere condotte che possono in egual misura essere realizzate dall'uomo e che sono regolate da una normativa specifica quale il Codice della Strada.

[15] M.G. LOSANO, *Giuscibernetica: macchine e modelli cibernetici nel diritto*, Torino, 1969; V. FROSINI, *Cibernetica, diritto e società*, Milano, 1973; S. RODOTÀ, *Elaboratori elettronici e controllo sociale*, Bologna, 1973.

[16] G. MARINUCCI, *Innovazioni tecnologiche e scoperte scientifiche: costi e tempi di adeguamento delle regole di diligenza*, in *Riv. it. dir. proc. pen.*, 2005, pp. 29 ss. In chiave attualizzante G. MOBILIO, *L'intelligenza artificiale e i rischi di una* "disruption" *della regolamentazione giuridica*, in *BioLaw Journal*, 2020, p. 411 richiama di concetti di "tecnicizzazione del diritto" e "globalizzazione giuridica".

[17] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 9.

[18] L. STORTONI, *Angoscia tecnologica ed esorcismo penale*, in *Riv. it. dir. proc. pen.*, 2004, p. 71. Sul tema più in generale v. V. MILITELLO, *Rischio e responsabilità penale*, Milano, 1988.

[19] I. SALVADORI, *Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale*, in *Riv. it. dir. proc. pen.*, 2021, p. 116; G. MOBILIO, *L'intelligenza artificiale*, cit., p. 424; A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 15.

[20] P. SEVERINO, *Intelligenza artificiale*, cit., p. 531.

[21] G. ROMANO, *Diritto, robotica e teoria dei giochi: riflessioni su una sinergia*, in G. ALPA, *Diritto e intelligenza artificiale*, cit., p. 105.

[22] M. PAPA, *Future crimes: intelligenza artificiale e rinnovamento del diritto penale*, in *Criminalia*, 2019, ora anche in *disCrimen* dal 4.3.2020, p. 1.

materia[23] in grado di rispondere compiutamente alle "sfide adattive"[24] che l'intelligenza artificiale porterà con sé. In questo senso si renderà necessario dar vita a un'intelligenza artificiale umano-centrica e affidabile[25] nonché conforme ai valori della convivenza civile[26]. Parte della dottrina, evidenziando l'ormai attuale interazione tra uomo e macchine, ha parlato di *homo sapiens 2.0*[27], di un uomo, cioè, che si serve dei sistemi intelligenti per migliorare le sue performance e ottimizzare i suoi processi produttivi delegando alle macchine le mansioni più usuranti, rimarcando un ulteriore fondamentale principio tipico di un'IA antropocentrica: sono le macchine ad essere al servizio dell'uomo, non il contrario[28].

Avremo modo di vedere come il fenomeno dell'intelligenza artificiale sia in grado di combinare questioni etiche e giuridiche[29]. Volendoci per il momento concentrare su tale ultimo aspetto, resta da chiederci come si muoverà l'IA in un ambiente antropocentrico come il nostro, concepito a misura d'uomo[30].

2. Qualche considerazione di politica criminale.

Una volta risposto in modo positivo alla domanda concernente la politica del diritto, possiamo porre la seconda questione e chiederci perché il diritto penale dovrebbe occuparsi dell'intelligenza artificiale. Essa è stata efficacemente definita come una *dual-use technology*[31], ossia una tecnologia capace di essere utilizzata tanto per fini leciti quanto per fini illeciti[32]. Il potenziale uso ambivalente di questi

---

[23] L. STORTONI, *Angoscia tecnologica*, cit., pp. 71-72, l'A. evidenzia chiaramente come questi due approcci, quello catastrofista e quello rassicurante, siano invero due modi con cui l'uomo, da sempre, guarda al futuro.

[24] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 4.

[25] Come d'altro canto evidenziano anche alcuni dei testi normativi europei presi in esame nelle pagine precedenti: *Piano coordinato sull'intelligenza artificiale,* cit., Cap. I Sez. I, Par. 5; *Orientamenti etici per un'IA affidabile,* cit., Cap. I Sez. I, Par. 7; *Libro Bianco sull'intelligenza artificiale,* cit., Cap. I Sez. I, Par. 8; *Proposta di Regolamento,* cit., Cap. I Sez. II, Par. 9.1; *Proposta di Risoluzione del Parlamento europeo sull'intelligenza artificiale nel diritto penale,* cit., Cap. I Sez. II, Par. 9.2.

[26] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 6.

[27] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 44.

[28] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 52.

[29] P. SEVERINO, *Intelligenza artificiale*, cit., p. 531.

[30] M.B. MAGRO, *Decisione umana e decisione robotica. Un'ipotesi di responsabilità da procreazione robotica*, in *La Legislazione Penale*, 10.5.2020, p. 1. Parte della dottrina parla a tal proposito della «questione problematica della *conditio humana* (se e come gli esseri umani possono trovare la loro strada in un mondo di macchine) sia quella della *conditio automata* (se e come le macchine intelligenti possono trovare la loro strada in una società umana)» C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale? Sulla trasformazione algoritmica della società*, in *Riv. it. dir. proc. pen.*, 2020, p. 1916.

[31] I. SALVADORI, *Agenti artificiali*, cit., p. 86.

[32] Sul duplice uso dei robot, ossia sulla possibilità che essi vengano adoperati a vantaggio degli individui oppure in modo distorto, v. A. TURANO, *Robotica e roboetica*, cit., p. 134. In dottrina è stato proposto un esempio calzante di uso dell'IA per scopi illeciti: un'«area criminale in cui viene ampiamente sfruttata l'intelligenza artificiale è il traffico di droga detto *business to business*, che si avvale di droni e sottomarini controllati a distanza. I sottomarini senza equipaggio offrono un chiaro esempio del potenziale duplice utilizzo, positivo e negativo, dell'Intelligenza Artificiale: sono stati ideati per scopi legittimi (difesa, protezione delle frontiere, pattugliamento delle acque), rivelandosi tuttavia funzionali anche ad attività illegali» R. BORSARI, *Intelligenza Artificiale e responsabilità penale: prime considerazioni*, in *MediaLaws*, 20.11.2019, p. 263.

sistemi o, per meglio dire, più in generale, di ogni forma di progresso scientifico, porta con sé una forse ineliminabile porzione di "*ambiguità del processo tecnologico*"[33], cioè la possibilità di servirsi di questi sistemi contro o a favore dell'uomo. Inoltre al diritto penale spetta occuparsi dell'intelligenza artificiale in quanto essa è in grado di integrare condotte penalmente rilevanti (sia essa considerata come autore o mezzo di un reato) e di cagionare danni a terzi. Si tratterebbe insomma di comportamenti di cui un soggetto umano sarebbero indubbiamente chiamato a rispondere sul piano giuridico[34].

Il mondo dell'intelligenza artificiale ci porterà a guardare con un nuovo sguardo all'intero sistema penale ed alle possibili ricadute sulla teoria generale del diritto[35]. Tocca dunque capire quale possa essere il ruolo del diritto penale di fronte ai pericoli, più o meno incogniti, di questa tecnologia[36].

In tale ottica, si renderà necessario dotarsi di un "diritto penale effettivo", condizione imprescindibile non solo per tutelare i consociati che, in un modo o nell'altro, si troveranno a interagire con le diverse forme di IA, ma anche per fornire certezza agli operatori del settore. Questi ultimi necessiteranno di regole, standard e principi chiari per operare in sicurezza ma anche, e soprattutto, di criteri d'imputazione certi per gestire il proprio lavoro senza inammissibili presunzioni di colpevolezza[37].

Come su accennato, l'arrivo dell'intelligenza artificiale nelle nostre vite porterà con sé l'insorgere di nuove forme di aggressione[38] (anche di nuovi beni giuridici)[39] e la rielaborazione dei diritti preesistenti[40]. Per fronteggiare tali innovative condotte lesive di beni giuridicamente tutelati dall'ordinamento si potrebbe procedere secondo tre percorsi distinti:

- servirci degli strumenti classici del nostro codice penale e applicare le fattispecie di reato in esso già contenute, chiaramente entro i limiti dell'interpretazione estensiva di queste ultime;
- rielaborare fattispecie criminose già esistenti, in modo da renderle conformi alle peculiarità dell'IA;

[33] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 24.
[34] P. MORO, *Libertà del robot? Sull'etica delle macchine intelligenti*, in R. BRIGHI, S. ZULLO, *Filosofia del diritto e nuove tecnologie. Prospettive di ricerca tra teoria e pratica*, Roma, 2015, p. 526.
[35] Analizza tali fenomeni, alla luce dei mutamenti generati dall'evoluzione tecnologica, L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., p. 36.
[36] L. STORTONI, *Angoscia tecnologica*, cit., p. 86, il quale invero si riferisce ai pericoli della tecnologia, in generale.
[37] L. STORTONI, *Angoscia tecnologica*, cit., p. 87-88.
[38] I. SALVADORI, *Agenti artificiali*, cit., p. 87; L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., p. 37.
[39] Basti pensare al «diritto delle persone umane di sapere se e quando stanno interagendo con una macchina o con un altro essere umano, e di decidere se, come e quando attribuire determinati compiti ad un sistema artificiale autonomo o ad una persona» A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 7.
[40] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 405.

- introdurre nuove ipotesi di reato, non solo al fine di garantire una tutela *ad hoc* ai beni giuridici coinvolti, ma anche con l'obiettivo di muoversi nel pieno rispetto del principio di legalità[41].

Gli sviluppi della scienza e della tecnica sono destinati a creare lacune di tutela (*rectius*, di responsabilità penale), che spetterà al legislatore colmare[42]. In capo a quest'ultimo ricadrà inoltre il difficile compito di individuare e delineare il fatto tipico costituente reato: queste nuove forme di aggressione sono destinate a spostarsi sempre di più dalla sfera empirico-materiale a quella immateriale, in ragione del trasmigrare dei nostri centri di interesse in rete: «alle difficoltà di dar forma ad un mondo sempre più smaterializzato, si aggiungono quelle dovute al moltiplicarsi degli interessi meritevoli di tutela ma tra loro in conflitto; interessi che occorre dunque bilanciare in concreto»[43].

I problemi non si arrestano di certo in punto di tecniche di incriminazione e di tutela di nuovi o rielaborati interessi meritevoli di tutela. Le maggiori difficoltà si riscontrano in ordine all'imputazione della responsabilità: chi sarà ritenuto responsabile di un evento lesivo materialmente realizzato dal sistema intelligente?

Il diritto penale è disegnato *per* l'uomo[44], incentrato sulla persona umana[45], è «edificato sull'uomo, sul rimprovero personale e colpevole, sul grado di responsabilità e di rimproverabilità per una azione *umana*»[46]. In realtà non potrebbe che essere così: il diritto penale è un prodotto dell'uomo e, conseguentemente, non può che essere costruito a misura d'uomo[47].

In realtà è il caso di osservare come il concetto stesso di "persona" si sia col tempo esteso fino ad includere nel suo ambito applicativo le "persone giuridiche", ossia entità non umane. Il diritto si è mostrato elastico ed in grado di adattarsi all'emersione di nuovi centri di imputazione giuridica, estendendo la responsabilità penale anche a entità diverse dalla persona in carne ed ossa[48]. Il diritto penale si è degnamente confrontato con la c.d. *corporate liability* che, come anticipavamo, ha consentito al diritto penale di andare a sanzionare anche entità non umane, ossia le persone giuridiche, grazie al disposto del d.lgs. n. 231 del 2001[49].

---

[41] Valuta di introdurre nuove fattispecie di reato o di intervenire a rimodellare quelle già esistenti anche F. BASILE, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Diritto Penale e Uomo*, 29.9.2019, p. 27.

[42] C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1928.

[43] M. PAPA, *Future crimes*, cit., p. 10.

[44] A. CAPPELLINI, Machina delinquere non potest? *Brevi appunti su intelligenza artificiale e responsabilità penale*, in *Criminalia*, 2018, ora anche in *disCrimen* dal 27.3.2019, p. 2.

[45] S. RIONDATO, *Robot: talune implicazioni di diritto penale*, in P. MORO, C. SARRA, *Tecnodiritto. Temi e problemi di informatica e robotica giuridica*, Milano, 2017, p. 90.

[46] V. MANES, *L'oracolo algoritmico*, cit., p. 547. «Il sistema penale è un sistema personocentrico e personologico, pensato per l'uomo ed affidato al giudizio dell'uomo, come tale *fallibile* ma pur sempre *controllabile* secondo un determinato iter argomentativo e i criteri che lo guidano» p. 548.

[47] S. RIONDATO, *Robotica e diritto penale (robot, ibridi, chimere, "animali tecnologici")*, in D. PROVOLO, S. RIONDATO, F. YENISEY, *Genetics, robotics, law, punishment*, Padova, 2014, p. 602.

[48] S. RIONDATO, *Robot: talune implicazioni di diritto penale*, cit., p. 90. L'A. evidenzia nel prosieguo che le società possono essere considerate anche come vittime del reato e, pertanto, destinatarie di un risarcimento danni. Lo stesso A. evidenzia anche in altro scritto che, «la qualità di vittima di reato non presuppone gli stessi requisiti psicologici della responsabilità penale» ID., *Robotica e diritto penale*, cit., p. 603.

[49] V. MANES, *L'oracolo algoritmico*, cit., p. 548. L'A. sottolinea che il diritto penale è riuscito a destreggiarsi tra le problematicità della responsabilità degli enti «e ciò, nonostante fosse difficile

Muovendoci sulla falsariga della responsabilità degli enti, potremmo arrivare a riconoscere, in capo ai sistemi dotati di intelligenza artificiale, una vera e propria personalità giuridica elettronica[50]? O sarebbe più utile continuare a considerare i sistemi intelligenti come meri oggetti e, pertanto, chiamare in causa le problematiche della responsabilità penale da prodotto[51]?

In dottrina si comincia già a parlare di una nuova forma di responsabilità: la c.d. responsabilità da algoritmo. Quest'ultimo, essendo integrato nel prodotto, può essere considerato come una «componente immateriale caratterizzante (…) idonea a qualificare l'ideatore come autore-fornitore di quella componente»[52], esponendo costui alla responsabilità per i danni derivanti a terzi connessi al malfunzionamento del prodotto intelligente.

Si è cercato, con queste sintetiche battute, di delineare la complessità dei problemi che il rapporto tra intelligenza artificiale e diritto penale porta con sé e che ci si propone di approfondire nel prosieguo. Fatta questa breve premessa sui motivi per i quali il diritto (*rectius*, il diritto penale) dovrebbe occuparsi del tema, ci sia consentito di fare un passo indietro e cercare di comprendere esattamente di cosa stiamo parlando.

SEZIONE PRIMA
L'INTELLIGENZA ARTIFICIALE

3. Non conosciamo completamente una scienza se non conosciamo la sua storia[53].

Il seme dell'intelligenza artificiale viene piantato nel 1950 da Alan Turing, il quale per primo si pose l'interrogativo "Can machines think?"[54]. Per tentare di affrontare al meglio la questione Turing propose di rimodulare la domanda sotto forma di gioco: l'*imitation game*. «Questo viene giocato da tre persone, un uomo

---

identificare, al cospetto di un *corporate crime*, una "condotta" secondo il concetto tradizionale di *azione*, nonostante fosse arduo rintracciare un rimprovero secondo una nozione personalistica di colpevolezza, e nonostante mancasse una "persona" da risocializzare mediante la pena». Sul punto v. anche C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 134.

[50] Tra gli altri, L.B. SOLUM, *Legal Personhood for Artificial Intelligences*, cit., pp. 1231-1287.

[51] Diffusamente sul punto C. PIERGALLINI, *Danno da prodotto e responsabilità penale. Profili dommatici e politico-criminali*, Milano, 2004. Sul tema v. anche U. CARNEVALI, *La responsabilità del produttore*, Milano, 1979; A. BERNARDI, *La responsabilità da prodotto nel sistema italiano: profili sanzionatori*, in *Riv. trim. dir. pen. econ*., 2003, pp. 1 ss.; D. CASTRONUOVO, *Responsabilità da prodotto e struttura del fatto colposo*, in *Riv. it. dir. proc. pen*., 2005, pp. 301 ss.; ID., *La responsabilità colposa nell'esercizio di attività produttive. Profili generali in tema di omicidio o lesioni per violazione delle discipline sulla sicurezza del lavoro e dei prodotti*, in A. CADOPPI, S. CANESTRARI, M. PAPA, *I delitti contro la persona*, I, Torino, 2006, pp. 579 ss.; L. FOFFANI, *Responsabilità per il prodotto e diritto comunitario: verso un nuovo diritto penale del rischio? Note comparatistiche sugli ordinamenti italiano e spagnolo*, in M. DONINI, D. CASTRONUOVO (a cura di), *La riforma dei reati contro la salute pubblica. Sicurezza del lavoro, sicurezza alimentare, sicurezza dei prodotti*, Padova, 2007, pp. 145 ss. Nella dottrina civilistica E. AL MUREDEN, *La sicurezza dei prodotti e la responsabilità del produttore: casi e materiali*, Torino, 2017.

[52] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning; dalla* machinery *produttiva all'auto* driverless*: verso una "responsabilità da algoritmo"?*, in U. RUFFOLO, *Intelligenza artificiale e responsabilità*, cit., p. 22. Lo stesso A. parla anche di responsabilità "da ideazione", p. 17.

[53] Il richiamo è ad A. COMTE, *Cours de philosophie positive*, 1 et 2 leçons, 1830-1842, p. 74.

[54] A.M. TURING, *Computing machinery and intelligence*, in *Mind*, 1950, pp. 433-460.

(A), una donna (B) e l'interrogante (C), che può essere dell'uno o dell'altro sesso. L'interrogante viene chiuso in una stanza, separato dagli altri due. Scopo del gioco per l'interrogante è quello di determinare quale delle altre due persone sia l'uomo e quale la donna (…) Le risposte, in modo che il tono di voce non possa aiutare l'interrogante, dovrebbero essere scritte, o, meglio ancora, battute a macchina. (…) "Che cosa accadrà se una macchina prenderà il posto di A nel gioco?"»[55]. La macchina sarà in grado di superare il test di Turing (e, dunque, di "pensare") se l'interrogatore, sapendo che uno dei suoi interlocutori non è umano, non sarà in grado di distinguere quale dei due sia la macchina o l'uomo[56].

Parte della dottrina ha mosso qualche riserva nei confronti di questo test, in particolare sulla sua *controfattualità*: il test di Turing «verifica la capacità di una macchina di svolgere un compito in modo tale che il *risultato* sarebbe indistinguibile dal risultato di un agente umano che lavora per raggiungere lo stesso compito»[57]. In altre parole, si cerca di mettere in evidenza che, nonostante la macchina si comporti *come se* fosse intelligente, non significa che essa lo sia davvero, essendo quest'ultima priva di coscienza[58]. Il Test di Turing, in altri termini, si fonda su una interpretazione meramente esteriore della coscienza e della volontà, accogliendo una nozione di "antropologia riduzionistica e positivista"[59]. Sulla scorta di queste argomentazioni il Test di Turing, ad avviso di alcuni, sarebbe stato confutato da John Searle e dal suo esperimento della stanza cinese[60]: egli contesta che la mente umana sia assimilabile ad un programma in

[55] A.M. TURING, *Computing machinery*, cit., p. 433.

[56] Per un approfondimento sul test di Turing v. G. ITALIANO, *Intelligenza Artificiale*, cit., p. 208; M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 24 ss.; M.B. MAGRO, *Robot, cyborg e intelligenze artificiali*, in A. CADOPPI, S. CANESTRARI, A. MANNA, M. PAPA, *Cybercrime*, cit., p. 1183; EAD., *Biorobotica, robotica e diritto penale*, in D. PROVOLO, S. RIONDATO, F. YENISEY, *Genetics*, cit., p. 510; P. MORO, *Biorobotica e diritti fondamentali. Problemi e limiti dell'intelligenza artificiale*, in D. PROVOLO, S. RIONDATO, F. YENISEY, *Genetics*, cit., p. 539. Per qualche cenno al potenziale superamento del test di Turing da parte di alcuni sistemi particolarmente sviluppati v. G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, in *agendadigitale.eu*, 11.6.2019, p. 2; M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale, responsabilità e accountability. Verso nuovi paradigmi?*, F. PIZZETTI, *Intelligenza artificiale*, cit., p. 335. *Contra* C. SALAZAR, *Umano, troppo umano...o no? Robot, androidi e cyborg nel "mondo del diritto" (prime notazioni)*, in *BioLaw Journal*, 2014, p. 260.

[57] L. FLORIDI, *What the Near Future of Artificial Intelligence Could Be*, in *Philos. Technol*., 2019, p. 2.

[58] M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 25.

[59] P. MORO, *Macchine come noi. Natura e limiti della soggettività robotica*, in U. RUFFOLO, *Intelligenza artificiale*, cit., p. 55. Tale punto di vista viene chiaramente esplicitato da G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 108, nota 16, il quale, richiamando il pensiero di Turing, chiarisce che per quest'ultimo non fosse tanto importante soffermarsi sulla possibilità che le macchine potessero pensare quanto, piuttosto, sulla loro capacità di riprodurre i risultati dell'"*ars cogitandi umana*". Per evidenziare quanto la sperimentazione di questi sistemi sia fondata e influenzata dal pensiero degli scienziati che vi lavorano, parte della dottrina mette in evidenza che in realtà la scienza non è davvero neutrale. In questo senso viene chiarito che «la costruzione del cosiddetto *imitation game* da parte di Alan Turing è indubbiamente condizionata da un'interpretazione della coscienza e della volontà della macchina basata sugli effetti della condotta reattiva agli impulsi esterni: tale interpretazione deriva dalla teoria del comportamentismo (o psicologia comportamentale), sviluppata dallo psicologo John Watson agli inizi del Novecento sull'assunto che il comportamento esplicito sia l'unica unità di analisi scientificamente studiabile, in quanto direttamente osservabile, mentre l'introspezione non potrebbe fornire alcun dato affidabile» P. MORO, *Libertà del robot?*, cit., pp. 533-534.

[60] J.R. SEARLE, *Minds, brains and programs*, in *Behavioral and Brain Sciences*, 1980, p. 419.

quanto, nonostante il computer sia in grado di elaborare ciò che può sembrare un'operazione mentale, in realtà non comprende realmente ciò che sta facendo[61].

Il fenomeno germogliò qualche anno più tardi, nell'estate del 1956, con il *Dartmouth Summer Research Project on Artificial Intelligence*[62], il primo workshop sull'intelligenza artificiale organizzato da John McCarhty, Marvin Minsky, Claude Shannon e Nathaniel Rochester. L'idea alla base del progetto era quella di prendere le mosse dalla supposizione che ogni aspetto dell'apprendimento o qualsiasi altra caratteristica dell'intelligenza potesse essere descritta in modo talmente tanto preciso da poter costruire una macchina in grado simularlo[63]. La paternità del termine "intelligenza artificiale" viene attribuita a John McCarhty[64] il quale, da ultimo, la definì come "la scienza e l'ingegneria per realizzare macchine intelligenti"[65].

La conferenza di Dartmouth diede ufficialmente il via ad uno sviluppo altalenante dell'intelligenza artificiale[66]. I successi ottenuti in questo settore generarono un'ondata di ottimismo che incentivò i finanziamenti della ricerca scientifica. Questo eccesso di positività[67], probabilmente, portò a un irragionevole innalzamento delle aspettative che, ben presto, finì per scontrarsi con le difficoltà tecniche dell'epoca[68]. Negli anni '70 iniziò il primo inverno dell'intelligenza

---

[61] P. MORO, *Macchine come noi*, cit., p. 59; ID., *Biorobotica e diritti fondamentali*, cit., p. 541; M.B. MAGRO, *Robot*, cit., p. 1184; EAD., *Biorobotica*, cit., p. 511.

[62] «We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer» J. MCCARTHY, M. MINSKY, N. ROCHESTER, C. SHANNON, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, 31.8.1955.

[63] G. ITALIANO, *Intelligenza Artificiale*, cit., p. 209.

[64] J. KAPLAN, *Intelligenza artificiale. Guida al futuro prossimo*, Roma, 2017, p. 37.

[65] J. MCCARTHY, *What is artificial intelligence?*, Stanford University, 12.11.2007, p. 2. Dello stesso A. v. anche J. MCCARTHY, P. HAYES, *Some Philosophical Problems From the Standpoint of Artificial Intelligence*, in *Machine Intelligence*, 1969. Tale definizione è riportata anche da A. VESPIGNANI, *L'algoritmo e l'oracolo. Come la scienza predice il futuro e ci aiuta a cambiarlo*, Milano, 2019, p. 65.

[66] C. ZHANG, Y. LU, *Study on artificial intelligence: The state of the art and future prospects*, in *Journal of Industrial Information Integration*, 2021, p. 1; X. YANG, Y. WANG, R. BYRNE, G. SCHNEIDER, S. YANG, *Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery*, in *Chem. Rev.*, 2019, p. 10521, la definiscono come una storia fatta di *boom* e *bust*.

[67] «L'intelligenza artificiale nasceva, insomma, con obiettivi molto ambiziosi, che apparvero ad alcuni, fin da subito, irrealistici, confusi e anche pericolosi. Ciò fece sì che, nel corso del tempo, non furono poche le polemiche, anche molto accese, che ne accompagnarono – e ne limitarono – lo sviluppo» M.B. MAGRO, *Robot*, cit., p. 1183.

[68] Il limitato potere di calcolo, l'irrisolvibilità di alcuni problemi e l'incapacità di gestire grandi quantità di dati erano alcuni degli scogli che le forme di intelligenza artificiale dell'epoca non riuscivano a superare. Tra i problemi che si registravano vi era il c.d. "paradosso di Moravec": alcuni compiti, anche particolarmente difficili per l'uomo (come dimostrare un teorema matematico) risultavano abbastanza semplici per un computer; di contro alcune azioni veramente semplici per l'uomo, come riconoscere un volto, risultavano molto complessi per le tecnologie dell'epoca, G. ITALIANO, *Intelligenza Artificiale*, cit., pp. 213 ss.

artificiale[69]: quando i ricercatori si trovarono davanti all'impossibilità di realizzare alcuni dei loro ambiziosi progetti i generosi finanziamenti si interruppero, rallentando drasticamente lo sviluppo dell'IA per quasi un decennio. Il settore dell'intelligenza artificiale vide un nuovo periodo di ripresa agli inizi degli anni '80, con l'introduzione dei primi sistemi esperti e degli investimenti ad essi destinati da parte di diverse industrie. L'eccessivo costo di questi sistemi li rese ben presto poco appetibili sul mercato portando a un secondo inverno dell'IA.

Si dovrà attendere intorno agli anni 2000 per assistere a una nuova estate dell'IA, dovuta non tanto al miglioramento delle tecniche nel settore quanto, piuttosto, all'aumento di potenza dei computer e della relativa capacità di calcolo computazionale[70], nonché alla crescita significativa (ad oggi potremmo dire illimitata) dei dati digitalizzati[71]. La più attenta dottrina ha sintetizzato i tre fondamentali fattori che spiegano come mai, proprio oggi, ci troviamo in una fase di florido sviluppo dell'intelligenza artificiale[72].

Il primo fattore è rinvenibile nell'arrivo degli algoritmi di *machine learning*, ossia modelli di calcolo in grado di imparare autonomamente a individuare relazioni precise fra dati senza avere alle spalle un modello pre-programmato: «l'algoritmo apprende dai dati adattando sé stesso man mano che impara dalle informazioni che sta elaborando»[73]. Il secondo fattore riguarda la disponibilità di hardware con una considerevole potenza di calcolo a prezzi non preclusivi. Il terzo ed ultimo fattore (probabilmente il più importante) è connesso all'illimitata quantità di dati oggi disponibile (iniziamo a familiarizzare con il concetto di "big data")[74].

Sembra di poter affermare, senza tema di smentita, che ci troviamo in una nuova estate dell'IA. Fatta questa breve premessa di carattere "storico" ci accingiamo adesso a cercare di fare chiarezza sul concetto stesso di intelligenza artificiale e sulle sue principali caratteristiche.

---

[69] Tale periodo storico può probabilmente coincidere con la pubblicazione nel 1969 di uno scritto, pubblicato nel volume *Perceptrons* da Marvin Minsky e Seymour Papert, che metteva in evidenza i limiti delle reti neurali artificiali, R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 29.

[70] R. ROVATTI, *Il processo di apprendimento algoritmico e le applicazioni nel settore legale*, in U. RUFFOLO, *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, cit., p. 32.

[71] A. VESPIGNANI, *L'algoritmo e l'oracolo,* cit., p. 66. G. ITALIANO, *Intelligenza Artificiale*, cit., p. 220; U. PAGALLO, *Intelligenza Artificiale e diritto. Linee guida per un oculato intervento normativo*, in *Sistemi intelligenti*, 2017, p. 615.

[72] M. GABBRIELLI, *Dalla logica al deep learning*, cit., pp. 26 ss.; R. ROVATTI, *Il processo di apprendimento algoritmico*, cit., pp. 32-33.

[73] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 66.

[74] L'importanza dirompente di questi tre elementi è stata confermata anche dagli studiosi del settore «The current AI boom started in the late 20th and early 21st centuries, driven by the rapid growth of stored data ("big data"), a concomitant increase in computing power (graphics processing units (GPUs), Google's tensor processing units (TPUs), etc.), and the continuous optimization of machine learning algorithms (e.g., deep learning)» X. YANG, Y. WANG, R. BYRNE, G. SCHNEIDER, S. YANG, *Concepts of Artificial Intelligence*, cit., p. 10523.

4. Di cosa stiamo parlando? Per una definizione di "intelligenza artificiale".

Cos'è l'intelligenza artificiale? Probabilmente trovare una definizione univoca di IA non è possibile[75], come (allo stato attuale) non è possibile trovare una "definizione giuridica" di intelligenza artificiale[76].

Dopo una prima definizione fornita da McCarhty – la quale, secondo gli esperti del settore, può a tutt'oggi essere complessivamente considerata valida[77] –, un passo in avanti venne fatto da Roger Schank nel 1987 il quale, prendendo in esame alcune caratteristiche tipiche dell'intelligenza umana e rapportandole all'intelligenza artificiale, affermò che «un programma di intelligenza artificiale che non apprende non è un programma di intelligenza artificiale»[78].

In realtà però la ricerca di una definizione comune di intelligenza artificiale non si è arrestata qui (e probabilmente non si arresterà). Come si è già avuto modo di anticipare, l'Unione Europea – nell'affrontare la materia – ha avvertito l'esigenza di individuare una definizione omnicomprensiva di IA, ritenendo che debbano considerarsi come dotati di intelligenza artificiale quei «sistemi software (ed eventualmente hardware) progettati dall'uomo che, dato un obiettivo complesso, agiscono nella dimensione fisica o digitale percependo il proprio ambiente attraverso l'acquisizione di dati, interpretando i dati strutturati o non strutturati raccolti, ragionando sulla conoscenza o elaborando le informazioni derivate da questi dati e decidendo le migliori azioni da intraprendere per raggiungere l'obiettivo dato. I sistemi di IA possono usare regole simboliche o apprendere un modello numerico, e possono anche adattare il loro comportamento analizzando gli effetti che le loro azioni precedenti hanno avuto sull'ambiente. Come disciplina scientifica, l'IA comprende diversi approcci e diverse tecniche, come l'apprendimento automatico (di cui l'apprendimento profondo e l'apprendimento per rinforzo sono esempi specifici), il ragionamento meccanico (che include la pianificazione, la programmazione, la rappresentazione delle conoscenze e il ragionamento, la ricerca e l'ottimizzazione) e la robotica (che comprende il controllo, la percezione, i sensori e gli attuatori e l'integrazione di tutte le altre tecniche nei sistemi ciberfisici)»[79].

---

[75] G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 2, il quale a sua volta cita M. IENCA, *Intelligenza*^{2}*. Per un'unione di intelligenza naturale e artificiale*, Torino, 2019; C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 1; F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 4; A. TURANO, *Robotica e roboetica*, cit., p. 128. Nella letteratura straniera R. CALO, *Artificial Intelligence Policy: a Primer and Roadmap*, in *SSRN*, 2017, p. 4.

[76] G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 107.

[77] X. YANG, Y. WANG, R. BYRNE, G. SCHNEIDER, S. YANG, *Concepts of Artificial Intelligence*, cit., p. 10521.

[78] R.C. SCHANK, *What's IA, Anyway?*, in *IA Magazine*, 1987, p. 64. L'A. nell'incipit del suo lavoro, prende in considerazione alcune caratteristiche che ci si aspetterebbe di trovare in un'entità intelligente (anche alternativamente e non cumulativamente) verificando se queste ultime potessero essere possedute dall'IA. Segnatamente l'A. si riferisce alle capacità comunicative, alla coscienza di sé, alla conoscenza della realtà esterna, all'intenzionalità dell'agire e, infine, alla creatività. Tale passaggio viene ripreso anche da G. HALLEVY, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, in *Akron Intellectual Property Journal*, 2010, pp. 175 ss.

[79] *A definition of AI: main capabilities and disciplines,* cit., p. 5. Un rimando a tale documento è contenuto anche nel Glossario degli *Orientamenti etici per un'IA affidabile*, Bruxelles, 8.4.2019, p. 45. Una precedente definizione di intelligenza artificiale era stata fornita dall'Unione Europea con il documento *L'intelligenza artificiale per l'Europa,* cit., p. 1. Per un approfondimento si rimanda al Cap. I, Sez. I, Parr. 3-7.

Gli studiosi (tanto delle discipline scientifiche quanto di quelle ascientifiche) hanno tentato di proporre definizioni più sintetiche e dalla portata più generica, in grado di essere comprese anche dai non addetti ai lavori.

Gli studiosi del settore[80] hanno definito l'intelligenza artificiale come una materia interdisciplinare[81] la cui attività è sintetizzabile come «l'uso della tecnologia per automatizzare attività che "normalmente richiedono l'intelligenza umana"»: in altre parole, tali sistemi condividono la caratteristica di svolgere compiti che, per essere realizzati dall'uomo, richiedono l'uso di processi cognitivi generalmente associati all'intelligenza umana[82]. L'IA «si serve dei computer per *simulare* i comportamenti intelligenti umani e addestra i computer ad apprendere comportamenti umani»[83]. Il riferimento alla simulazione del comportamento sembra aver definitivamente confermato le obiezioni mosse al test di Turing ed aver affermato che un sistema intelligente non può essere in grado di pensare autonomamente, bensì soltanto di imitare il pensiero umano. Oggi l'intelligenza artificiale può essere considerata una strategia di sviluppo per ogni Paese del mondo[84], nonché un campo di ricerca necessario in ogni settore della vita umana, incluso quello del diritto[85].

La dottrina giuridica che si è avvicinata al tema, muovendosi sulla falsariga di quanto sancito dagli esperti del settore, ha riconosciuto l'interdisciplinarietà della materia[86] definendo l'intelligenza artificiale come un "*umbrella term*", ossia un "contenitore" in grado di includere non solo diverse tecniche e processi di funzionamento[87] ma anche differenti settori tecno-scientifici[88].

---

[80] Oltre alla dottrina citata nel prosieguo si veda anche J.P. HATON, *A brief introduction to artificial intelligence*, in *IFAC*, 2006, pp. 8 ss.; M. RUPALI, P. AMIT, *A Review Paper on General Concepts of "Artificial Intelligence and Machine Learning"*, in *International Advanced Research Journal in Science, Engineering and Technology*, 2017, pp. 79 ss.

[81] «As a multidisciplinary field, AI involves integrating insights from diverse disciplines such as computer science, mathematics, psychology, linguistics, philosophy, neuro-science, artificial psychology, and many others» X. YANG, Y. WANG, R. BYRNE, G. SCHNEIDER, S. YANG, *Concepts of Artificial Intelligence*, cit., p. 10521; «AI is a compilation of computer science, logic, biology, psychology, philosophy, and many other disciplines, and it has achieved remarkable results in applications such as speech recognition, image processing, natural language processing, the proving of automatic theorems, and intelligent robots» C. ZHANG, Y. LU, *Study on artificial intelligence*, cit., p. 1.

[82] H. SURDEN, *Artificial Intelligence and Law: An Overview*, in *Georgia State University Law Review*, 2019, 1307. Sulla stessa scia si pongono altre definizioni fornite dagli studiosi del settore: «Artificial intelligence is the study of how to make computers perform intelligent tasks that, in the past, could only be performed by humans» C. ZHANG, Y. LU, *Study on artificial intelligence*, cit., p. 1; «The word "AI" is used in general when a computer simulation works, like thinking and problem solving, which humans associate with another human mind» H. VARSHNEY, R. A. KHAN, U. KHAN, R. VERMA, *Approaches of Artificial Intelligence and Machine Learning in Smart Cities: Critical Review*, in *IOP Conf. Ser.: Mater. Sci. Eng.*, 2021, p. 1.

[83] C. ZHANG, Y. LU, *Study on artificial intelligence*, cit., p. 1, corsivo nostro. Il concetto dell'"imitazione" delle funzioni cognitive umane viene ripreso da A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 65.

[84] C. ZHANG, Y. LU, *Study on artificial intelligence*, cit., p. 1.

[85] H. VARSHNEY, R. A. KHAN, U. KHAN, R. VERMA, *Approaches of Artificial Intelligence*, cit., p. 2.

[86] C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 98; G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 8.

[87] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 8; R. CALO, *Artificial Intelligence Policy*, cit., p. 5.

Nella branca dei saperi non strettamente scientifici lo scopo dell'intelligenza artificiale è stato individuato nel «far sì che una macchina si comporti in modi che sarebbero definiti intelligenti se un essere umano si comportasse così»[89]. In altre parole, i sistemi intelligenti devono essere in grado di porre in essere comportamenti "*human like*", ossia in riproduzione dell'intelligenza umana[90].

È stata avvertita, specie dalla dottrina giuridica, l'esigenza di una definizione di IA che sia flessibile e non ostacoli l'evoluzione scientifica[91], pur dovendo

---

[88] L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., p. 2.

[89] L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., p. 2. Sulla stessa scia v. M.B. MAGRO, *Biorobotica*, cit., p. 508 ad avviso della quale l'intelligenza artificiale «indica sia un termine che una disciplina, ed esprime l'idea che si possa emulare ogni aspetto dell'intelligenza umana. Sebbene siano state proposte parecchie, talora anche contraddittorie, definizioni di tale disciplina, ciò che le accomuna e connota lo spirito della IA è quello di imitare, riprodurre per mezzo di macchine elettroniche l'attività mentale umana, ovvero quella che costituisce la facoltà più essenziale dell'uomo»; C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 102, richiama la seguente definizione di IA, intesa come «the science of making machines do things that would require intelligence if done by men»; analogamente M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 334, richiamano la definizione dell'*Oxford Dictionary of Philosophy* che definisce l'IA come «the science of making machines that can do of thing that humans can do». Per ulteriori proposte definitorie avanzate in materia di IA v. U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 615; R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 262; G. MOBILIO, *L'intelligenza artificiale*, cit., p. 401; C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1914; C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 100. Per un'analisi delle definizioni fornite negli Stati Uniti v. *Fundamentally Understanding The Usability and Realistic Evolution of Artificial Intelligence Act of 2017*, anche noto come *Future of Artificial Intelligence Act of 2017* «The term "artificial intelligence" includes the following: (A) Any artificial systems that perform tasks under varying and unpredictable circumstances, without significant human oversight, or that can learn from their experience and improve their performance. Such systems may be developed in computer software, physical hardware, or other contexts not yet contemplated. They may solve tasks requiring human-like perception, cognition, planning, learning, communication, or physical action. In general, the more human-like the system within the context of its tasks, the more it can be said to use artificial intelligence. (B) Systems that think like humans, such as cognitive architectures and neural networks. (C) Systems that act like humans, such as systems that can pass the Turing test or other comparable test via natural language processing, knowledge representation, automated reasoning, and learning. (D) A set of techniques, including machine learning, that seek to approximate some cognitive task. (E) Systems that act rationally, such as intelligent software agents and embodied robots that achieve goals via perception, planning, reasoning, learning, communicating, decision making, and acting». Per una definizione più recente fornita negli USA v. *John S. McCain National Defense Authorization Act For Fiscal Year 2019* «the term "artificial intelligence" includes the following: (1) Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets. (2) An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action. (3) An artificial system designed to think or act like a human, including cognitive architectures and neural networks. (4) A set of techniques, including machine learning, that is designed to approximate a cognitive task. (5) An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision making, and acting». Per le definizioni accolte nel Regno Unito v. *AI in the UK: ready, willing and able?*, in *Select Committee on Artificial Intelligence, House of Lords, Report Sessions*, 2017-19, pp. 11 ss.

[90] M.B. MAGRO, *Robot*, cit., p. 1182.

[91] C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 7, la quale richiama il Considerando C della *Risoluzione del Parlamento europeo*, cit.

ammettere che verosimilmente tale definizione sarà destinata a mutare col tempo. L'intelligenza artificiale potrebbe rientrare appieno nel concetto di "*generative technology*", ossia quelle tecnologie che offrono molteplici possibilità di innovazione e che, per ciò stesso, sono destinate a cambiare nel tempo «man mano che le persone lavorano con e attraverso la nuova tecnologia»[92].

Prendendo atto dell'"ambiguità concettuale" di questa nuova forma di intelligenza, parte della dottrina ha anche proposto di fornire una definizione in negativo di IA, «il cui fine è l'eliminazione di un qualsiasi tentativo di assimilazione dell'intelligenza artificiale al pensiero e alla razionalità – per non dire emotività – umana»[93].

Per fugare ogni dubbio derivabile da un errato accostamento dell'intelligenza artificiale a quella umana è stato altresì proposto di parlare, piuttosto, di *razionalità artificiale*: «laddove per "razionalità" si intende la capacità di scegliere la migliore azione da intraprendere per conseguire un determinato obiettivo alla luce di alcuni criteri di ottimizzazione delle risorse a disposizione»[94].

Infine, da un punto di vista che potremmo definire "umanistico", l'intelligenza artificiale è stata intesa anche come *costrutto sociale*, proprio per la sua capacità di influire sulle concezioni classiche dell'ordinamento, tanto in ottica sociale quanto in ottica giuridica, anche penale[95].

5. Tipologie e caratteristiche dell'IA: il machine learning.

Dopo aver fornito una panoramica delle proposte definitorie avanzate in materia di intelligenza artificiale, tentando di effettuare una sorta di *reductio ad unum*, potrebbe essere utile adesso cercare di individuarne le principali caratteristiche.

Preliminarmente occorre chiarire che l'IA può essere intesa sia come *software* (ossia un sistema privo di un supporto fisico) che come *hardware* (sistema, invece, dotato di una componente materiale)[96]. In tal senso si usa distinguere macchine *embodied*, ossia dotate di un corpo, e *non embodied*, sprovviste di un

---

[92] J.B. BALKIN, *The Path of Robotics Law*, in *California Law Review Circuit*, 2015, p. 47.

[93] G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 107. In tal sede l'A. dà conto del fatto che alcuni studi hanno proposto 8 diverse definizioni di intelligenza artificiale, divisibili in 4 possibili approcci, intendendo l'IA come quella scienza che studia sistemi in grado di: pensare in modo umano, agire come un umano, pensare razionalmente e agire razionalmente. L'A. si riferisce alla sistematica proposta da Russell e Norvig i quali hanno recentemente pubblicato la quarta edizione del celebre, S. RUSSEL, P. NORVIG, *Artificial Intelligence. A modern approach*, US, 2020.

[94] F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 5 il quale, a sua volta, richiama S. RUSSELL, P. NORVIG, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 3ª edizione, 2009, pp. 36 ss. Menzionano il concetto di "razionalità" anche M.B. MAGRO, *Robot*, cit., p. 1181, la quale parla di «capacità di seguire una logica consequenziale e capacità di modellazione» e G. CAPILLI, *I criteri di interpretazione*, cit., p. 465.

[95] C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1940.

[96] Tra i sistemi software si ricordano i programmi per eseguire transazioni finanziarie, per la traduzione automatica di testi, per il riconoscimento facciale o per realizzare forme di giustizia predittiva. Tra i sistemi hardware si annoverano le auto a guida autonoma, i veivoli a pilotaggio remoto o i bracci robotici. Sul punto v. I. SALVADORI, *Agenti artificiali*, cit., p. 91; F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 6; A. TURANO, *Robotica e roboetica*, cit., p. 130.

corpo fisico[97]. In realtà la distinzione più nota che si è soliti fare all'interno della categoria dell'intelligenza artificiale è quella fra sistemi di IA deboli e forti[98].

L'IA *debole* si riferisce a quei sistemi che svolgono compiti ben definiti, secondo regole pre-programmate, per raggiungere obiettivi specifici in modo efficiente[99]. Tale concezione di IA riprende l'orientamento secondo cui le macchine non sarebbero in grado di "pensare autonomamente", limitandosi a *riprodurre* i processi intellettivi tipici dell'uomo[100]. L'intelligenza artificiale debole mantiene dunque un ruolo subalterno rispetto all'uomo, in quanto sprovvista di coscienza[101] e di intenzionalità. Parte della dottrina ha molto insistito, in particolare, sulla portata di tale ultimo elemento, ritenendo che l'intenzionalità sia una declinazione fondamentale dell'intelligenza umana e che, non essendo posseduta dalle macchine, non consentirebbe di omologare la mente umana alla "mente artificiale"[102]. In altri termini «in una prospettiva "*debole*", le macchine si comportano *come se fossero effettivamente pensanti*»[103].

Parlando, invece, di IA *forte* ci si riferisce a sistemi che non sono meri strumenti nelle mani dell'uomo, bensì entità capaci di porre in essere veri e propri processi cognitivi[104]. I sostenitori dell'esistenza di un'IA forte ritengono che non vi sia una reale differenza qualitativa tra intelligenza umana e artificiale[105]. Ciò che diversificherebbe l'uomo dalla macchina sarebbe soltanto il supporto fisico esterno: il corpo in carne e ossa dell'uomo da un lato e l'hardware della macchina dall'altro[106]. In parole povere, «nella versione "*forte*", l'IA punta a conseguire un

---

[97] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 25.

[98] Sulla distinzione classica tra *strong AI* e *weak AI* v. J.R. SEARLE, *Minds, brains and programs*, cit., p. 417. M.B. MAGRO, *Robot*, cit., pp. 1183-1184 distingue queste due categorie in intelligenza artificiale generale (forte) e circoscritta (debole). Una distinzione a livello normativo è rinvenibile nel Glossario della *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, cit., p. 47 nei seguenti termini «intelligenze artificiali "forti" (capaci di contestualizzare problemi specializzati di varia natura in maniera completamente autonoma) e intelligenze artificiali "deboli" o "moderate" (alte prestazioni nel loro ambito di addestramento)». Critico su tale distinzione G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 3.

[99] R. ROVATTI, *Il processo di apprendimento algoritmico*, cit., p. 36.

[100] M.B. MAGRO, *Robot*, cit., 1184; EAD., *Biorobotica*, cit., p. 511.

[101] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1759. Altra dottrina chiarisce che «la coscienza umana non è riducibile ad una sola delle sue funzioni e, dunque, il tentativo di paragonare la mente umana al *computer* è destinato a fallire. La coscienza è una proprietà dell'intera persona e deve essere considerata nella totalità che contraddistingue l'intelligenza umana, sicché il tentativo di riprodurne una semplice parte attraverso modelli computazionali anche molto evoluti è destinato a fallire, anche perché genera un problema che non è computabile» P. MORO, *Libertà del robot?*, cit., p. 530.

[102] Da ultimo su questo tema M.B. MAGRO, *Robot*, cit., p. 1184. Invero il tema era già stato arato da J.R. SEARLE, *Minds, brains and programs*, cit., pp. 421 ss.

[103] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1747, corsivo dell'A.

[104] J.R. SEARLE, *Minds, brains and programs*, cit., p. 417. Tale ultimo A. viene richiamato anche da M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 21, nota 2, il quale afferma che Searle, con il suo esperimento della stanza cinese, abbia negato la possibilità di costruire una macchina davvero pensante e, quindi, la possibilità di creare intelligenze artificiali "forti". Si dedica al tema anche P. MORO, *Libertà del robot?*, cit., pp. 531-532.

[105] M.B. MAGRO, *Biorobotica*, cit., p. 510, l'A. individua proprio in Alan Turing il padre dell'intelligenza artificiale forte; F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 231.

[106] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1759.

livello di intelligenza pari a quello dell'uomo, dunque *macchine pensanti (human like)*»[107].

Tale ripartizione ha portato ad un'ulteriore divisione in dottrina, che vede contrapporsi la teoria strutturalista alla teoria funzionalista.

La teoria *strutturalista*, associata all'IA debole, sostiene «la non autenticità del pensiero meccanico e la diversità ontologica-qualitativa tra intelligenza artificiale e intelligenza naturale: se è differente l'hardware, è diverso anche il software»[108]. L'intelligenza artificiale non può manifestare l'intenzionalità tipica dell'intelligenza umana, essa si comporta "come se" capisse ma non vi è, alla base del suo agire, una reale comprensione[109]. Secondo questa tesi ciò che conta non è tanto la funzione dell'agire del sistema intelligente, quanto la sua struttura, ossia dove si svolge e si articola il "pensiero"[110].

La teoria *funzionalista*, accostata invece all'IA forte, accoglie una soluzione inversa: essa non guarda alla struttura quanto, piuttosto, alla funzione. Fondamento strutturale di tale teoria è l'"isomorfismo mente-computer" in senso funzionale[111]: «un computer è paragonabile ad un essere umano quanto ad intelligenza, se le prestazioni svolte dal computer non possono essere distinte da quelle svolte dagli esseri umani»[112]. Sarebbe dunque impossibile distinguere le funzioni cognitive della macchina da quelle dell'uomo[113], indipendentemente dal fatto che l'articolazione del pensiero avvenga all'interno di una scatola cranica o di un meccanismo fatto di metallo e circuiti.

Per quanto ci si senta di condividere in tal sede l'orientamento dottrinale che nega una piena assimilazione dell'intelligenza artificiale a quella umana[114] –

---

[107] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1747.
[108] M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 13.
[109] M.B. MAGRO, *Robot,* cit., p. 1184.
[110] M.B. MAGRO, *Biorobotica*, cit., p. 512.
[111] M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 6.
[112] M.B. MAGRO, *Biorobotica*, cit., p. 510.
[113] M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 7.
[114] «Le macchine intelligenti, anche le più raffinate ed evolute (grazie al crescere delle informazioni a cui possono accedere) non possono essere paragonate all'uomo, senza ridurre quest'ultimo ad un profilo comportamentale monodimensionale e rigido» M.B. MAGRO, *Robot*, cit., p. 1195; le macchine sono inconsapevoli, non hanno coscienza di sé, P. MORO, *Macchine come noi*, cit., p. 59; esse non sono per natura munite di creatività o intuito, ID., *Biorobotica e diritti fondamentali*, cit., p. 541-542; le macchine non sono dotate di empatia C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 124; esse, ancora, sono sprovviste del c.d. "senso comune", «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» M.B. MAGRO, *Biorobotica*, cit., p. 512, EAD., *Robot,* cit., p. 1195; sul concetto di "senso comune" pocanzi citato, sulle caratteristiche tipiche dell'intelligenza umana e sulla ripartizione del tipo di attività che possono considerarsi "intelligenti" (incluse quelle non governate da regole formali e, pertanto, non riproducibili dall'IA) v. G. FORNERO, *Intelligenza artificiale e filosofia*, in N. ABBAGNANO, *Storia della filosofia*, Vol. IV, G. FORNERO, F. RESTAINO, D. ANTISERI (a cura di), *La filosofia contemporanea*, Milano, 2013, pp. 536 ss. «L'uomo è dotato di una mente colorata, come ricorda Omero a proposito di Ulisse, basata sulla multiformità dell'ingegno e su plurime dimensioni della razionalità che, come ha accertato la psicanalisi del Novecento, provengono anche dall'inconscio. Secondo Sigmund Freud, la vita psichica si svolge in modalità performative anche nelle sue forme secondarie – *lapsus*, dimenticanze, sogni – che malcelano un'organizzazione alternativa della razionalità soggettiva: tali forme del pensare sono del tutto ignote all'intelligenza meccanica» P. MORO, *Libertà del robot?*, cit., p. 530.

proprio in considerazione delle peculiarità esclusive di quest'ultima – si è ritenuto comunque opportuno dar conto di questa distinzione in quanto, ad una preliminare disamina, il dubbio può legittimamente insinuarsi. I sistemi intelligenti di nuova generazione[115], infatti, sono in grado di palesare un'intelligenza in tutto e per tutto analoga a quella umana[116] (prescindendo dall'ormai affrontata questione se sia una reale intelligenza o una mera simulazione), di percepire l'ambiente esterno, di adattarsi e di imparare da esso. Cerchiamo di approfondire il tema prendendo in esame le caratteristiche tipiche dell'IA.

La dottrina ha efficacemente sintetizzato le principali caratteristiche dell'intelligenza artificiale nei seguenti punti «a) l'uso di grandi quantità di dati e informazioni; b) una elevata capacità logico-computazionale[117]; c) l'uso di nuovi algoritmi, come quelli del *deep learning* e dell'*auto-apprendimento*, che definiscono metodi per estrarre conoscenza dai dati per dare alle macchine la capacità di prendere decisioni corrette in vari campi di applicazione»[118]. Nel prosieguo della trattazione prenderemo in esame i requisiti di maggior interesse, ossia il ruolo rivestito dai dati nello sviluppo dell'IA e l'incidenza delle tecniche di apprendimento automatico.

In tal sede vorremmo concentrarci – senza alcuna pretesa di esaustività – sull'ultimo dei requisiti appena elencati, ossia l'uso di algoritmi di ultima generazione per programmare i sistemi intelligenti e sulle conseguenze del loro uso.

Preliminarmente occorre chiarire che col termine *algoritmo* si intende riferirsi a «sequenze di istruzioni, linee di codice, che possiamo scrivere e che possiamo leggere»[119]. Gli algoritmi "classici" sono ideati per affrontare e risolvere un dato problema[120] attraverso un numero predefinito di passaggi[121]. Il vero passo in avanti, ossia ciò che rende l'agire dei sistemi intelligenti tanto simile all'agire umano, è determinato dagli algoritmi di *machine learning*.

Il *machine learning* (d'ora in poi anche ML), noto anche come *apprendimento automatico*, è uno dei principali sottocampi dell'IA[122]. In realtà il ML annovera al suo interno una vera e propria famiglia di tecniche di intelligenza artificiale che condividono alcune caratteristiche: la maggior parte dei metodi di *machine learning* funziona rilevando e individuando schemi utili contenuti

---

[115] A. CAPPELLINI, Machina delinquere non potest?, cit., pp. 6-7 parla di entità artificiali di seconda generazione; C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 135 parla di IA di quarta generazione.

[116] Nonostante si tratti di processi volti solo a "riprodurre" il pensiero umano, ormai i sistemi più evoluti hanno raggiunto un tale livello di complessità da rendere difficile distinguere il pensiero vero e proprio dall'imitazione del pensiero medesimo, F. CAROCCIA, *Soggettività giuridica dei robot?*, p. 220.

[117] Per qualche ulteriore considerazione sui sistemi di *cognitive computing* v. M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 7; M.B. MAGRO, *Robot*, cit., p. 1187.

[118] M.B. MAGRO, *Robot*, cit., p. 1182, richiamata anche in G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 4.

[119] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 3.

[120] M. DI FLORIO, *Il diritto penale che verrà*, cit., p. 7, il quale richiama a sua volta C. TOFFALORI, *Algoritmi*, Bologna, 2015.

[121] A. TURANO, *Robotica e roboetica*, cit., p. 129, nota 18, il quale a sua volta richiama A. BUTTERFIELD, G.E. NGONDI, *A Dictionary of Computer Science*, Oxford, 2016.

[122] H. VARSHNEY, R. A. KHAN, U. KHAN, R. VERMA, *Approaches of Artificial Intelligence*, cit., p. 3.

all'interno di grandi quantità di dati[123]. Obiettivo del ML è quello di «consentire ai computer di apprendere autonomamente. L'idea chiave del *machine learning* è che l'algoritmo impara da solo a identificare relazioni precise nei dati osservati, senza avere regole e modelli espliciti pre-programmati. In questo modo, invece di scrivere algoritmi con milioni di istruzioni specifiche per svolgere un compito particolare, l'algoritmo apprende dai dati adattando sé stesso man mano che impara dalle informazioni che sta elaborando»[124]. Tali sistemi, in altri termini, imparano dalla loro esperienza, adeguandosi agli stimoli ricevuti e modificando di conseguenza il loro comportamento[125].

L'apprendimento automatico si riferisce ad algoritmi in grado di "apprendere" e di migliorare le prestazioni delle attività cui sono applicati, attività che si presumeva, un tempo, essere ad esclusivo appannaggio della cognizione umana (basti pensare alle auto a guida autonoma, agli algoritmi predittivi, al riconoscimento vocale e, financo, al campo del diritto)[126]. Gli algoritmi di *machine learning* apprendono mediante un'analisi dinamica[127] di un'enorme moltitudine di dati provenienti, anche, dalle nostre interazioni con i dispositivi tecnologici che ne fanno da hardware[128]. I sistemi di *machine learning* si dividono in:

- *Supervised learning* (apprendimento supervisionato): all'algoritmo vengono forniti un insieme di esempi sotto forma di relazioni tra input e output in modo che quest'ultimo riesca a trarre dal legame tra dati in entrata e dati in uscita una regola generale da poter applicare anche in casi non predeterminati[129]. Gli esempi forniti su cui l'algoritmo viene addestrato per riconoscere i modelli ricorrenti vengono denominati "set di formazione" e l'obiettivo di tale *training* è, in altre parole, consentire all'algoritmo di creare un modello informatico interno di un dato fenomeno che possa essere generalizzato per essere poi applicato a nuovi esempi mai visti di quel fenomeno[130]. L'apprendimento si definisce "supervisionato" in quanto l'algoritmo formula le sue previsioni e viene corretto dal supervisore quando sbaglia, così apprendendo dall'esperienza e non ripetendo più l'errore[131];
- *Unsupervised learning* (apprendimento non supervisionato): viene rimesso agli algoritmi medesimi il compito di classificare i dati (ad essi non viene fornita

[123] H. SURDEN, *Artificial Intelligence and Law*, cit., p. 1311, l'A. chiarisce che questi sistemi (e i loro schemi) vengono applicati in varie attività (ad es. nel guidare un'auto o nel rilevamento di frodi), con modalità tali da produrre risultati utili e nonché comportamenti apparentemente intelligenti.
[124] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 66.
[125] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 5.
[126] H. SURDEN, *Machine Learning and the Law*, in *Washington Law Review*, vol. 89, n. 1, 2014, p. 88.
[127] A. TURANO, *Robotica e roboetica*, cit., p. 131.
[128] G. ITALIANO, *Intelligenza Artificiale*, cit., p. 217.
[129] G. CAPILLI, *I criteri di interpretazione*, cit., p. 465.
[130] H. SURDEN, *Machine Learning and the Law*, cit., p. 93. Per far comprendere al lettore il succitato meccanismo l'A. riporta l'esempio delle e-mail spam. All'algoritmo vengono sottoposte una serie di e-mail che un essere umano ha già in precedenza predeterminato come spam. In tal modo, analizzando gli esempi fornitigli, il sistema apprende le caratteristiche delle e-mail spam, riconoscendole e distinguendole dalle e-mail regolari.
[131] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 67.

una "etichetta" dall'uomo)[132] e scoprire il legame che intercorre tra essi. Per far ciò si utilizzano algoritmi di *clustering*, volti cioè ad individuare diversi sottogruppi all'interno di un insieme di dati[133];

- *Reinforced learning* (apprendimento con rinforzo): l'algoritmo viene lasciato libero di fare tentativi, provare schemi differenti e imparare dai suoi errori[134]. Infatti ad ogni tentativo fatto dal sistema seguirà un segnale (in termini di sanzione o ricompensa)[135] che indichi se la scelta è stata o meno efficace[136], così consentendo al sistema di imparare e minimizzare i suoi errori.

Per ognuna di queste metodologie possono essere sviluppati algoritmi dal diverso funzionamento: non esiste un algoritmo migliore di un altro, esistono piuttosto specifici algoritmi più o meno adeguati in base alle peculiarità del problema che si deve risolvere[137].

La differenza tra gli algoritmi di *machine learning* e quelli tradizionali appare intuitiva anche per un non addetto ai lavori: mentre questi ultimi vengono usati dal programmatore per impartire un'indicazione chiara alla macchina, gli algoritmi di ML si servono di modelli "sub-simbolici", ove cioè il sistema è capace di apprendere dall'analisi di un cospicuo insieme di dati[138]. Mentre gli algoritmi tradizionali sono trasparenti e spiegabili, gli algoritmi di *machine learning* funzionano diversamente. La differenza principale è che gli algoritmi di *machine learning* consentono «a un sistema automatico di imparare e di creare un modello del problema che bisogna risolvere. Un modello che non è però spesso facilmente interpretabile come le linee di codice degli algoritmi tradizionali»[139]. La svolta dell'apprendimento automatico rispetto ai classici metodi di programmazione appare, pertanto, di tutta evidenza[140], potendo tali evoluti sistemi prendere "decisioni" autonomamente e controllare il proprio agire[141].

---

[132] M. RUPALI, P. AMIT, *A Review Paper on General Concepts of "Artificial Intelligence and Machine Learning"*, in *International Advanced Research Journal in Science, Engineering and Technology*, 2017, p. 81.

[133] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 68.

[134] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 68.

[135] M. RUPALI, P. AMIT, *A Review Paper*, cit., p. 81.

[136] G. CAPILLI, *I criteri di interpretazione*, cit., p. 466.

[137] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 69. L'A. spiega infatti che non sarà possibile servirsi dello stesso algoritmo per il riconoscimento delle immagini e per comprendere i gusti cinematografici di una persona.

[138] M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 26.

[139] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit. p. 4. L'A. propone un esempio particolarmente forte ma in grado di spiegare perfettamente l'opacità del sistema: «per fare un esempio molto approssimativo, e forse anche un po' esagerato, potremmo pensare al modello prodotto da un algoritmo di machine learning come a un cervello umano. Non possiamo dissezionare il cervello per capire cosa è successo, perché è stata presa una certa decisione» p. 4.

[140] M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 2, «L'apprendimento automatico è un nuovo approccio all'automazione e costituisce una svolta epocale nell'evoluzione delle IA. Mentre nella programmazione classica si costruisce un programma per risolvere un problema sulla base di una pregressa ed esatta conoscenza della soluzione del problema (o delle informazioni da cui ottenere la conoscenza necessaria per risolvere il problema), con l'apprendimento automatico, invece, il programmatore costruisce un modello che "trova" e "impara" la soluzione a quel problema stesso, senza che lo stesso programmatore conosca la soluzione o persino abbia previsto il problema in termini matematici. Questa nuova tecnologia supera o elude la mancanza di conoscenza dell'uomo in due sensi: in termini di ideazione del problema e della sua soluzione e in

Tra le tecniche più evolute di *machine learning* si ricordano le *neural networks* e il *deep learning*[142]. Le prime, ossia le reti neurali, simulano il funzionamento di un insieme di neuroni che, stimolati da un input in entrata, restituiscono un output in uscita[143]. Il secondo tipo di algoritmo (l'apprendimento profondo o rete neurale profonda) si fonda su un fitto insieme di reti neurali dislocate su un numero più elevato di livelli, in grado di produrre piccoli miglioramenti della rete durante la fase di *training*[144]. Tale ultima tecnologia si candida a diventare il futuro dell'IA[145], essendo invero già alla base dei sistemi di riconoscimento vocale e di immagini[146].

Iniziamo a comprendere che i sistemi informatici di cui ci stiamo interessando sono dotati di un sempre più crescente grado di autonomia, il che rende il loro comportamento non completamente specificato dal programmatore né completamente prevedibile, essendo piuttosto il risultato dell'implementazione delle capacità del sistema. Occorre considerare però che parte della dottrina ha opportunamente ricordato come l'autonomia di questi sistemi, in realtà, non sia assoluta bensì relativa, in quanto limitata ai compiti loro assegnati[147].

La graduale perdita di controllo da parte del programmatore che stiamo riscontrando si acuisce maggiormente se pensiamo al fatto che le intelligenze artificiali possono apprendere non solo dalla propria esperienza, ma anche da quella dei loro "simili", grazie alle tecnologie di *cloud computing*. «Collegare tra loro più soggetti artificiali e renderli capaci di scambiarsi informazioni in cloud permette infatti di "sommare" le "esperienze di vita" di una moltitudine di macchine intelligenti, sottoposte agli scenari più diversi»[148]. In questo modo verrà incrementata non soltanto la capacità di apprendimento di questi sistemi[149] ma anche la rapidità di questo apprendimento[150]. Tali forme di *interactive learning*[151] danno vita a un insieme potenzialmente illimitato di scenari e a nuove forme di rischio, nonché ad una pericolosa diminuzione del grado di controllo che l'uomo è in grado di esercitare sulla macchina[152].

---

termini di conoscenza degli strumenti con cui la soluzione potrebbe essere prodotta automaticamente».

[141] M.B. MAGRO, *Robot*, cit., p. 1202.

[142] H. SURDEN, *Artificial Intelligence and Law*, cit., p. 1311.

[143] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 70.

[144] G. ITALIANO, *Intelligenza Artificiale*, cit., pp. 219-220.

[145] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 3.

[146] A. TURANO, *Robotica e roboetica*, cit., p. 131, nota 24.

[147] M.B. MAGRO, *Robot*, cit., p. 1202. Altra parte della dottrina ha chiarito che tali macchine, in realtà, non sono realmente autonome in quanto non sono capaci di regolarsi autonomamente. Esse sono piuttosto "automatiche", ossia in grado di "muoversi da sole", ma ciò dà loro solo una "parvenza di autonomia": «ogni loro comportamento è scritto nel loro DNA artificiale, o meglio prescritto dai loro costruttori, programmatori o utilizzatori umani» A. CAPPELLINI, Machina delinquere non potest?, cit., p. 5.

[148] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 6. Un cenno ai sistemi di collegamento in *cloud* è rinvenibile anche in G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 6.

[149] R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 265.

[150] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 6.

[151] P. MORO, *Macchine come noi*, cit., p. 46.

[152] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 6.

5.1. L'imprevedibilità dell'IA: il black box effect.

Dalla disamina appena svolta sembra possibile trarre alcuni punti fermi. I sistemi di intelligenza artificiale e di *machine learning* portano con sé il vantaggio di individuare schemi e correlazioni che l'essere umano non sarebbe in grado di cogliere e di farlo anche con una certa rapidità[153]. Tali capacità rendono questi sistemi reattivi (stante la loro capacità di reagire agli stimoli esterni), proattivi (ossia capaci di determinare e realizzare il loro agire) e interattivi[154] (in grado, cioè, di interagire tra di loro e con gli uomini)[155]. Si tratta, a ben vedere, di caratteristiche tipiche dell'essere umano, il che giustifica il dubbio precedentemente proposto sull'assimilabilità dell'intelligenza artificiale e di quella umana.

Di contro, il potenziale livello di autonomia di tali sistemi è in grado di generare problemi concernenti la trasparenza del sistema e il controllo che il programmatore è in grado di esercitare su di essi. Quest'ultimo, in un certo senso, si limita a creare l'algoritmo il quale impara grazie alla propria esperienza, potendo così pervenire a decisioni non originariamente prevedibili da parte del suo ideatore[156]. L'interazione con il mondo esterno rende tendenzialmente illimitato lo "spettro di variabili" che possono incidere sul comportamento del sistema intelligente[157]. Ad avviso di parte della dottrina, infatti, piuttosto che di programmazione, sarebbe più corretto parlare di "educazione", vista la capacità del sistema di imparare in autonomia[158].

Quindi, il comportamento delle più evolute forme di IA può essere *ex ante* imprevedibile[159]: la condotta del sistema, per quanto riconducibile al programma inserito dall'uomo, potrebbe non essere stata prevista da quest'ultimo[160]. Come è stato condivisibilmente osservato dalla dottrina, l'imprevedibilità dell'agire dell'IA è un fattore che va considerato giuridicamente: «sarebbe profondamente iniquo, in tal caso, ritenere i creatori dell'algoritmo responsabili per i danni causati dalla loro invenzione ma, al contempo, una tale soluzione lascerebbe un delicato vuoto normativo», in quanto i soggetti danneggiati dall'agire illecito

---

[153] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 72.

[154] A proposito del c.d. *interactive learning*, inteso come il processo di apprendimento con cui un sistema intelligente impara operando in sinergia con l'uomo, v. P. MORO, *Libertà del robot?*, cit., pp. 526-527. Per quanto l'A. si riferisca invero alla c.d. *interaction human-robot*, riteniamo che queste considerazioni siano valevoli anche a proposito della più ampia categoria dell'intelligenza artificiale. Cercheremo di delineare la distinzione tra IA e robotica nel Par. 7 della presente Sezione.

[155] P. MORO, *Biorobotica e diritti fondamentali*, cit., p. 540; ID., *Libertà del robot?*, cit., p. 536. Sulla reattività e interattività di tali sistemi v. anche M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 2.

[156] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1746.

[157] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 83.

[158] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 10.

[159] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 7; M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 357.

[160] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 344. Altra dottrina osserva che «i possibili eventi lesivi prodotti dall'intelligenza artificiale, pertanto, potranno dipendere sempre da come questa è stata progettata per essere autonoma e dal fatto che l'autonomia fa sì che il robot sia fuori dal controllo dell'essere umano» G. CAPILLI, *I criteri di interpretazione*, cit., p. 482.

dell'IA resterebbero privi di ristoro, ferme restando le consistenti difficoltà in punto di ricostruzione del nesso causale[161].

Se, in un certo senso, possiamo affermare che l'output finale fornito dal sistema sia sempre noto ad un osservatore esterno, ciò che non è né visibile né comprensibile è l'iter logico che lo ha generato[162]. Non solo le risposte fornite da questi sistemi sono connotate da un ineliminabile grado di autonomia ma, inoltre, gli uomini che li hanno programmati non sono completamente in grado di spiegarne il funzionamento[163]. Veniamo dunque al secondo problema che connota i sistemi di *machine learning*: la mancanza di trasparenza. La complessità del movimento dei dati all'interno delle reti neurali, specie ove si tratti di reti neurali profonde, rende sostanzialmente impossibile «tracciare i passaggi logico-argomentativi che permettono alla macchina di decidere in autonomia»[164].

Tale problema è universalmente conosciuto come *black box*: questi «algoritmi prendono un dato in entrata e ne producono uno in uscita, passando attraverso un processo di apprendimento che è una scatola nera non interpretabile dall'esterno»[165]. Tra i dati inseriti in input dall'uomo e gli esiti forniti in output dal sistema v'è un "vuoto di comprensione"[166], il procedimento che li lega resta avvolto da un alone di opacità[167] incidendo significativamente sulla piena comprensione del sistema da parte dell'uomo[168]. L'effetto della *black box* è quello di «ostacolare la verifica sulle singole fasi interne al procedimento, impedendo la possibilità di un controllo sulla congruità delle motivazioni alla base della decisione»[169].

Il funzionamento di questi sistemi, insomma, non è spiegabile[170]. Tale problematica genera interrogativi di non secondario rilievo: come possiamo esser certi dell'affidabilità di questi sistemi? Quanto possiamo fidarci della bontà dei loro output? Ancora, «gli algoritmi producono davvero conoscenza se in realtà non approfondiscono la nostra comprensione di come funziona il mondo?»[171].

Il *black box effect* non solo rende impossibile ricostruire il nesso causale tra un dato antecedente e la sua conseguenza[172] (*black box* causale)[173] ma impedisce,

---

[161] G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., pp. 108-109.
[162] C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 122.
[163] M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 30.
[164] C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 122.
[165] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 74.
[166] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 7.
[167] R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 265.
[168] I. SALVADORI, *Agenti artificiali*, cit., p. 89.
[169] C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 122. L'A. riporta in tal sede l'esempio della vittoria di AlphaGo Master (un macchinario ideato per giocare al gioco di Go) contro il campione mondiale del medesimo gioco: i programmatori del sistema non sono riusciti a comprendere i diversi passaggi eseguiti dalla macchina e a capire come questa abbia potuto vincere. Altro esempio di opacità dei sistemi di *deep learning* viene riportato dall'A. a proposito dell'abbandono di un esperimento da parte di Facebook che consisteva nel far dialogare tra loro due programmi di IA servendosi di un linguaggio da loro stessi creato e incomprensibile per i loro stessi programmatori.
[170] A proposito degli studi volti ad "aprire" questa scatola nera e a renderla *explainable* v. G. MOBILIO, *L'intelligenza artificiale*, cit., p. 407.
[171] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 74.
[172] Se in ambito scientifico viene esposto a chiare lettere che «non si dovrebbero trarre conclusioni solo sulla base della correlazione tra due eventi: la correlazione non è causalità» A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 64, in ambito giuridico sappiamo che non è sempre così. Ci sono

altresì, di individuare il responsabile di un potenziale evento lesivo cagionato dall'IA che, in astratto, potrebbe essere imputabile al programmatore, al costruttore, all'utente finale o, financo, alla macchina stessa. Il tutto a discapito dell'eventuale persona offesa, la quale non sarebbe in grado di individuare il soggetto cui rivolgersi per chiedere un risarcimento del danno o contro cui agire in giudizio[174].

Il problema della mancata trasparenza dei sistemi di IA spesso dipende anche da un intricato sistema di brevetti e segreti industriali che coprono i software in questione[175] e che sono necessari per tutelare il copyright[176]. Inoltre l'opacità del sistema mette in risalto il c.d. "differenziale di conoscenze"[177] tra gli ideatori dei sistemi intelligenti e gli utenti che se ne servono, i quali non posseggono gli strumenti per aprire la scatola nera.

La mancanza di trasparenza di questi sistemi diventa tanto più pericolosa quando essi vengono utilizzati in contesti ove sono coinvolti diritti fondamentali della persona[178]. La trasparenza dovrebbe costituire un requisito fondamentale di tutti i sistemi di IA: «la decisione, umana, di lasciare che macchine o algoritmi producano effetti su altri esseri umani dovrebbe essere condizionata all'intelligibilità del modo in cui agiscono e alla possibilità di ascrivere tali effetti alla responsabilità di esseri umani. L'alternativa, di automatizzare alcuni processi decisionali, delegandoli a sistemi di intelligenza artificiale (…) che non includono le più elementari regole del buon senso umano, comporta la conseguenza di una perdita di prevedibilità e di controllo su quei processi, i cui costi e i cui benefici dovranno essere oggetto, di volta in volta, di una attenta valutazione comparativa»[179].

---

branche del diritto penale in cui l'accertamento del nesso causale risulta particolarmente complesso, specie quando viene in gioco il meccanismo della sussunzione sotto leggi scientifiche di tipo statistico. Basti pensare, a titolo meramente esemplificativo, al complesso accertamento del nesso causale tra l'insorgenza delle malattie professionali e la condotta del datore di lavoro che ha esposto i suoi dipendenti a una sostanza tossica o che ha omesso le dovute cautele per mettere in sicurezza la struttura (quindi anche alla difficoltà di comprendere se ci si trovi davanti ad un modello imputativo commissivo o omissivo), così cagionando l'insorgere della malattia. Per tutti Cassazione penale, Sez. I, 20.04.2006, n. 20370, in *DeJure*, Cassazione Porto Marghera. Analoghe problematiche sorgono in tema di responsabilità medica: per tutti, Cassazione penale, SS.UU., 10.07.2002, n. 30328, Sezioni Unite Franzese, in *DeJure*.

[173] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1760.

[174] C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 123.

[175] F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 14; U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 629; E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale,* cit., p. 80. Più in generale sul tema v. M. RICOLFI, *Il futuro della proprietà intellettuale nella società algoritmica*, in *Giur. it., Speciale 170 anni*, 2019, pp. 10 ss.

[176] C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 8; G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 104; I. SALVADORI, *Agenti artificiali*, cit., p. 89.

[177] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1772.

[178] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 21. Sulle conseguenze della *black box* in ambito processualpenalistico v. V. MANES, *L'oracolo algoritmico,* cit., p. 560.

[179] D. TAFANI, *Sulla moralità artificiale. Le decisioni delle macchine tra etica e diritto*, in *Rivista di filosofia*, 2020, p. 100.

Possediamo adesso gli strumenti necessari per comprendere un'ulteriore distinzione tra sistemi di intelligenza artificiale avanzata dalla dottrina, ossia quella tra sistemi simbolici e sub-simbolici[180].

La prima categoria, ossia l'IA *simbolica*, è tipica degli algoritmi tradizionali ed è caratterizzata dalla c.d. *explainability*. La seconda categoria (invero già accennata) è quella dell'IA *sub-simbolica*, tipica degli algoritmi di *machine learning* e connotata dalla c.d. *inexplicability*. Vorremmo in tal sede proporre un'ulteriore bipartizione da associare a quella appena proposta e che possa servire a connotare le diverse tipologie di IA e a creare due macro-aree del settore. Potremmo ricondurre la prima tipologia di IA ai sistemi dotati di *automazione*, ossia meccanismi classici che richiedano ancora, in misura più o meno incisiva, l'intervento umano, agendo in modo prestabilito. La seconda tipologia di IA, invece, potrebbe essere riferita ai sistemi dotati di *autonomia*, ossia i sistemi di cui ci siamo fino ad ora maggiormente occupati e che si caratterizzano per la possibilità di agire anche senza la supervisione umana, adattando il loro processo decisionale al fine di migliorare il raggiungimento degli obiettivi prefissati dai programmatori[181].

6. Il ruolo dei dati.

Altro elemento di fondamentale importanza per garantire il funzionamento dei sistemi intelligenti è costituito dai dati. Non dobbiamo infatti dimenticare – come forse si avrà già avuto modo di intuire dalle pagine precedenti – che i sistemi di *machine learning* dipendono dalla disponibilità dei dati. Essi apprendono dai dati che ricevono e migliorano progressivamente i loro risultati[182].

L'ascesa dell'apprendimento automatico è stata determinata e alimentata da un massiccio aumento della disponibilità di dati online[183], dati che siamo principalmente noi stessi a produrre mediante le interazioni tra di noi (dati "*people to people*") o con le macchine (dati "*people to machine*"). Esiste poi una terza tipologia di dati (dati "*machine to machine*") che vengono generati automaticamente dai nostri dispositivi, indipendentemente dall'intervento umano[184].

L'efficacia e la buona riuscita di un apprendimento automatico dipendono dalla disponibilità di grandi quantità di dati di alta qualità (i sistemi di *machine learning* spesso non funzionano adeguatamente in ambienti in cui ci sono pochi dati o dati inadeguati)[185]. La qualità e la completezza dei dati utilizzati per addestrare l'IA è condizione imprescindibile per garantire l'accuratezza del

---

[180] M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 27; R. ROVATTI, *Il processo di apprendimento algoritmico*, cit., pp. 32-33.

[181] Sulla distinzione tra agenti automatizzati e agenti autonomi v. I. SALVADORI, *Agenti artificiali*, cit., p. 92.

[182] L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., p. 4.

[183] H. SURDEN, *Artificial Intelligence and Law*, cit., pp. 1315-1316.

[184] G. ITALIANO, *Intelligenza Artificiale*, cit., pp. 220-221, l'A. si riferisce a tutti i sistemi del c.d. *Internet of Things* che oggi fa parte delle nostre vite, dai dispositivi di geolocalizzazione agli *smart devices*. Una ulteriore e diversa classificazione dei dati è stata fornita da L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., pp. 3 ss. il quale parla di dati storici, dati sintetici e dati ibridi. Sulla capacità di tali algoritmi di implementare automaticamente la propria banca dati mediante la codifica di nuove informazioni, così sviluppando un'elevata capacità di *problem solving*, v. M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 2.

[185] H. SURDEN, *Artificial Intelligence and Law*, cit., p. 1316.

risultato finale[186]. Nel contesto che stiamo cercando di delineare appare chiaro che, ad oggi, le aziende con una maggiore disponibilità di dati detengono un vantaggio competitivo sul mercato, contribuendo ad alimentare i progressi dell'IA[187]. Non è un caso che i dati siano stati efficacemente definiti come il "nuovo petrolio"[188], proprio a voler sottolineare il valore che acquistano nella società attuale[189].

Ciò di cui stiamo parlando, ossia l'avvento di un quantitativo massiccio di dati digitalizzati e delle nuove tecnologie, è stato definito dalla dottrina di settore con la parola "datificazione"[190] e l'enorme quantitativo di dati di cui essa abbisogna per funzionare e progredire è noto a tutti come Big Data[191]. Questo è il termine che riassume la rivoluzione tecnologica che stiamo cercando di delineare ma che, a tutt'oggi, non ha forse trovato una compiuta definizione. L'Unione Europea ha tentato di fornirne un'approssimativa definizione affermando che con «il termine "big data" (megadati) si riferisce a grandi quantità di tipi diversi di dati prodotti da varie fonti, fra cui persone, macchine e sensori»[192]. Qualcuno ha proposto di descriverli secondo la "Regola delle tre V": Volume (che evidentemente si riferisce al quantitativo di informazioni che caratterizza i big data), Velocità (riferita alla rapida acquisizione e analisi dei dati) e Varietà (per indicare la pluralità delle informazioni raccolte)[193]. Altri ancora hanno ritenuto che sarebbe sbagliato tentare di racchiudere i big data in un dato numerico,

---

[186] F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 6; M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 28.

[187] G. ITALIANO, *Intelligenza Artificiale*, cit., p. 221.

[188] Non a caso, il 6 maggio 2017 veniva pubblicato sull'*Economist* un articolo dal titolo *The world's most valuable resource is no longer oil, but data*.

[189] L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., p. 3; R. ROVATTI, *Il processo di apprendimento algoritmico*, cit., p. 33.

[190] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 52.

[191] M. DELMASTRO, A. NICITA, *Big Data. Come stanno cambiando il nostro mondo*, Bologna, 2019. L'influenza che i Big Data possono avere è limpidamente espressa dalla seguente espressione: «sono loro, i Big data, che pensano e amministrano coloro da cui hanno avuto origine» R. CALASSO, *L'innominabile attuale*, Milano, 2017, citato da A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 9. Riteniamo opportuno riportare qui l'opinione di parte della dottrina ad avviso della quale sarebbero i dati "piccoli" ma di elevata qualità a costituire il futuro dell'intelligenza artificiale insieme alla sua crescente capacità di generare i propri dati e alla capacità di progettazione, L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., pp. 4-7-13.

[192] *La riforma della protezione dei dati dell'UE e i big data*, Scheda informativa, gennaio 2016, p. 1. L'Unione Europea prosegue chiarendo che «Alcuni esempi sono i dati sul clima, le immagini satellitari, le immagini e i video digitali, le registrazioni di operazioni o i segnali GPS. I big data possono comprendere dati personali: ad es. informazioni riguardanti una persona, come un nome, una fotografia, un indirizzo e-mail, estremi bancari, messaggi postati sui siti delle reti sociali, informazioni cliniche o l'indirizzo IP di un computer».

[193] D. LANEY, *3D Data Management: Controlling Data Volume, Velocity, and Variety*, in *META Group*, 6.2.2001, pp. 1 ss. A questa impostazione di base sono seguite altre proposte volte ad aumentare il numero delle "V" fino a 5, aggiungendo «Variabilità: il contenuto dei dati muta di significato a seconda dell'analisi a cui è sottoposto; - Valore. Idoneità ad estrarre un valore/significato dai dati o dalla loro analisi», M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, *AI: profili tecnologici Automazione e Autonomia: dalla definizione alle possibili applicazioni dell'Intelligenza Artificiale*, in *BioLaw Journal*, 2019, p. 5.

individuando il loro reale valore nella "novità" concernente il fatto che ad oggi «il vero *big* dei dati è l'essere carburante per le macchine della conoscenza»[194].

È chiaro che questa enorme mole di dati sarebbe priva di valore se considerata isolatamente: al fine di valorizzarli si rende necessario l'uso di sofisticati meccanismi di intelligenza artificiale per poterli processare e analizzare[195]. I sistemi intelligenti consentono oggi di gestire in modo rapido ed efficace un quantitativo di dati non umanamente governabile.

Da queste prime indicazioni possiamo conseguentemente dedurre che, se ciò che conta è la qualità del dato, del pari avrà preminente rilevanza la provenienza di esso[196]. Ancora, se i dati forniti al sistema intelligente non sono affidabili, o sono addirittura malevoli, gli algoritmi che li hanno recepiti non potranno che produrre conclusioni inaffidabili[197].

Ad avviso della dottrina specialistica, la parte più complessa della programmazione dell'IA attiene proprio alla «creazione del modello, che viene costruito a partire dai dati. Questa è la fase di *training*: gli algoritmi di machine learning usano i dati per imparare, cercano di trovare segnali nascosti all'interno dell'enorme quantità di dati su cui lavorano», per questo è così importante che questi sistemi lavorino con dati di qualità. In tutti i settori in cui l'intelligenza artificiale e il *machine learning* hanno trovato applicazione, hanno «risolto in modo nuovo molti problemi difficili, che prima non si sapeva risolvere affatto. Ma hanno anche creato problemi completamente nuovi, e anche loro di non facile risoluzione»[198].

Uno dei principali problemi sollevati dall'alimentazione dell'IA mediante i dati è stato individuato nella potenziale lesione del *principio di non discriminazione*[199]. Questa problematica viene maggiormente in rilievo nel momento in cui ci si confronta con i c.d. algoritmi predittivi, così denominati in quanto, «sulla base di un'enorme quantità di dati e della loro elaborazione, consentono alla macchina di fornire una risposta probabilistica ai dubbi sull'accadimento di un evento incerto perché futuro»[200]. A tal proposito, quando un algoritmo predittivo è alimentato con dati discriminatori, si usa parlare del c.d. fenomeno di "*garbage in garbage out*"[201]: da un input di cattiva qualità non può che derivare un output di scarsa qualità. I dati reali, inseriti dall'uomo nella macchina, sono inevitabilmente viziati da pregiudizi che sono tipici dell'essere umano (a ciascuno i suoi) e che non vengono filtrati dalla macchina, tutt'altro, vengono anzi amplificati dai sistemi di *machine learning*. Gli "errori" (se così vogliamo definirli) presenti in questi dati sono spesso insidiosi e non facilmente riconoscibili. Il punto è che, in questo contesto (ma forse in qualsiasi altro contesto in cui vengono utilizzate forme di intelligenza artificiale), un output insoddisfacente non è necessariamente frutto del malfunzionamento della

[194] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 58.
[195] G. ITALIANO, *Intelligenza Artificiale*, cit., p. 221.
[196] L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., p. 5.
[197] G. ITALIANO, *Intelligenza Artificiale*, cit., pp. 223-224.
[198] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit. p. 4.
[199] L'IA, infatti, andrebbe alimentata con dati neutri, privi di qualsivoglia pregiudizio o discriminazione, così C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 99.
[200] G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 10.
[201] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 85.

macchina quanto, piuttosto, delle disfunzioni della nostra società e dei dati che la rappresentano[202].

Da queste poche battute riusciamo a comprendere come la neutralità dell'intelligenza artificiale sia, in realtà, un falso mito[203]: l'algoritmo ripete le impostazioni mentali e i preconcetti (in termini di discriminazione e disuguaglianza) di cui sono portatori coloro che lo programmano[204]. L'IA non può essere neutrale[205] perché il primo a non esserlo è l'uomo che la programma, creando il c.d. fenomeno di "bias in bias out"[206].

Altro problema concerne il fatto che gli algoritmi di ML, per loro natura, fondano i loro modelli su grandi quantità di dati che, però, sono inevitabilmente *dati del passato*[207]: appare intuitivo che imparare da dati troppo lontani nel tempo sarebbe controproducente[208]. L'uso dei c.d. proxy data, ossia dati riferibili a tempi passati – perciò dedotti indirettamente – rischia (al pari di un input discriminatorio), di fornire un output non accurato[209]. Servirsi di dati statici e non dinamici, nel lungo periodo, potrebbe riproporre i pregiudizi storici e discriminatori cui facevamo pocanzi cenno, specie ove questi algoritmi vengano usati in settori in cui possono incidere sui diritti della persona[210], financo in materia penale[211].

---

[202] M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 28.

[203] Sull'equivoco della neutralità della tecnologia v. G. MOBILIO, *L'intelligenza artificiale*, cit., p. 406. *Contra* S. CHIARLONI, *Riflessioni minime*, cit., p. 8 il quale afferma che «la tecnologia in sé è neutra: dipende dall'uso che se ne fa».

[204] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 407.

[205] «Non si può dire che le logiche algoritmiche o quelle di autoapprendimento siano di per sé più neutrali dei ragionamenti umani che pure stanno alla base della loro costruzione e implementazione» C. CASONATO, *Potenzialità e sfide dell'intelligenza artificiale*, in *BioLaw Journal*, 2019, pp. 179-180.

[206] C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1932.

[207] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit. p. 5. «Nel diritto, l'intelligenza artificiale non sarà mai capace di vivere nel presente. essa attribuirà sempre significato all'accadere storico secondo un meccanismo di reminiscenza» M. PAPA, *Future crimes*, cit., p. 12.

[208] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 68.

[209] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 407.

[210] Pensiamo agli algoritmi che vengono utilizzati da molte aziende al fine di assumere personale per una certa posizione lavorativa. Se forniamo all'algoritmo informazioni sui precedenti soggetti che hanno ricoperto quella carica i quali sono stati in larga maggioranza uomini, a meno di non riuscire (compito non semplice) ad eliminare ogni traccia di legame col sesso del dipendente, l'algoritmo sarà portato a preferire un soggetto di sesso maschile per quel compito, in discriminazione del genere femminile. A proposito di un uso dell'IA che possa andare a incidere sulla vita delle persone, da più parti viene proposto l'esempio dell'uso dei sistemi intelligenti per valutare la concessione o il diniego di mutui o prestiti alle persone interessate, cfr. M.B. MAGRO, *Robot*, cit., p. 1206, nota 45; G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit. p. 5.

[211] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit. p. 6. Per un chiaro esempio delle possibili ricadute in ambito penalistico delle conseguenze discriminatorie degli output forniti dagli algoritmi predittivi v. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Tale aspetto è stato richiamato anche al Cap. I, Sez. I, Par. 4.

Ulteriore e non certamente secondario problema è determinato dall'impatto che questo uso (e abuso) di dati può avere sulla tutela della *privacy*[212]. Nella moderna società in cui viviamo stiamo assistendo alla sostituzione dell'*habeas corpus* in favore dell'*habeas data*, proprio in ragione della proiezione di una parte di noi nei dati rinvenibili in rete[213]. Occorre preliminarmente tenere presente che le garanzie individuali andrebbero rispettate sin dalla fase della raccolta dei dati personali[214] che poi alimenteranno l'IA, il che ci rimanda alla più generale esigenza di rispettare la privacy a trecentosessanta gradi. Si usa parlare in tal senso di *privacy by design* e *privacy by default*, «ovverosia l'interiorizzazione preventiva dei valori di tutela nella creazione stessa degli algoritmi»[215].

Il principale testo normativo europeo[216] che oggi si occupa del trattamento dei dati personali è il *Regolamento Generale sulla Protezione dei Dati*, Reg. (UE) 2016/679, meglio noto come GDPR[217]. Quest'ultimo prevede espressamente all'art. 25 che la protezione dei dati personali debba avvenire sin dalla fase di progettazione dell'algoritmo (per l'appunto, *by design*) e per impostazione predefinita (*by default*)[218]. Tale Regolamento in realtà, al Considerando 6[219], si

[212] F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 10; G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit. p. 5; A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 74.

[213] Questo passaggio significativo viene richiamato da G. MOBILIO, *L'intelligenza artificiale*, cit., p. 405 il quale richiama sua volta S. RODOTÀ, *Dal soggetto alla persona*, Napoli, 2007 e S. PIETROPAOLI, *Habeas Data. I diritti umani alla prova dei big data*, in S. FARO, T. EDOARDO, G. PERUGINELLI, *Dati e algoritmi. Diritto e diritti nella società digitale*, Bologna, 2020, pp. 97 ss.

[214] G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 6.

[215] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 80.

[216] Non il solo, invero. Per una più approfondita panoramica a proposito della Dir. 27 aprile 2016 n. 2016/680/UE, attuata dall'ordinamento italiano con il d.lgs. 18 maggio 2018 n. 51, v. G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 7.

[217] Per un approfondimento sul tema v., in particolare, F. PIZZETTI, *La protezione dei dati personali e la sfida dell'Intelligenza Artificiale*, in F. PIZZETTI, *Intelligenza artificiale*, cit., pp. 5 ss. Sul tema anche A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 75.

[218] *Regolamento Generale sulla Protezione dei Dati*, cit., art. 25: «1. Tenendo conto dello stato dell'arte e dei costi di attuazione, nonché della natura, dell'ambito di applicazione, del contesto e delle finalità del trattamento, come anche dei rischi aventi probabilità e gravità diverse per i diritti e le libertà delle persone fisiche costituiti dal trattamento, sia al momento di determinare i mezzi del trattamento sia all'atto del trattamento *stesso il titolare del trattamento mette in atto misure tecniche e organizzative adeguate*, quali la pseudonimizzazione, *volte ad attuare in modo efficace i principi di protezione dei dati*, quali la minimizzazione, e a integrare nel trattamento le necessarie garanzie al fine di soddisfare i requisiti del presente regolamento e tutelare i diritti degli interessati. 2. *Il titolare del trattamento mette in atto misure tecniche e organizzative adeguate per garantire che siano trattati, per impostazione predefinita, solo i dati personali necessari per ogni specifica finalità del trattamento*. Tale obbligo vale per la quantità dei dati personali raccolti, la portata del trattamento, il periodo di conservazione e l'accessibilità. In particolare, dette misure garantiscono che, per impostazione predefinita, non siano resi accessibili dati personali a un numero indefinito di persone fisiche senza l'intervento della persona fisica», corsivo nostro.

[219] *Regolamento Generale sulla Protezione dei Dati*, cit., Considerando 6: «La rapidità dell'evoluzione tecnologica e la globalizzazione comportano nuove sfide per la protezione dei dati personali. La portata della condivisione e della raccolta di dati personali è aumentata in modo significativo. La tecnologia attuale consente tanto alle imprese private quanto alle autorità pubbliche di utilizzare dati personali, come mai in precedenza, nello svolgimento delle loro attività. Sempre più spesso, le persone fisiche rendono disponibili al pubblico su scala mondiale informazioni personali che le riguardano. La tecnologia ha trasformato l'economia e le relazioni sociali e dovrebbe facilitare ancora di più la libera circolazione dei dati personali all'interno

concentra sul ruolo delle nuove tecnologie nell'ambito del trattamento dei dati personali, chiarendo più avanti, segnatamente all'art. 35, che quando un certo tipo di trattamento preveda l'uso di nuove tecnologie, spetterà al titolare del trattamento effettuare in via preliminare una valutazione d'impatto del suddetto trattamento sulla protezione dei dati personali[220]. Il Regolamento prevede un complessivo meccanismo di responsabilizzazione del titolare del trattamento al quale spetterà porre in essere un aggiornamento e un riesame delle misure tecnico-organizzative adottate per tutelare il diritto alla privacy[221].

Vorremo in tal sede evidenziare come i classici metodi di tutela del diritto alla riservatezza paiono cedere di fronte all'avanzare delle nuove tecnologie, dato di cui il succitato Regolamento dovrebbe tenere conto. Ci stiamo riferendo al più classico degli strumenti di tutela del diritto alla privacy: il *consenso informato*. Tale categoria si sta dimostrando sempre più inadeguata a proteggere il titolare dei dati dalle nuove applicazioni dei sistemi informatici, genericamente intesi[222]. La libertà del consenso (dis)informato diventa oggi una mera *fictio*[223], quasi un consenso obbligato. Non prestare il consenso ogniqualvolta un servizio online lo richiede comporterebbe inevitabilmente essere tagliati fuori dal mondo: social media, servizi di mail, di acquisti online e qualsiasi altro servizio offerto in rete. Il consenso informato, da strumento di tutela della riservatezza della persona, diventa la porta d'accesso attraverso la quale ogni giorno consentiamo l'accesso ai nostri dati al titolare di qualunque attività svolta online[224]. Questo insieme di attività svolte in rete vengono comunemente definite come "tecnologie dell'informazione e della comunicazione" (anche note come ICT: *Information and Communication Technology*)[225].

Appare intuitivo quanto sia rischioso che questa enorme quantità di dati prodotta quotidianamente venga gestita da un ristretto numero di società[226]. Ci stiamo riferendo ai c.d. "baroni dei big data"[227], a quei "giganti dei servizi digitali"[228] che gestiscono la quasi totalità delle attività e dei servizi rinvenibili in

---

dell'Unione e il loro trasferimento verso paesi terzi e organizzazioni internazionali, garantendo al tempo stesso un elevato livello di protezione dei dati personali».

[220] *Regolamento Generale sulla Protezione dei Dati*, cit., art. 35: «Quando un tipo di trattamento, allorché prevede in particolare l'uso di nuove tecnologie, considerati la natura, l'oggetto, il contesto e le finalità del trattamento, può presentare un rischio elevato per i diritti e le libertà delle persone fisiche, il titolare del trattamento effettua, prima di procedere al trattamento, una valutazione dell'impatto dei trattamenti previsti sulla protezione dei dati personali».

[221] Per un commento critico, concernente il dato per cui con una simile previsione ci si espone a dover porre in essere aggiornamenti continui per cercare di stare al passo con l'evoluzione tecnologica v. G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 7.

[222] C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 107.

[223] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 80.

[224] C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 107.

[225] L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., p. 35; A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 66; C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 107.

[226] C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 109.

[227] S. CHIARLONI, *Riflessioni minime*, cit., p. 8.

[228] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 60. L'A. evidenzia, ad esempio, come una ricerca effettuata su Google lasci dietro di sé una scia di dati utili a identificare il soggetto che sta svolgendo la ricerca (es. ora, data, luogo, indiritto IP del dispositivo). Si tratta delle c.d. "briciole digitali" che ciascuno di noi produce navigando in internet e che, una volta aggregate, formano uno straordinario volume di dati socio-economici, p. 52.

rete. Stiamo parlando di Google, Amazon, Facebook, Apple, Microsoft, IBM e di tutte le loro appendici[229]. Le piattaforme digitali dei *Big Tech* «sono tutt'altro che neutrali, essendo prodotte da poche (e ricchissime) società commerciali che realizzano larghissimi profitti attraverso la raccolta, l'elaborazione e la vendita dei dati personali degli utenti che le utilizzano (…) Esse, dunque, sono in grado di influenzare (o addirittura determinare) le decisioni individuali e collettive attraverso varie forme di *nudging*, di "spinte gentili", meno visibili ma non per questo meno efficaci»[230]. Tali società gestiscono di fatto il mondo dell'informazione sulla base di opache logiche di profitto[231].

In questo contesto si colloca quella "filosofia iperliberalista" che propone di lasciare il mercato libero di governare questa nuova economia, fatta di tecnologia e dati[232], secondo un sistema di *self-regulation* delle aziende del settore le quali, evidentemente, preferirebbero una regolamentazione che avvenga esclusivamente mediante standard etici e non mediante norme giuridiche vincolanti che prevedano specifiche sanzioni in caso di una loro violazione[233]. Una simile soluzione non appare condivisibile.

Questo strapotere nelle mani di pochi potrebbe tramutarsi in un rischio per la riservatezza, la democrazia, la libertà di espressione e il pluralismo[234]. Ciò accade già con i sistemi di "pubblicità adattiva"[235] i quali segnalano alle persone offerte sempre più ritagliate sui propri gusti. Il vero problema è che ciò avviene anche con riferimento alle informazioni di natura politica, specie elettorale: «il rischio è che la persona venga esposta a notizie molto limitate e ritagliate su misura sulla base delle sue presunte preferenze»[236], così impoverendo il dibattito pubblico e acuendo il pericolo di dar vita a una sorta di "*microtargeting* degli elettori"[237]. Il destinatario di tali informazioni confiderà nella loro neutralità, esponendosi così al

[229] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 403, nota 16.

[230] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 70. L'A. dimostra che non si tratti di mera fantascienza ricordando il caso Cambridge Analytica recentemente occorso e che ha visto la suddetta società accusata di aver interferito, per conto della Russia, nell'elezione del presidente Donald Trump mediante la creazione di account di Facebook falsi. Sul tema v. anche S. CHIARLONI, *Riflessioni minime,* cit., p. 8.

[231] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 404, l'A. parla a tal proposito di "cattura del regolatore".

[232] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 403.

[233] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 410; A. CELOTTO, *I robot possono avere diritti?*, in *BioLaw Journal*, 2019, p. 99 parla di "ordine spontaneo" riscontrando che «in nome del tradizionale principio di base *ubi societas, ibi ius*, il diritto degli Stati è sempre molto lento e si adatta con difficoltà alle novità. Ancor più quando sono novità che crescono e si sviluppano secondo il principio di singolarità tecnologica, cioè con una accelerazione oltre ogni capacità di comprensione previsione umana».

[234] C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 110.

[235] C. SALAZAR, *Umano, troppo umano,* cit., p. 264. Sul punto v. anche G. ITALIANO, *Intelligenza Artificiale,* cit., p. 217; M.B. MAGRO, *Robot,* cit., p. 1193; A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 60 il quale parla di *consumer analytics* (ossia la scienza che studia il comportamento dei clienti online) e di *market basket* (ossia un paniere d'acquisto, per definire il profilo di ciascun cliente). A proposito del c.d. *opinion mining*, ossia la ricerca e l'esame delle opinioni espresse online, v. P. MORO, *Macchine come noi*, cit., p. 57.

[236] C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 110.

[237] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 22.

rischio di imbattersi in qualche *fake news*[238] e di alimentare il fenomeno della c.d. *post-verità*[239] e della c.d. *bubble democracy*[240], una sorta di confinamento degli elettori in sistemi social autoreferenziali che impediscono il confronto tipico del dibattito democratico[241].

Come correttamente osservato dalla dottrina, ciò che viene in gioco in un simile scenario non è più tanto la tutela della libera manifestazione del pensiero[242] quanto, piuttosto, la stessa "libertà di pensiero"[243]: «oggi la questione dirimente non è più garantire la circolazione delle idee, bensì la loro formazione, la loro genuina concezione»[244]. Ci troviamo nell'era della "disintermediazione dell'informazione"[245] operata mediante internet. In altre parole, le conoscenze vengono lasciate libere, favorendo l'eliminazione di qualsivoglia filtro o distinzione tra esperti e inesperti, facendo perdere autorità alla voce dei primi[246].

Quelle appena riportate sono solo alcune delle caratteristiche principali dei dati e delle possibili conseguenze derivabili da un loro uso distorto mediante le nuove tecnologie. Si è ritenuto opportuno in tal sede dar conto di questa e delle altre tipicità dei sistemi intelligenti in quanto convinti del fatto che approfondire il tema della responsabilità penale che può derivare dall'uso di questi sistemi non possa prescindere da una – seppur sommaria – conoscenza degli stessi.

7. Le differenze con la robotica.

Vorremmo brevemente sgombrare il campo da un comune fraintendimento: sistemi intelligenti e sistemi robotici non sono assimilabili e non sempre coincidono. L'assenza di una nozione univoca di "robot"[247] rende difficile individuare il rapporto intercorrente tra intelligenza artificiale e robotica[248]. Parte della dottrina ritiene che tra i due concetti non vi sia un rapporto di genere a specie[249] e che il campo della robotica non esaurisca il novero dei sistemi dotati di

---

[238] G. ITALIANO, *Intelligenza Artificiale*, cit., p. 224. Sul tema cfr. Report of the independent High level Group on fake news and online disinformation, *A multi-dimensional approach to disinformation*, marzo 2018.

[239] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 70, il quale richiama la definizione fornita dal Dizionario della lingua vivente di Oxford che, con tale termine, si riferisce alle «circostanze in cui i fatti oggettivi sono meno influenti nel plasmare l'opinione pubblica rispetto agli appelli alle emozioni e alle convinzioni personali», https://www.lexico.com/definition/post-truth.

[240] Sul tema delle "*filter bubbles*" e delle "*echo chambers*" v. M. FASAN, *Intelligenza artificiale e pluralismo: uso delle tecniche di profilazione nello spazio pubblico democratico*, in A. D'ALOIA, *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, Milano, 2020, pp. 345 ss.

[241] C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 111.

[242] Sui problemi della libertà di espressione online v. A. GULLO, *Nuove frontiere tecnologiche e sistema penale*, cit., p. VIII.

[243] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 23.

[244] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 71 il quale, a sua volta, cita M. AINIS, *Il regno dell'Uroboro: benvenuti nell'epoca della solitudine di massa*, Milano, 2018, pp. 11-12.

[245] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 70; sul tema v. anche A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 23.

[246] L. MEZZETTI, *Introduzione*, in U. RUFFOLO, *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, cit., p. 12.

[247] G. CAPILLI, *I criteri di interpretazione*, cit., p. 462.

[248] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 335.

[249] A. TURANO, *Robotica e roboetica*, cit., p. 128.

intelligenza artificiale[250]. In tal senso parte degli studiosi usa distinguere robot *cognitivi* e *deterministici*: i primi sono capaci di apprendere dalla loro esperienza e di regolare di conseguenza il loro comportamento; i secondi, invece, agiscono in modo prevedibile sulla base di schemi predeterminati[251]. Entrambi tali dispositivi, più o meno complessi, godono di una certa "autonomia operativa", ossia un'autonomia non illimitata, bensì circoscritta a un determinato compito[252].

L'esigenza di fornire una definizione comune di "robot", che sia flessibile e che non ostacoli l'innovazione tecnologica, è stata avvertita anche a livello europeo, ove si constata altresì che oggi i robot sono in grado di svolgere attività che, un tempo, erano ad esclusivo appannaggio dell'uomo[253]. Ferma restando la succitata differenziazione tra robot dotati o meno di intelligenza artificiale, ci sembra efficace la definizione sintetica fornita da parte della dottrina ad avviso della quale «il robot è oggi un sistema complesso che imita attività umane integrando i risultati ottenuti dall'intelligenza artificiale in vari ambiti del comportamento e del ragionamento»[254]. Esso richiama l'idea «di una macchina artificiale costruita dall'uomo per svolgere precise funzioni legate soprattutto al mondo del lavoro»[255].

Altro comune fraintendimento che vorremmo in tal sede sfatare concerne la coincidenza dei robot con macchine dotate di fattezze umane. Esistono robot sprovvisti di una forma antropomorfa[256] e robot, invece, muniti di sembianze

---

[250] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 369; A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 8.

[251] A. TURANO, *Robotica e roboetica*, cit., p. 130, nota 22. Altra parte della dottrina, nello stesso senso, distingue tra Robot e Robot intelligenti, M.B. MAGRO, *Robot*, cit., pp. 1190 ss. Altri ancora distinguono robot tele-operati (o autonomi) ossia in grado di svolgere, previo input iniziale, i loro compiti da soli, e robot cognitivi, qualora l'IA che li integra gli consenta di apprendere dall'esperienza e di regolare di conseguenza il loro comportamento, C. SALAZAR, *Umano, troppo umano*, cit., p. 260.

[252] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto: quale ruolo per il controllo umano?*, in *BioLaw Journal*, 2019, p. 36. Vediamo come le considerazioni svolte sul finire del Par. 5 in ordine all'autonomia relativa e non assoluta dei sistemi di IA si ripropongono negli stessi termini a proposito dei sistemi robotici.

[253] *Risoluzione del Parlamento europeo*, cit., Considerando C e Z. La Risoluzione afferma, al Considerando AA, il concetto di "autonomia di un robot" già menzionato nel Cap. I, Sez. I, Par. 2 e che qui riportiamo per completezza: «l'autonomia di un robot può essere definita come la capacità di prendere decisioni e metterle in atto nel mondo esterno, indipendentemente da un controllo o un'influenza esterna; che tale autonomia è di natura puramente tecnologica e il suo livello dipende dal grado di complessità con cui è stata progettata l'interazione di un robot con l'ambiente».

[254] P. MORO, *Biorobotica e diritti fondamentali*, cit., p. 533.

[255] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 24. Gli AA. si concentrano anche sull'origine del termine robot, da far risalire al romanzo "R.U.R." dello scrittore Karel Čapeck, il quale riprendere l'espressione dalla sua lingua madre e in particolare dalla parola "robota" che in ceco significa "lavoro forzato". Sulle origini del termine v. anche C. SALAZAR, *Umano, troppo umano*, cit., p. 258; A. TURANO, *Robotica e roboetica*, cit., p. 126; G. CAPILLI, *I criteri di interpretazione*, cit., p. 462, nota 15.

[256] Per una panoramica v. C. SALAZAR, *Umano, troppo umano*, cit., pp. 258 ss. la quale parla del robot "Belt Timeless Climbing", un piccolo veicolo costruito ricalcando la dinamica della scalata dei gechi sui piani verticali; i robot-serpenti, utilizzabili sia per la sorveglianza dei luoghi protetti dal segreto militare sia per operazioni di recupero dei superstiti dopo calamità naturali; il "K5", un robot cilindrico e semovente, munito di sensori, telecamere, di un programma di riconoscimento facciale, di infrarossi e di una strumentazione in grado di creare dettagliate mappe 3D dell'area in cui si trova; l'esempio probabilmente più noto è rappresentato da "Aibo", un cane-robot in grado

umane[257]. Parte della dottrina distingue in tal senso i robot umanoidi dagli androidi: i primi, per quanto dotati di alcune fattezze umane, sono inequivocabilmente ed immediatamente riconoscibili come macchine; i secondi, invece, costituiscono all'apparenza una perfetta replica di un essere umano[258]. Gli studiosi si sono soffermati sul punto, ritenendo che ai robot con fattezze umane potrebbero essere attribuite caratteristiche umane, creando una sorta di processo di "antropomorfizzazione"[259] che potrebbe pericolosamente condurre a relazionarsi con i robot nel medesimo modo in cui ci si relazionerebbe con un soggetto umano.

Su tale scia parte della dottrina ha rilevato come l'immissione sul mercato di macchine con fattezze perfettamente umane potrebbe incontrare qualche ostacolo non solo dal punto di vista giuridico[260] ma anche sociale[261]. Esistono già disposizioni nel nostro ordinamento volte a sanzionare la creazione di esseri non-umani che siano, però, dotati di caratteristiche tipiche dell'uomo. Ci stiamo riferendo all'art. 13 della legge n. 40 del 19 febbraio 2004, il quale vieta la sperimentazione sugli embrioni umani. La più attenta dottrina ha condivisibilmente osservato che «se la *ratio* dell'articolo 13 della legge n. 40/2004 è quella di impedire la creazione di esseri che abbiano allo stesso tempo qualità animali ed umane (come la forza di una tigre e l'intelletto dell'uomo), allora la norma potrebbe ben preludere ad un analogo divieto rispetto alla creazione di esseri dotati, allo stesso tempo, di qualità robotiche ed umane (come la durata e la potenza di una macchina e il cervello di un uomo), da imporsi una volta che la robotica sarà in grado di realizzare una tale creazione»[262]. Tale circostanza appare tanto più di interesse ai nostri fini in quanto la violazione del succitato divieto è sanzionata mediante il diritto penale, il quale entra in gioco al fine di «prevenire (e reprimere) la creazione di esseri (non-umani) che condividono caratteristiche umane»[263].

L'attualità e la problematicità del tema hanno portato alla nascita della c.d. roboetica[264]. Un primissimo approccio in tal senso, ad oggi valevole come base su cui edificare un solido e definitivo impianto etico della robotica, può essere

---

di imparare i comandi e che possa fungere sia da animale di compagnia sia da giocattolo intelligente per i più piccoli, prodotto dalla Sony e ritirato dal mercato dopo qualche anno in ragione dell'esiguo numero di esemplari venduti. Un cenno ad Aibo è rinvenibile anche in M.B. MAGRO, *Biorobotica*, cit., p. 509; G. CAPILLI, *I criteri di interpretazione*, cit., p. 475.

[257] Si ricordano in tal sede i robot umanoidi "Asimo" (Advanced Step in Innovative MObility), prodotto dalla Honda; "Atlas", in grado di spostarsi su terreni accidentati e di resistere agli urti ritrovando autonomamente l'equilibrio; "I-Cub", progettato a Genova, in grado di riconoscere gli oggetti di uso domestico che riesce a tenere nelle sue mani robotiche, di trasportarli e di "comprendere" le indicazioni che gli vengono date comportandosi di conseguenza, C. SALAZAR, *Umano, troppo umano,* cit., pp. 258 ss.; "TEO", creato da un team dell'Università Carlos III di Madrid, L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., p. 12.

[258] C. SALAZAR, *Umano, troppo umano,* cit., p. 261.

[259] M.B. MAGRO, *Robot*, cit., p. 1198.

[260] C. SALAZAR, *Umano, troppo umano,* cit., p. 262, a proposito delle implicazioni evidenziate dai membri del progetto RoboLaw, sul quale v. nota 460 del presente capitolo.

[261] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 54, a proposito delle implicazioni di un robot antropomorfo su soggetti che potrebbero da esso sviluppare forme di dipendenze e dell'impatto estetico di tali macchine a livello emotivo.

[262] S. RIONDATO, *Robotica e diritto penale*, cit., p. 605.

[263] S. RIONDATO, *Robotica e diritto penale*, cit., p. 606.

[264] La paternità del termine viene fatta risalire allo studioso Gianmarco Veruggio. Per una completa bibliografia di quest'ultimo cfr. A. TURANO, *Robotica e roboetica*, cit., pp. 132-133.

rappresentato dalle leggi di Asimov[265]. Il Parlamento europeo, nel corpo della Risoluzione recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica, ricorda che esse sono indirizzate a coloro i quali progettano e usufruiscono dei suddetti robot, ritenendo che tali leggi non siano suscettibili di essere convertite in codici con cui istruire le macchine[266]. Tali regole appaiono ispirate alla logica del *neminem laedere*[267] o, più in generale, alla regola del *primum non nocere*, in quanto essere ruotano intorno all'idea secondo cui le macchine non debbano recar danno all'uomo e all'intera umanità, ponendo le basi per la creazione di un codice etico per disciplinare la c.d. *human-robot interaction*[268].

Il rispetto di tali regole potrebbe servire a prevenire possibili usi illeciti dei sistemi robotici, nonché a porre specifici limiti (che potremmo definire "etici") volti a precludere – specie per quei robot dotati di intelligenza artificiale e, pertanto, di capacità di autoapprendimento – ogni evoluzione che possa sfociare in comportamenti lesivi di terzi[269]. Parte della dottrina ha però evidenziato la contraddittorietà delle leggi di Asimov. Esse, infatti, non sarebbero in grado di risolvere dilemmi etici e bilanciamenti di interessi tipici, ad esempio, di una situazione di legittima difesa: «cosa accade se si ordini ad un *robot* di ferire un uomo per il bene di qualcun altro?»[270].

Nonostante la suggestione suscitata dal richiamo alle leggi di Asimov, occorre dare atto del fatto che, invero, gli studi sul tema hanno iniziato a svilupparsi intorno agli anni 2000: uno dei primi documenti[271] ufficiali sul tema

---

[265] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 46 definisce le leggi di Asimov come «l'antesignano di ogni codice etico valido per i robot», specificando comunque come esse possano, al più, costituire un «riferimento culturale». A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 6, l'A. chiarisce in tal sede che, a Suo avviso, il Parlamento europeo, richiamandole nella Risoluzione recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica, ha inteso porre una base etico-deontologica che possa guidare i ricercatori del settore. Abbiamo già elencato le Leggi di Asimov al Capitolo I, Par. 2 insieme ad alcune considerazioni critiche in nota sul loro uso in ambito giuridico. Le riportiamo qui di seguito per completezza: «1) Un robot non può recar danno a un essere umano né può permettere che, a causa del proprio mancato intervento, un essere umano riceva danno; 2) Un robot deve obbedire agli ordini impartiti dagli esseri umani, purché tali ordini non contravvengano alla Prima Legge; 3) Un robot deve proteggere la propria esistenza, purché questa autodifesa non contrasti con la Prima o con la Seconda Legge; 0) Un robot non può recare danno all'umanità, né può permettere che, a causa del proprio mancato intervento, l'umanità riceva danno».

[266] *Risoluzione del Parlamento europeo*, cit., Considerando T, concetto richiamato anche in P. MORO, *Macchine come noi*, cit., p. 47.

[267] P. MORO, *Libertà del robot?*, cit., p. 541; M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 340.

[268] A. TURANO, *Robotica e roboetica*, cit., p. 128.

[269] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 23.

[270] Per questo e molti altri interrogativi del genere v. M.B. MAGRO, *Robot*, cit., p. 1199; EAD., *Biorobotica*, cit., p. 515. Sull'opportunità di rivisitare tali leggi G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 111. Propone la questione anche U. PAGALLO, *Saggio sui robot e il diritto penale*, in S. VINCIGUERRA, F. DASSANO (a cura di), *Scritti in memoria di Giuliano Marini*, Napoli, 2010, p. 599. Sulla contraddittorietà delle leggi di Asimov v. anche G. HALLEVY, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, cit., p. 173.

[271] Per un'analisi approfondita degli ulteriori testi che affrontano il tema della Roboetica v. A. TURANO, *Robotica e roboetica*, cit., pp. 133 ss.

risale infatti al 2006 ed è individuabile nella *Roboethics Roadmap*[272], volta ad incoraggiare le riflessioni sul tema. Ad oggi la roboetica può essere definita come quel «settore di studi che si occupa degli aspetti etici e sociali delle tecnologie robotiche nella loro interazione con l'uomo e con la società nel suo insieme (…) Oggetto specifico della disciplina non è il robot e la sua etica artificiale, bensì l'etica umana dei programmatori, ingegneri e utilizzatori dei robot»[273]. Invero altra parte della dottrina differenzia l'etica robot-umani (concernete il modo in cui i robot si comportano nei nostri riguardi), l'etica umani-robot (riferita al nostro comportamento nei confronti dei robot) e l'etica robot-robot (in ordine a come i robot si comportano fra loro)[274].

La questione concernente le modalità del comportamento umano in un contesto di coesistenza con agenti robotici e la regolamentazione del loro uso in modo da garantire il rispetto dei diritti e delle libertà fondamentali[275], ci conduce direttamente davanti alla possibilità concreta che i robot possano trovarsi difronte a veri e propri dilemmi morali. Basti pensare, a titolo esemplificativo, ai robot utilizzati per l'assistenza ad un soggetto anziano e al rifiuto da parte di quest'ultimo di assumere un farmaco: potrebbe il robot obbligarlo? Si propone in tal senso di implementare i metodi di apprendimento di questi sistemi in modo che facciano proprie regole etiche in grado di aiutarli a distinguere un comportamento giusto da uno sbagliato[276], per quanto tale prospettiva sembri attualmente ancora irrealistica[277].

Ci sentiamo in tal sede di poter affermare che, forse, tale tentativo di implementazione dovrebbe camminare a fianco di un ulteriore sforzo da parte dei programmatori di questi sistemi, ossia quello di rendere il loro comportamento non solo conforme all'etica ma anche conforme al diritto e alle sue regole.

8. Per un'IA etica e conforme al diritto.

Il campo dell'etica, menzionato in chiusura del paragrafo precedente, è in realtà stato oggetto di interesse anche nella più generale materia dell'intelligenza artificiale. Che quest'ultima debba essere etica appare ormai dato acquisito, specialmente a livello europeo[278]. Per (quantomeno provare a) rendere etico il comportamento di un sistema intelligente occorre programmarlo di conseguenza,

---

[272] G. VERUGGIO, *Euron Roboethics Roadmap*, 2007, consultabile al sito http://www.roboethics.org/index_file/Roboethics%20Roadmap%20Rel.1.2.pdf.

[273] A. TURANO, *Robotica e roboetica*, cit., pp. 132-133, il quale richiama a sua volta una completa bibliografia in tema di Roboetica. Sul tema v. anche P. MORO, *Libertà del robot?*, cit., p. 534.

[274] M.B. MAGRO, *Robot*, cit., pp. 1197 ss.

[275] M.B. MAGRO, *Robot*, cit., p. 1197.

[276] P. MORO, *Macchine come noi*, cit., p. 53.

[277] P. MORO, *Libertà del robot?*, cit., p. 353-541, in tal sede l'A. evidenzia che è stato riscontrato che tale questione supererebbe i confini di una qualsivoglia controllabilità dei sistemi di intelligenza artificiale. Se mai un giorno sarà possibile insegnare a questi sistemi come distinguere il giusto dall'ingiusto, a quel punto si potrà davvero parlare di "libertà delle macchine".

[278] Lo si è a lungo evidenziato nel Capitolo I quando ci si è occupati dell'analisi dei testi europei di riferimento. Basti pensare, a titolo esemplificativo, al Par. 4 sulla Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi e al Par. 7 sugli Orientamenti etici per un'IA affidabile. A proposito delle linee guida etiche redatte dalla Società giapponese per l'intelligenza artificiale (JSAI) v. C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 3.

secondo un approccio di *ethics by design*[279]. È stato proposto in dottrina (sulla falsariga di quanto già avviene con i succitati sistemi di *reinforced learning*) di adottare un approccio di bonus/malus etico, volto dunque ad orientare il comportamento dell'IA in modo che esso possa valutare il proprio agire e discernere tra una condotta corretta e una scorretta[280]. In poche parole, ci troviamo di fronte una nuova e diversa sfida: allineare gli algoritmi ai valori europei e ai diritti umani[281].

A dimostrazione dell'interesse che questa materia ha suscitato, l'IEEE (*Institute of Electrical and Electronics Engineers*), una delle più grandi organizzazioni professionali per il progresso della tecnologia, ha promosso l'*IEEE Global Initiative on Ethics of Autonomous and Intelligent System*, coinvolgendo centinaia di esperti del settore al fine di trovare in sinergia posizioni comuni in ambito etico. L'efficace sintesi del loro lavoro è oggi rinvenibile nell'*Ethical Aligned Design* (EAD)[282], documento che si pone l'obiettivo di fornire raccomandazioni che possano costituire un riferimento chiave per il lavoro dei tecnici del settore, individuando cinque principi cardine: il rispetto dei diritti umani, la priorità del benessere, la responsabilità, la trasparenza e la consapevolezza di un possibile uso improprio di tali tecnologie[283].

In un certo senso tutto questo non dovrebbe sorprendere: l'avvento delle nuove tecnologie tende ad essere accompagnato dalla nascita di nuove problematiche e di nuovi ambiti di riflessione[284], la *tecnoetica*[285] è uno di questi. Ci troviamo, a bene vedere, nel campo della c.d. *machine ethics* il quale, in estrema sintesi, studia come far sì che i sistemi autonomi intelligenti si comportino in modo eticamente responsabile[286]. La rilevanza degli aspetti etici degli algoritmi si fa tanto più evidente quanto più tali sistemi vengono caricati di responsabilità, ad esempio quando viene loro richiesto di effettuare transazioni finanziarie, di guidare un'auto o di fornire supporto ad una decisione importante, magari in sede giudiziaria. In tutti questi ambiti assume particolare importanza avere la possibilità di spiegare agli utenti perché il sistema ha assunto una data decisione e se quest'ultima possa essere considerata a tutti gli effetti "etica". Abbiamo avuto modo di vedere quanto questi sistemi, specie se dotati di *machine learning*, siano impenetrabili a causa del c.d. *black box effect*. Ciò rende estremamente arduo non solo ripercorrere il processo logico seguito dall'IA, ma anche individuare il responsabile della decisione assunta dal sistema intelligente. Che questi problemi suaccennati tornino attuali in questa sede, ove stiamo

[279] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 416.
[280] M.B. MAGRO, *Robot*, cit., p. 1200.
[281] L. MEZZETTI, *Introduzione*, cit., p. 7.
[282] IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, Version 2. IEEE, 2017, consultabile al sito https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf. È possibile consultare la prima edizione del testo al seguente link https://ethicsinaction.ieee.org/wp-content/uploads/ead1e.pdf.
[283] IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, *Ethically Aligned Design*, cit., pp. 22 ss. Per un approfondimento v. R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., pp. 48 ss.
[284] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 408.
[285] E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale,* cit., p. 76.
[286] H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law,* cit., p. 343.

cercando di fornire qualche spunto a proposito dell'etica, dimostra da un lato che non si tratta di problemi esclusivamente di tipo tecnologico e, dall'altro, che essi non possono essere risolti soltanto dai tecnici[287].

Si pone, a questo punto, un altro problema: chi "educa" le macchine? Torna nuovamente in rilievo l'illusione di una tecnologia totalmente neutra. Le diverse programmazioni, a ben vedere, dipenderanno dalle diverse logiche e dai diversi principi che ispirano colui il quale programma la macchina, proprio in quanto un'etica universalmente condivisa non sembra poter esistere. Anzi, ci si chiede come, «in un tempo caratterizzato da un esasperato pluralismo etico»[288], sia possibile impartire alla macchina un criterio morale di comportamento, creare algoritmi che rispecchino la dimensione etica desiderata oppure, ancora, spiegare al sistema come risolvere un conflitto di interessi giuridici. Sembra fuor di dubbio che i programmatori di questi sistemi intelligenti debbano possedere anche una formazione di tipo etico, il che potrà avvenire solo abbracciando la cultura della cooperazione tra branche del sapere[289].

Il tema dell'etica diverrà ancor più rilevante col progredire degli studi sul c.d. *affective computing*: «tentativo di programmare e realizzare macchine emotive, basate su software in grado di riconoscere ed esprimere emozioni»[290]. Tali sistemi, tramite il *deep learning*, sono in grado di elaborare le espressioni del viso ma, per quanto essi all'apparenza sembrino mostrare una inedita componente emotiva, dobbiamo ricordare che si tratta pur sempre di una simulazione[291], mancando in capo alle macchine il c.d. "senso dell'intero"[292].

Come abbiamo accennato alla fine del paragrafo precedente, non basta chiedersi come rendere l'IA conforme all'etica (anche se forse, a seguito di quanto abbiamo detto, sarebbe meglio parlare di "etiche" al plurale), occorre piuttosto chiedersi come far sì che l'IA si conformi alle regole del diritto.

Queste ultime devono restare saldo punto di riferimento, è proprio per questo che non sembra possibile condividere la preferenza accordata agli standard etici da parte dei *Big Tech* (cui abbiamo su accennato) rispetto ad un assetto di norme giuridiche. L'idea di una *self-regulation* in cui i controllati diventano i controllori farebbe venir meno il primato del diritto. Il rapporto tra etica e diritto non va, infatti, inteso come un rapporto conflittuale bensì come un lavoro sinergico, cumulativo e non alternativo[293]. Posta l'irrinunciabilità al ruolo del diritto[294], sarà opportuno per i programmatori dei sistemi di IA affrontare il c.d. rapporto tra "*code and law*"[295], o forse sarebbe più opportuno parlare di "*code as law*"[296]: in

[287] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit. p. 5.
[288] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 11.
[289] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., pp. 13-17.
[290] P. MORO, *Macchine come noi*, cit., p. 57. A proposito della c.d. "computazione emotiva" v. anche M.B. MAGRO, *Robot*, cit., p. 1193.
[291] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 26; M.B. MAGRO, *Biorobotica*, cit., p. 511.
[292] P. MORO, *Macchine come noi*, cit., p. 58.
[293] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 410.
[294] Sulla necessità di operare, mediante il diritto, un controllo su chi detiene i Big Data C. CASONATO, *Potenzialità e sfide dell'intelligenza artificiale*, cit., p. 178; sulla "decentralizzazione dei dati" A. CELOTTO, *I robot possono avere diritti?*, cit., p. 98.
[295] Per un approfondimento sul tema v. G. MOBILIO, *L'intelligenza artificiale*, cit., pp. 416 ss.
[296] R. LEENES, F. LUCIVERO, *Laws on Robots, Laws by Robots, Laws in Robots: Regulating Robot Behaviour by Design*, in *Law, Innovation and Technology*, 28.2.2014, p. 198.

altre parole, capire come far sì che i sistemi autonomi intelligenti si comportino in modo conforme alla legge[297].

Nell'affrontare tale compito ci si troverà davanti sfide tradizionali, come classificare giuridicamente un comportamento, gestire un conflitto tra norme, decifrare l'ambigua formulazione di alcune norme giuridiche, comprendere concetti giuridici (a volte) troppo vaghi o fronteggiare eccezioni imprevedibili da gestire sulla scorta dei principi generali[298]. V'è però anche da considerare che un comportamento socialmente condivisibile non dipende soltanto dall'osservanza delle norme di legge, potendo prevalere su queste ultime diversi fattori come le convenzioni sociali o il buon senso[299]. Dunque, al non semplice compito di tradurre la flessibilità tipica della legge in codici di comportamento da inserire negli algoritmi, si aggiunge l'ulteriore problema (probabilmente insormontabile) di far comprendere al sistema intelligente quando sarebbe più corretto non osservare quella legge (basti pensare a una manovra stradale d'emergenza che violi il codice della strada ma che sia funzionale a salvare una vita umana).

Da questa sintetica disamina possiamo dunque notare come non servano solo norme etiche e giuridiche che regolino i possibili conflitti generati dall'interazione uomo macchina[300], ma anche norme, tanto etiche quanto giuridiche, da cui i sistemi intelligenti possano imparare e che possano, conseguentemente, osservare e rispettare[301]. Occorre però ricordare sempre che, i reali destinatari di queste norme non saranno i sistemi di IA quanto, piuttosto, i programmatori e i produttori di queste ultime[302].

9. Considerazioni conclusive.

Nelle pagine precedenti abbiamo cercato di tracciare una cornice entro cui poter sviluppare le nostre riflessioni in punto di giuridicità penale. Possiamo dunque adesso cercare di trarre le preliminari fila conclusive di quanto esposto finora.

Abbiamo imparato a conoscere l'intelligenza artificiale come una materia interdisciplinare[303], la quale richiede il lavorio sinergico di competenze molto diverse fra loro al fine di affrontare le sfide che l'IA porta (e porterà) con sé e che abbiamo cercato di accennare nelle pagine precedenti[304]. Se da un lato appare di

---

[297] H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law,* cit., p. 343.

[298] H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law,* cit., pp. 343-344.

[299] H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law,* cit., p. 345.

[300] P. MORO, *Macchine come noi*, cit., p. 46; P. MORO, *Libertà del robot?*, cit., p. 527.

[301] G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 111.

[302] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 98.

[303] C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 103; P. MORO, *Libertà del robot?*, cit., p. 540; M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 29; R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 34; A. CELOTTO, *I robot possono avere diritti?*, cit., p. 99.

[304] «Tutti questi problemi, come discriminazioni, privacy, responsabilità degli algoritmi, sono nuovi e molto complicati. Problemi che non sono soltanto tecnologici, che non sono soltanto economici, che non sono soltanto tipici delle discipline sociali, e che quindi non possono essere affrontati con approcci tradizionali di queste discipline. Richiedono sempre più una stretta collaborazione e contaminazione tra esperti provenienti da discipline completamente diverse, che devono confrontarsi e lavorare insieme, con molto impegno e apertura mentale. (…) In questo

tutta evidenza che il giurista, da solo, non possegga gli strumenti e le conoscenze tecniche per fronteggiare il fenomeno[305], di contro risulta altrettanto evidente che il tecnico non possa essere del tutto conscio delle implicazioni giuridiche di ciò che progetta. Servirebbe, a ben vedere, una figura "ibrida" in grado di sapersi destreggiare in entrambe queste branche del sapere[306]: tra «le *hard sciences* (come l'elettronica, la genetica o la chimica), ma anche le *humanities* (come il diritto o la letteratura)»[307].

L'intelligenza artificiale si fonda, altresì, sulla simulazione del pensiero e dell'agire umano, grazie ai sofisticati sistemi di *machine learning*. Nel campo dell'intelligenza artificiale tali tecniche di apprendimento automatico sono in una fase di rapida crescita[308] tanto da costituire attualmente l'approccio più significativo all'IA. Il ML oggi «è alla base della maggior parte dei principali sistemi di intelligenza artificiale che incidono sulla società odierna, inclusi i veicoli autonomi, l'analisi predittiva, il rilevamento delle frodi e gran parte dell'automazione in medicina»[309]. Si tratta, come abbiamo cercato di evidenziare, di sistemi dotati di un certo grado di autonomia, il che solleva nuove e peculiari problematiche: «quando vengono utilizzati tali sistemi autonomi, le norme giuridiche non possono più essere considerate come regole del comportamento umano, poiché non sono gli esseri umani ad agire bensì le macchine. Ciò solleva il problema di come i sistemi autonomi possano essere progettati in modo tale che il loro comportamento sia conforme alla legge»[310].

Le peculiarità dell'IA, che abbiamo cercato di evidenziare nei passaggi precedenti, impongono non soltanto un adeguato meccanismo di test del relativo funzionamento[311] – proprio al fine di prevenire i rischi in essa insiti[312] – ma anche un sistema di revisione, per controllare e fronteggiare la velocità dello sviluppo di

---

scenario, è molto importante lavorare tutti insieme sulle nuove sfide che gli algoritmi stanno creando, soprattutto sui loro aspetti etici, di responsabilità, di discriminazione, di trasparenza, di equità, di organizzazione del lavoro e di governo nella nostra società. Per fare questo sono necessarie competenze fortemente interdisciplinari, che sappiano dialogare e lavorare insieme, in modo aperto, a 360 gradi, su tecnologie digitali, scienze sociali, economia, diritto, scienze politiche, cultura e società» G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., pp. 5-8.

[305] A proposito della "non autosufficienza" del diritto M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 370.

[306] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 419.

[307] P. MORO, *Biorobotica e diritti fondamentali*, cit., p. 537.

[308] H. VARSHNEY, R.A. KHAN, U. KHAN, R. VERMA, *Approaches of Artificial Intelligence*, cit., p. 2.

[309] H. SURDEN, *Artificial Intelligence and Law*, cit., p. 1315.

[310] H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law*, cit., pp. 342-343.

[311] Ciò è invero specificato anche al Punto 23 della *Risoluzione del Parlamento europeo*, cit., ove si legge che «testare i robot in condizioni reali è essenziale per individuare e valutare i rischi che potrebbero comportare, nonché il loro sviluppo tecnologico successivo alla fase puramente sperimentale di laboratorio; sottolinea, a tale proposito, che testare i robot in situazioni reali, in particolare nelle città e sulle strade, solleva un gran numero di problemi, tra cui ostacoli che rallentano lo sviluppo di queste fasi di collaudo, e richiede una strategia e un meccanismo di monitoraggio efficaci; invita la Commissione a elaborare criteri uniformi in tutta l'Unione, che i singoli Stati membri dovrebbero utilizzare per identificare le aree in cui autorizzare gli esperimenti con robot, nel rispetto del principio di precauzione».

[312] I. SALVADORI, *Agenti artificiali*, cit., p. 88.

questi sistemi[313]. Manifestano la loro utilità in tal senso le c.d. *regulatory sandboxes*, ossia «ambienti delimitati e protetti entro cui le imprese possono sperimentare per un determinato periodo di tempo la diffusione di prodotti innovativi da porre sul mercato, senza però essere obbligate a soddisfare ogni previsione normativa, ma sottoponendosi volontariamente al monitoraggio del regolatore. In questo modo le imprese possono verificare in un ambiente reale l'impatto di un prodotto, mentre il regolatore può meglio stabilire quali sono gli interessi da tutelare e abbattere le barriere e i costi della regolazione»[314].

In tal senso pare pionieristica l'esperienza giapponese, ove sono stati creati degli appositi spazi giuridicamente deregolamentati ove poter svolgere in libertà diversi esperimenti con un certo margine di tolleranza. Tale innovativo laboratorio è stato soprannominato *Tokku* e, invero, ne sono stati dislocati diversi in tutto il Paese: in questi spazi viene consentito sia agli scienziati che ai comuni cittadini di operare in sinergia sperimentando una preliminare coesistenza tra uomini e macchine, anche al fine di prevenire e risolvere le potenziali controversie giuridiche che potrebbero sorgere[315]. Tale meccanismo porta con sé diversi vantaggi, tra cui migliorare la comprensione di questi sistemi e di come possano reagire in diversi contesti, individuare e prevenire i rischi che la possibile perdita di controllo su di essi può generare nonché i relativi comportamenti indesiderati, definire lo spazio operativo di futuri sistemi intelligenti nonché, da ultimo, affrontare razionalmente gli aspetti burocratici di questo tipo di sperimentazione per evitare che essi costituiscano un ostacolo allo sviluppo dell'IA[316]. Ciò potrebbe anche incentivare la creazione di uno "spazio giuridico europeo" che, partendo da una sperimentazione a livello locale, sfrutti un sistema di cooperazione rafforzata[317] per implementare quelli che risultano essere, a tutti gli effetti, luoghi dove si esercita il futuro.

Ad ogni modo, uno dei problemi che sembra emergere da questa preliminare disamina è una sorta di «*personificazione* dell'intelligenza artificiale»: si parla spesso dell'IA in prima persona, come autore di condotte e comportamenti anche giuridicamente rilevanti. Ciò rischia di fare di essa un'espressione abusata al punto da svuotarla di significato[318]. Assistiamo, in altri termini, a una «*proiezione linguistica* dei caratteri umani»[319] con riferimento all'intelligenza artificiale. Per parlare di essa ci serviamo di un "vocabolario antropomorfo", forse non tanto per riferirci alla macchina in sé quanto, piuttosto, al «modo in cui noi vediamo la macchina, e indirettamente noi stessi»[320]. Tale abitudine genera a sua volta delle "confusioni antropomorfe": servirsi indifferentemente della medesima terminologia per parlare degli uomini e delle macchine induce ad equipararli, a credere che le macchine siano in grado di pensare come noi e di avere il nostro

[313] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 27.
[314] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 420.
[315] U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 631; I. SALVADORI, *Agenti artificiali*, cit., p. 117, nota 104; E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale,* cit., p. 89.
[316] U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 631.
[317] E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale,* cit., p. 92.
[318] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 3, corsivo nostro.
[319] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1766, corsivo nostro.
[320] M.B. MAGRO, *Robot,* cit., p. 1198.

stesso grado di consapevolezza[321], insomma, che non vi sia tra di essi alcuna differenza significativa[322]. In tal modo si rischia di legittimare l'applicazione ai sistemi dotati di intelligenza artificiale di schemi concettuali impropri, in quanto di esclusiva derivazione umana.

Ad esempio, dire che i sistemi di IA "prendono decisioni" è fuorviante: esse non prendono decisioni bensì generano output. Anche questa è una forma di antropomorfismo che connota la concezione sociale che l'IA sta pian piano assumendo nell'odierna società[323]. Si potrebbe al più parlare di un'opzione effettuata tra una limitata serie di scelte, piuttosto che di una "decisione" vera e propria[324]. Si dovrà col tempo procedere ad una "*alfabetizzazione* della società" con riguardo ai profili fondamentali che connotano l'IA, ad una sorta di "*drenaggio culturale*" da effettuare mediante un vero e proprio processo di educazione continua (un "*long life learning*") al fine di porre l'uomo al timone e non al traino di tali nuove tecnologie[325].

Un altro problema afferisce alla *regolamentazione giuridica* della materia, la quale al momento risulta limitata e frammentata[326]. Si registra una prevalenza di normative settoriali[327] che, ad avviso di parte della dottrina, non sono in grado di fronteggiare le nuove sfide dell'IA[328]. Altri evidenziano come, probabilmente, non sia del tutto possibile abbandonare la via della "pluralità di regolamentazioni"[329], dovendo prendere atto che la rapidità di sviluppo di questi sistemi potrebbe imporre un ripensamento dell'efficacia delle classiche e statiche fonti del diritto[330]. Ciononostante, si avverte nel settore la necessità di agire a livello condiviso[331], di realizzare un quadro normativo solido, omogeneo, semplice, unitario e flessibile[332]. In questo contesto, l'Unione Europea rappresenta terreno fertile ove poter elaborare standard legislativi comuni e colmare gli esistenti vuoti normativi con nuove cornici regolamentari[333]. Come abbiamo cercato di evidenziare nel Capitolo I, l'UE si sta già muovendo in questa direzione, a cominciare dalla Risoluzione del Parlamento europeo concernente norme di diritto civile sulla robotica[334].

La dottrina che si è maggiormente concentrata sul tema della regolamentazione giuridica dell'intelligenza artificiale ritiene necessario un nuovo approccio alla materia, riscontrando come un ricorso esclusivo agli strumenti di

---

[321] A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., p. 65.
[322] D. TAFANI, *Sulla moralità artificiale,* cit., p. 99.
[323] C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?,* cit., p. 1923, nota 50.
[324] M.B. MAGRO, *Robot,* cit., p. 1181.
[325] L. MEZZETTI, *Introduzione,* cit., p. 11, corsivi nostri.
[326] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 403.
[327] A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 76.
[328] U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 624.
[329] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1772.
[330] E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale,* cit., p. 77.
[331] Dà conto dell'assenza di un unico quadro legislativo che disciplini la materia dell'intelligenza artificiale U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 624. Avverte della necessità di affrontare la questione a un livello comune C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 136.
[332] F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 10; P. SEVERINO, *Intelligenza artificiale*, cit., p. 544; A. AMIDEI, *Robotica intelligente e responsabilità,* cit., pp. 82-101.
[333] A. AMIDEI, *Robotica intelligente e responsabilità,* cit., pp. 64-65.
[334] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 4.

hard law rischierebbe di essere ineffettivo e inidoneo. Non sarebbe infatti pensabile una regolamentazione giuridica per prescrizioni generali che si adatti uniformemente a tutti i settori che si servono l'IA: sarebbe un "*one size fits all approach*": «occorre invece immaginare più modelli di regolazione "*tailor made*" che, pur entro una cornice unitaria, si dimostrino opportunamente flessibili, ovvero sufficientemente stringenti per offrire tutela agli interessi coinvolti di volta in volta ma, al contempo, adeguati alla diversità delle tecnologie regolate e adattabili con il mutare delle condizioni»[335]. Se però da un lato si è constatato che probabilmente la flessibilità degli strumenti di *soft law* si adatterebbe meglio alle esigenze dinamiche dell'IA, non si è potuto fare a meno di ravvisare che tale tecnica normativa potrebbe non essere idonea ad assicurare l'armonizzazione cui il settore aspira, facendo riemergere la necessità di un, seppur minimo, sostrato di *hard law*[336].

Chiaro è che non sarà compito semplice «adeguare i singoli ordinamenti nazionali assicurando spazi di disciplina uniforme in ambito unionale»[337], per non parlare della difficoltà di creare un'adeguata armonizzazione tecnico-normativa a livello transnazionale[338]. Altro interrogativo che ci si dovrà porre è quanta parte di queste nuove questioni potrà trovare soluzioni adeguate nell'*evoluzione interpretativa* delle normative vigenti e quanta, invece, richiederà una specifica *evoluzione normativa*[339].

Ragionando a livello europeo, altro fattore destinato a incidere sulla regolamentazione del settore dell'IA sarà connesso al tipo di ordinamento con cui ci si confronterà, se di *civil law* o di *common law*: l'approccio empirico di quest'ultimo, maggiormente legato al dato esperienziale e al precedente[340], probabilmente si adatterà meglio alla materia *de qua* rispetto all'approccio maggiormente sistemico degli ordinamenti di *civil law*[341].

Non è probabilmente casuale che i sistemi di *common law* (in particolare quello statunitense)[342] siano quelli che ad oggi vantano i più evoluti sviluppi nel settore dell'IA. Il ritardo dell'Europa[343] nella materia in esame potrebbe acuire il

---

[335] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 419. L'A. prosegue riscontrando una sorta di superamento della *hard law* da parte della *soft law*, prendendo altresì atto, da un lato, della scarsità della produzione di norme vincolanti in materia (più che altro concentrate in ambiti settoriali) e, dall'altro, di una maggiormente cospicua produzione di atti di *soft law*, p. 421.

[336] E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale,* cit., pp. 78-84.

[337] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 5.

[338] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1773.

[339] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 4.

[340] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 338.

[341] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 3.

[342] Ad oggi non risulta che gli Stati Uniti d'America abbiano adottato una normativa federale in materia di intelligenza artificiale, concentrandosi piuttosto su una normazione specializzata e di settore (ad es. in materia di droni o *self driving cars*), A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 67, nota 7. Il modello statunitense sembra per lo più basarsi su un approccio "auto-regolativo" della concorrenza e del mercato, U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 616. Detto in altri termini, su un modello *business oriented*, opposto a quello *regulation oriented* dell'Unione Europea, R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 36. Per un approfondimento sull'attuale quadro giuridico americano e, in particolare, sulla materia della *product liability law*, v. M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., pp. 348 ss.

[343] C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 98; A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 67.

divario fra i vari Paesi del mondo[344] e portare l'Unione a rischiare di «"subire" standard e norme eterodeterminate ed "imposte" da un mercato globale che rischia di funzionare su una sorta di *first come-first served basis*»[345]. Il ritardo di cui stiamo parlando potrebbe essere acuito anche dall'interno dell'Unione, nell'eventualità in cui alcuni Stati membri si dotassero di autonome normative in materia di IA[346], finendo così per "trascinare" l'intervento dell'UE e per creare il c.d. fenomeno di "osmosi a doppio senso" tra le normative nazionali e quella europea[347].

Ci troviamo al cospetto di tecnologie che non ci concedono il tempo di abituarci ad esse[348]: si tratta di un «problema di "*metabolizzazione*" dell'innovazione da parte di una società tarata su ritmi di sviluppo più lenti (…) Lo sviluppo, per essere sostenibile, deve dare al cittadino il tempo di metabolizzare l'innovazione, possibilmente senza rallentarla»[349]. Questo tempo non è sempre concesso al cittadino e ciò si fa tanto più evidente quando il progresso della ricerca tecnologica va ad incidere sul mondo del lavoro che lo

[344] E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale,* cit., p. 74. Sul passaggio da un *digital divide* a un *algorithmic divide* G. ITALIANO, *Intelligenza Artificiale*, cit., p. 222. Altra parte della dottrina si concentra, a livello più generale, sul c.d. *robotic divide*: «una forte disuguaglianza fra individui, fasce sociali, peasi, interi continenti, a causa dei costi e delle difficoltà di accesso alle nuove risorse tecnologiche» R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., pp. 44-45; sul punto v. anche A. TURANO, *Robotica e roboetica*, cit., p. 140. La Cina e gli Stati Uniti si contendono il primato per il predominio nel campo dell'intelligenza artificiale, G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 1. Nel 2017 il Consiglio di Stato cinese ha pubblicato il *Next Generation Artificial Intelligence Development Plan*, una strategia per raggiungere entro il 2030 il ruolo di centro globale per l'innovazione in materia di intelligenza artificiale, consultabile al sito https://www.mfa.gov.cn/ce/cefi/eng/kxjs/P020171025789108009001.pdf. Tale Piano viene richiamato anche da A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 68, nota 9, cui si rimanda per qualche ulteriore considerazione a proposito dello sviluppo della materia dell'IA in Giappone e in Corea del Sud.

[345] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 78.

[346] U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 623. La dottrina usa parlare a tal proposito di "federalismo sperimentale", il quale non mira a realizzare una regolamentazione unitaria bensì una varietà di approcci regolativi a livello statale, sulla falsariga di quanto accade già in America, ad esempio, in materia di *self-driving cars*, E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale,* cit., p. 89.

[347] «Più volte categorie giuridiche facenti parte degli ordinamenti nazionali degli Stati membri, anche qualora particolarmente specifiche e peculiari, siano confluite nel lessico del legislatore dell'Unione europea, dando vita ad un fenomeno di circolazione dei modelli normativi e concetti giuridici non soltanto "verticale dall'altro al basso" (ossia dalla normativa europea a quella dei singoli Stati membri, qual portato del primato del diritto dell'Unione europea su quello nazionale) e non soltanto "orizzontale" (tra i singoli Stati europei), ma anche "verticale dal basso all'alto"» A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 69, nota 11.

[348] Esse entrano nella nostra vita quotidiana senza una reale percezione da parte nostra, G. MOBILIO, *L'intelligenza artificiale*, cit., p. 402. «Il progresso irrompe, non chiede permesso», G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 8. Altra dottrina, pur riferendosi alla Robotica, parla di un'"esplosione cambriana" per definire la grande velocità di evoluzione e diversificazione in materia. Riportiamo questa definizione in quanto riteniamo che il riferimento all'esplosione cambriana possa perfettamente adeguarsi anche alla materia dell'intelligenza artificiale più genericamente considerata, R. CALO, *Robots in American Law*, in *University of Washington School of Law Legal Studies Research Paper*, n. 4, 2016, p. 39.

[349] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 40.

riguarda tanto da vicino[350]. Lo sfasamento temporale tra il lesto avanzamento della tecnica e le difficoltà sociali dello stare al passo con esso[351] sono destinate ad alimentare problemi preesistenti e a introdurne di nuovi. Ci troviamo in una nuova fase nuova dell'esperienza umana[352] in cui l'arrivo nelle nostre vite delle moderne tecnologie (in un generale contesto di incertezza) non è concretamente arrestabile[353] e che ci porterà a tentare di prevedere le possibili dinamiche che l'IA porterà con sé e a pensare a strategie idonee per tutelare i consociati[354].

Iniziare a riflettere seriamente su questi temi è importante per non aggravare il ritardo del diritto[355], anche penale[356]: ciò pare opportuno non solo al fine di «alimentare una razionale riflessione pubblica su questi temi»[357], ma anche per procedere ad una adeguata formazione dei membri della nuova società digitale[358]. Serve una «politica del diritto lungimirante (…) nella proposizione *de iure condendo* ma altresì alla interpretazione *de iure condito*»[359], un approccio umano-centrico idoneo a creare un clima di fiducia e sicurezza per i cittadini[360]. Proprio in considerazione del fatto che il diritto è probabilmente lo strumento migliore di cui dispongono gli organi di governo per regolare una materia di tal fatta, occorrerà averne ben presente criticità e rischi per creare un quadro normativo in grado di non soffocare lo sviluppo tecnologico ma, altresì, di non richiedere di essere frequentemente rivisto in ragione dello stesso sviluppo della tecnologia.

Ciò che stiamo cercando di dire, in sintesi, è che tradizionalmente «il diritto, inteso in senso ampio (…), ha accompagnato passo passo l'evoluzione delle nuove tecnologie, studiandone gli effetti e promuovendone lo sviluppo»[361], senza limitarsi ad intervenire solo dopo il sorgere di eventuali (e ineliminabili) problemi. Spetta dunque al diritto aprire questa nuova strada in quanto è il solo in grado di

---

[350] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 39, gli AA. evidenziano anche che l'avvento delle nuove tecnologie nel mondo del lavoro imporrà un *reskilling* dei dipendenti, riqualificandoli dal punto di vista professionale, p. 41; altra parte della dottrina chiarisce che l'avvento dell'intelligenza artificiale è destinato ad incidere sul mondo del lavoro, sostituendo l'uomo in mansioni noiose e pericolose ma anche creando nuove figure professionali: «pare necessario impostare politiche di intervento che possano essere costituzionalmente orientate, al fine di perseguire l'obiettivo di rendere la Repubblica, in termini concreti, "fondata sul lavoro"», C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., pp. 112-113. Riporta uno studio di Oxford sull'impatto dell'IA nel mondo del lavoro M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 29; si concentrano sul tema anche G. ITALIANO, *Intelligenza Artificiale*, cit., p. 223; C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 8; E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale*, cit., p. 74; A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 15; A. TURANO, *Robotica e roboetica*, cit., p. 137; P. MORO, *Libertà del robot?*, cit., p. 538.

[351] L. MEZZETTI, *Introduzione,* cit., p. 13.

[352] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 8.

[353] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 88.

[354] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 41 il quale efficacemente afferma che «l'innovazione si *autogenera*», corsivo nostro.

[355] C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 102.

[356] F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 3.

[357] P. SEVERINO, *Intelligenza artificiale*, cit., p. 536.

[358] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 8.

[359] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 3.

[360] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 424.

[361] G. ALPA, *Prefazione*, in U. RUFFOLO, *Intelligenza artificiale. Il diritto, i diritti, l'etica*, cit., p. XVII.

accompagnare – da sempre – processi di cui la società diviene protagonista[362]. Un diritto, insomma, che non faccia da chiudi fila bensì da apripista. Ci accingiamo, adesso, ad affrontare con lo strumento del diritto penale i punti problematici (per non dire deboli) dell'intelligenza artificiale[363].

SEZIONE SECONDA
QUESTIONI DI RESPONSABILITÀ PENALE

10. Quale ruolo per il diritto penale?

Parte della dottrina ha suggestivamente affermato che intelligenza artificiale e diritto penale avrebbero obiettivi comuni: «se il nostro "diritto penale umano tradizionale" (di qualsiasi provenienza teorica) può solo garantire la tutela dei beni giuridici in modo normativo e controfattuale, perché le violazioni delle norme rimangono all'ordine del giorno, l'IA persegue nel lungo periodo l'obiettivo della pratica impossibilità o almeno della sostanziale minimizzazione delle lesioni ai beni giuridici»[364].

Occorre a questo punto chiederci se le norme attualmente vigenti che regolano il diritto penale siano suscettibili di essere applicate anche alle problematiche penalistiche dell'IA[365]. V'è chi ritiene che, così come strutturato, il diritto penale non sia adeguato a fronteggiare le moderne sfide dell'IA[366] e che i suoi strumenti non siano in grado di governare i c.d. "*rischi da ignoto tecnologico*"[367]. È stato prospettato il rischio che il diritto penale, innanzi all'IA, appaia come uno strumento vecchio[368] in quanto i suoi canoni classici (in punto di condotta, nesso causale, colpevolezza, tutela dei beni giuridici e, più in generale, degli elementi costitutivi del reato) sarebbero incompatibili con le peculiarità dell'IA[369]. La dottrina più rigida sul punto rifiuta l'idea di un «diritto penale strutturato secondo i canoni classici ma interpretato alla luce di mutate categorie concettuali che quei canoni, di fatto, stravolgono»[370].

Altra parte della dottrina propone, più cautamente, di ripensare in ottica attualizzante gli istituti tradizionali e di creare, ove necessario, specifici modelli *ad hoc*[371]. Questa evoluzione tecnologica imporrà al diritto un adeguamento al fine di fronteggiare i nuovi fenomeni che verranno, mantenendo la propria funzione regolatoria. Per far ciò il penalista si servirà certamente, in un primo

---

[362] G. CORASANITI, *Intelligenza artificiale e diritto: il nuovo ruolo del giurista*, in U. RUFFOLO, *Intelligenza artificiale. Il diritto, i diritti, l'etica*, cit., p. 402.
[363] C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1929.
[364] C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1934.
[365] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1750.
[366] M.B. MAGRO, *Robot*, cit., p. 1205.
[367] L. STORTONI, *Angoscia tecnologica*, cit., p. 74.
[368] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1748.
[369] L. STORTONI, *Angoscia tecnologica*, cit., p. 74, il quale invero si riferisce al progresso tecnologico generalmente inteso.
[370] L. STORTONI, *Angoscia tecnologica*, cit., p. 83.
[371] P. SEVERINO, *Intelligenza artificiale*, cit., p. 532; G. CAPILLI, *I criteri di interpretazione*, cit., p. 458; G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 103 propone, ad esempio, di avvalersi di un sistema di interpretazione evolutiva contribuendo al progresso osmotico di queste nuove tecnologie, senza necessariamente far immediato ricorso ad una nuova normativa; A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 80, afferma che ad un problema inedito non devono necessariamente seguire soluzioni inedite.

momento, delle frecce che possiede già al suo arco, mediante un'interpretazione evolutiva degli strumenti che conosce (ove dotati di una *eadem ratio* che possa giustificare l'assimilazione del regime vigente all'IA)[372], per poi procedere, ove necessario, alla modifica delle norme preesistenti e, financo, alla creazione di norme nuove, senza mai abbandonare i principi fondamentali e le garanzie proprie del nostro diritto penale[373], nonché la tutela dei diritti fondamentali[374]. Rivisitare i paradigmi tipici della responsabilità penale non sarà semplice: «in un campo dove i principi di tassatività e tipicità, insieme al divieto di analogia *in malam partem*, impongono naturalmente cautele maggiori, se non veri e propri ostacoli, rispetto a ogni possibile operazione interpretativa che culmini nell'applicazione di *nuove* forme di responsabilità»[375]. Ciò che pare indubbio è che un approccio conservatore non sia adeguato per affrontare la materia *de qua*. Si dovrà delineare un rinnovato quadro giuridico-penale ove poter collocare rielaborati criteri d'imputazione della responsabilità (specie in considerazione dell'autonomia di cui sono dotati questi sistemi)[376] e definire, in questa nuova realtà, i connotati dei comportamenti penalmente rilevanti[377].

Altro punto su cui verosimilmente si registrerà in dottrina un acceso dibattito concernerà l'opportunità o meno di introdurre nuove fattispecie di reato aventi come protagonisti i sistemi intelligenti. C'è chi ritiene la non proficuità di una «perenne rincorsa legislativa»[378], ossia di un costante lavorio da parte del legislatore volto a introdurre nuove fattispecie penali per colmare ogni lacuna giuridica (più o meno presunta). Altri ritengono invece che un maggiore rispetto del principio di legalità imporrebbe l'introduzione di specifiche tipologie delittuose volte a regolamentare una nuova generazione di reati commessi non mediante l'IA ma da essa[379].

Queste nuove domande non devono sorprenderci: l'obsolescenza delle normative vigenti costituisce cifra stilistica del rapporto tra diritto e tecnologia. Il progresso tecnologico ha sempre messo in discussione l'efficacia e l'idoneità delle norme giuridiche, le quali mostrano in tal sede tutta la loro precarietà[380]. Ciò è già avvenuto con l'avvento di internet e con il passaggio dai *computer crimes* (reati informatici) ai *cybercrimes* (reati cibernetici)[381]. In via di mera esemplificazione basti pensare che, se il diritto non si adeguasse allo sviluppo tecnologico, ad oggi non avremmo una specifica circostanza aggravante per gli atti persecutori commessi mediante strumenti informatici o telematici, come non avremmo una normativa *ad hoc* sul captatore informatico.

L'intelligenza artificiale, dal canto suo, possiede alcune peculiarità penalistiche. Se in apertura abbiamo detto che l'IA mira a ridurre la lesione dei

---

[372] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 353.
[373] L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., pp. 40-94.
[374] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 370.
[375] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 340.
[376] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 19; M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 352.
[377] L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., pp. 42-43.
[378] L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., p. 42.
[379] M.B. MAGRO, *Robot*, cit., p. 1206; U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 622; F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 27.
[380] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 333.
[381] L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., p. 47.

beni giuridici, dobbiamo prendere atto che tale modo di operare rischia di alimentare il c.d. paradigma *zero trust*: ad esempio, gli *smart contracts* renderanno impossibili le frodi contrattuali[382] e i sistemi di *criminal compliance* nelle aziende aiuteranno a individuare i settori maggiormente a rischio e a intercettare eventuali condotte illecite[383], realizzando così una cooperazione pubblico-privata nella prevenzione della criminalità[384]. La diminuzione delle lesioni dei beni giuridici che ne deriverebbe comporterebbe altresì l'anticipazione della relativa soglia di tutela, propendendo per un diritto penale securitario volto a sacrificare la libertà dei cittadini[385] e a realizzare una sorta di "paternalismo tecnologico"[386]. Tale modo di ragionare rischierebbe di condurre a un controllo capillare dei consociati al fine di prevenirne possibili comportamenti illeciti, ponendo fine alla controfattualità del diritto penale liberale che siamo abituati a conoscere. D'altro canto, accogliere una prospettiva liberale garantirebbe le libertà dei cittadini[387] accettando, di contro, il possibile sacrificio di alcuni beni giuridici che sarebbero lesi da eventuali comportamenti illeciti e che troverebbero tutela *ex post*[388]. Probabilmente la via preferibile sarebbe quella della ricerca di una soluzione di compromesso, che miri a bilanciare la tutela dei beni giuridici suscettibili di lesione e il rispetto delle libertà democratiche.

Altro bilanciamento che occorrerà svolgere in materia riguarderà il rapporto tra un approccio paternalistico che, vietando ogni tecnologia foriera di rischi, rispetti il principio di precauzione al costo di limitare il progresso tecnologico, e un approccio liberale volto a diffondere l'IA indipendentemente dai rischi che questa possa comportare per le persone[389]. Anche in tal sede si dovrà, a nostro avviso, optare per una terza via che bilanci, da un lato, i benefici in termini di miglioramento della qualità di vita che l'IA porterà con sé e, dall'altro, i rischi che non verranno ritenuti socialmente accettabili. In tal senso si renderà fondamentale una struttura di norme giuridiche in materia di responsabilità che mirino a realizzare una soddisfacente allocazione delle conseguenze connesse alla

---

[382] C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1921.
[383] P. SEVERINO, *Intelligenza artificiale*, cit., p. 538.
[384] A. GULLO, *Nuove frontiere tecnologiche e sistema penale*, cit., p. VII.
[385] C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 119. Ad esempio, esistono sistemi di *criminal compliance* che, mediante un monitoraggio completo delle comunicazioni che avvengono all'interno di un'azienda, consentono di intercettare interazioni sospette e di intervenire in via preventiva, C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1921, nota 42. Al fine di prevenire possibili reati, dunque, si esercita un controllo capillare su tutte le comunicazioni dei dipendenti dell'azienda, incidendo sulla loro libertà.
[386] C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1937.
[387] «Solo in questo modo la mancata commissione di un reato potrebbe continuare ad essere considerata sotto il *profilo individuale* come una libera decisione a favore del rispetto della legge e la sua commissione come una libera decisione contro la legge. In un (ipotetico) mondo in cui l'intelligenza artificiale rende i fatti di reati *eo ipso* o *de facto* impossibili, questa libertà, la famosa capacità di agire diversamente – anche se questa fosse solo una finzione necessaria per una comunità liberale – non viene più in rilievo» C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1936. L'A. prosegue più avanti affermando che «il guadagno virtuale di libertà dei molti che non sono esposti alla sorveglianza dovrebbe essere preferito alla reale perdita di liberà dei pochi che diventano vittime di reati che (probabilmente) si sarebbe potuto prevenire attraverso l'impiego dell'IA» p. 1938.
[388] C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 119.
[389] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 353.

produzione di un illecito o, più generalmente, dei rischi accettati dalla società e derivanti dall'interazione con l'IA[390]. Riteniamo, dunque, che il diritto penale non dovrebbe limitarsi a predisporre una regolamentazione meramente reattiva, che si limiti cioè a contrastare le conseguenze negative dell'intelligenza artificiale, bensì proattiva, ossia volta a intercettarne i rischi e a gestirne le problematiche[391].

Prescindendo dall'accoglimento di un approccio securitario o liberista, di un modello di precauzione o liberale, reattivo o proattivo, ciò che sembra fuor di dubbio è che la materia del rapporto tra intelligenza artificiale e diritto penale vada affrontata in modo dinamico, nella consapevolezza che nonostante tutti i tentativi che si potranno fare per prevenire (o financo impedire) la commissione delle più disparate fattispecie delittuose mediante l'uso dei sistemi intelligenti, la criminalità sarà sempre in grado di reinventarsi e di trovare nuovi modi per delinquere che, semplicemente, non si erano in precedenza verificati.

11. Alcuni spunti penalistici dal diritto civile.

La nostra ricerca ruota intorno all'imputazione della responsabilità. L'interrogativo più rilevante che dobbiamo porci in questo ambito e che fino ad ora abbiamo solo lambito riguarda il problema di come attribuire la responsabilità per i danni cagionati da sistemi intelligenti dotati di una certa autonomia[392].

Come abbiamo già avuto modo di accennare, la Risoluzione del Parlamento europeo recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica[393] ha proposto due modelli di imputazione della responsabilità: quello della responsabilità oggettiva da un lato e quello della gestione dei rischi dall'altro[394].

Secondo il modello della *responsabilità oggettiva* è sufficiente individuare il nesso causale tra l'agire del sistema intelligente e il danno verificatosi, attribuendo la responsabilità al "danneggiante" (con le relative difficoltà nell'individuarlo) indipendentemente da un suo comportamento colpevole e, dunque, dalla presenza di un qualsivoglia elemento soggettivo[395]. Tale modello difficilmente troverebbe

---

[390] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 354.

[391] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1748; a proposito dell'utilità di un approccio proattivo v. G. MOBILIO, *L'intelligenza artificiale*, cit., p. 405.

[392] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 337; A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 12.

[393] Il problema cruciale della nostra ricerca è ben sintetizzato in un passaggio della suddetta Risoluzione ove si chiarisce che: «nell'ipotesi in cui un robot possa prendere decisioni autonome, le norme tradizionali non sono sufficienti per attivare la responsabilità per i danni causati da un robot, in quanto non consentirebbero di determinare qual è il soggetto cui incombe la responsabilità del risarcimento né di esigere da tale soggetto la riparazione dei danni causati» *Risoluzione del Parlamento europeo*, cit., Considerando AF. Passaggio riportato anche da A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 69; A. TURANO, *Robotica e roboetica*, cit., p. 149.

[394] *Risoluzione del Parlamento europeo*, cit., Punti 53 ss. «ritiene che il futuro strumento legislativo debba essere fondato su una valutazione approfondita della Commissione che stabilisca se applicare l'approccio della responsabilità oggettiva o della gestione dei rischi; osserva al contempo che la *responsabilità oggettiva* richiede una semplice prova del danno avvenuto e l'individuazione di un nesso di causalità tra il funzionamento lesivo del robot e il danno subito dalla parte lesa; constata che l'approccio di *gestione dei rischi* non si concentra sulla persona "che ha agito con negligenza" in quanto responsabile a livello individuale bensì sulla persona che, in determinate circostanze, è in grado di minimizzare i rischi e affrontare l'impatto negativo», corsivo nostro.

[395] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 87.

cittadinanza in materia penale ove, come è noto, la responsabilità oggettiva è fortemente marginalizzata[396].

Il modello della *gestione del rischio*, invece, ripartirebbe la responsabilità tra i soggetti deputati a minimizzare il rischio dell'evento lesivo, individuando il responsabile nell'operatore che, in base al tipo di danno verificatosi, sarebbe stato maggiormente in grado di prevenirlo[397]. L'individuazione dei soggetti potenzialmente responsabili di un dato evento lesivo avviene mediante un'analisi di *risk management*, volta dunque a individuare i soggetti "più vicini" al prodotto e, pertanto, dotati di un maggiore potere impeditivo di eventuali rischi. A ben vedere, il tipo di responsabilità cui sarà assoggettato colui che verrà individuato come responsabile del danno cagionato dall'IA sarà comunque una forma di *strict liability*[398], non essendo necessario che sussista in capo a quest'ultimo l'elemento soggettivo del dolo o della colpa[399].

La più attenta dottrina ha infatti notato come i modelli della responsabilità oggettiva e della gestione del rischio siano, in realtà, «due "diversi momenti" del medesimo fenomeno, ossia come due componenti parimenti necessarie per addivenire ad una effettiva allocazione della responsabilità tra soggetti coinvolti»[400]. Non si tratterà dunque di fare una scelta tra i due modelli in quanto essi non dovrebbero essere considerati alternativamente bensì in via cumulativa e integrativa[401].

Abbiamo ritenuto opportuno dar conto, seppur brevemente, di alcune considerazioni svolte da un punto di vista civilistico[402] in quanto ci sentiamo in tal

---

[396] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 364.

[397] A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 87; U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 11; A. TURANO, *Robotica e roboetica*, cit., p. 151.

[398] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 12; A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 87.

[399] A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 87.

[400] A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 86.

[401] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 12; G. CAPILLI, *I criteri di interpretazione*, cit., p. 478.

[402] La dottrina civilistica ha già indicato alcune possibili strade per fronteggiare gli eventi lesivi cagionati dall'IA. «Se ne potrebbe trattare, così, in termini di: responsabilità del produttore (ex art. 114 del codice del consumo) che chiama quest'ultimo a risarcire il danno cagionato dai difetti del suo prodotto; responsabilità per l'esercizio di attività pericolose (ex. art. 2050 c.c.) secondo cui chiunque cagioni il danno è tenuto al risarcimento se non prova di avere adottato tutte le misure idonee a evitarlo; responsabilità per la custodia di animali (ex art. 2052 c.c.) secondo cui il proprietario o chi se ne serve è comunque responsabile dei danni salvo che provi il caso fortuito; responsabilità in vigilando e in educando di genitori e maestri d'arte (ex art. 2048 c.c.) secondo cui i genitori sono responsabili del danno cagionato dal fatto illecito dei figli minori che abitano con essi, e i precettori e coloro che insegnano un mestiere o un'arte sono responsabili del danno cagionato dal fatto illecito dei loro allievi nel tempo in cui sono sotto la loro vigilanza, a meno che non provino di non aver potuto impedire il fatto. In altra prospettiva, si è proposta una assicurazione obbligatoria che permetta in ogni caso la copertura del danno risarcibile oppure una responsabilità diffusa fra quanti hanno partecipato alla produzione, all'assemblaggio, alla programmazione della macchina, sul modello della *corporation* di *common law*» C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 105. Il tema dell'assimilazione tra sistemi intelligenti e animali o minori è quello su cui si è registrato il più acceso dibattito. Con riferimento al rapporto tra animali e forme di IA si è constatato che entrambi sarebbero infatti dotati di un'intrinseca pericolosità, nonché in grado di porre in essere condotte non prevedibili. Con riguardo, invece, ai danni causati da minori o animali il codice civile prevede una responsabilità oggettiva che si può evitare fornendo una prova liberatoria se si dimostra, nel primo caso, di non

sede di condividere l'opinione di quella parte di dottrina la quale ritiene che in questa materia le ricostruzioni in punto di responsabilità civile e penale presentano considerevoli elementi di continuità[403]. L'idea alla base del sistema della gestione del rischio può infatti essere considerata valida anche in ottica penalistica. Individuare il soggetto più vicino al malfunzionamento, senza però considerarlo responsabile di default, potrebbe essere una soluzione penalisticamente orientata. Andare a ricercare la presenza della componente soggettiva, quantomeno in termini di colpa, infatti, legittimerebbe un'imputazione anche a titolo penale in quanto, in tal modo, verrebbe integrato anche l'elemento della colpevolezza. Certo, non sarebbe improbabile, in un contesto di rischio come quello in esame, non rinvenire alcun elemento di colpevolezza in capo all'operatore deputato al controllo e alla supervisione del sistema intelligente o, comunque, non riuscire a provarlo oltre ogni ragionevole dubbio, specie in considerazione dell'elevato grado di autonomia che può essere raggiunto dai sistemi intelligenti. In simili ipotesi la vittima del danno causato dall'IA si troverebbe sprovvista di tutela giuridico-penale, vedendosi così costretta a tentare una richiesta risarcitoria in sede civile, ove però l'esito positivo non sarebbe affatto scontato. Se tale scenario non ci sembra ammissibile, dobbiamo del pari rifuggire dal ragionamento opposto: sarebbe infatti massimamente iniquo considerare coloro che, a diverso titolo, hanno collaborato alla creazione dell'IA, sempre responsabili per i danni cagionati da quest'ultima nel corso del suo agire autonomo[404].

Sembra quasi che nella materia oggetto d'esame il binomio tra diritto civile e diritto penale non appaia del tutto inconciliabile, in quanto entrambi inizieranno a porsi interrogativi comuni: chi sono i nuovi custodi della cosa intelligente? Come si delineano i rapporti tra caso fortuito e autoapprendimento deviante dell'IA? Come si articolerà il meccanismo della responsabilità vicaria? Nell'ambito di attività pericolose come queste, in che modo si fornirà la prova liberatoria di aver adottato tutte le misure necessarie ad impedire la verificazione del danno?

---

aver potuto impedire il fatto (art. 2048 c.c.) e, nel secondo caso, l'intervento di un caso fortuito (art. 2052). In realtà parte della dottrina ha osservato come «l'assimilazione del regime di responsabilità per i danni prodotti dall'intelligenza artificiale a quello applicabile ai padroni di animali o di genitori di minori ha registrato consensi soltanto parziali. In dottrina si è spiegata la differenza tra animali e sistemi di Intelligenza Artificiale evocando l'esistenza, nei primi, di componenti che giocoforza non si ritrovano nei secondi: e, segnatamente, la capacità di autodeterminazione e l'istinto che consentono loro di svolgere anche azioni diverse da quelle per cui sono stati "addomesticati" o istruiti. È in presenza di questi elementi che gli animali possono incorrere in "deviazioni" che comportano conseguenze dannose, mentre i sistemi di Intelligenza Artificiale non possono svolgere o elaborare processi diversi da quelli per cui sono stati progettati: possono però computare in modo erroneo le variabili e addivenire a un calcolo sbagliato. Il grado di autonomia è quindi ben diverso, sebbene i critici di questa impostazione che equipara l'Intelligenza Artificiale (e più precisamente i robot) agli animali addomesticati non escludano la possibilità che il programmatore e/o l'utente possano essere gravati da responsabilità per effetto di un comportamento indesiderato» M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., pp. 360-361. L'assimilazione con animali e minori non convince neanche G. CAPILLI, *I criteri di interpretazione*, cit., p. 476. Sul tema v. anche U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 15.

[403] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 363.

[404] Concetto analogo rinvenibile anche in G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 108.

Iniziamo a chiederci se e in che misura gli strumenti classici del diritto possano essere in grado di governare gli scenari futuri che l'IA ci proporrà[405].

12. Evitare due eccessi: dalla deresponsabilizzazione alla responsabilità oggettiva.

Come pocanzi accennato, l'idea che i sistemi di intelligenza artificiale possano cagionare danni a terzi senza che esista un adeguato sistema di tutela crea un qualche disagio[406]. Dobbiamo dunque iniziare a chiederci chi può essere considerato giuridicamente responsabile delle decisioni (potenzialmente imprevedibili) assunte dall'IA e chi sarebbe chiamato a rispondere dei reati da questa commessi[407].

Si tratta di comprendere se gli attuali modelli di imputazione della responsabilità penale possano adattarsi a un reato causato dall'agire autonomo di una macchina e come debbano considerarsi ipotesi in cui la condotta penalmente rilevante sia opera condivisa dell'agire umano e artificiale (*culpa in interagendo*). Ci si dovrà inoltre chiedere se la causazione dell'evento lesivo sia da imputarsi a colui che ha ideato l'algoritmo, a colui che ha prodotto il sistema intelligente integrato dal suddetto algoritmo o, ancora, a colui che si è servito dell'IA[408].

I soggetti potenzialmente responsabili del danno causato dal sistema intelligente sono molteplici. Nella materia in esame si pone il c.d. *many hands problem*[409], espressione che in tal sede palesa la moltitudine di soggetti coinvolti durante tutto il processo di vita dell'IA (basti pensare alla fase di ideazione, realizzazione o programmazione). Si pone, conseguentemente, un problema di riparto di compiti e poteri[410], nonché della relativa «*individuazione delle singole responsabilità all'interno delle organizzazioni complesse*»[411]. Occorre inoltre considerare che la produzione di un sistema intelligente potrebbe non essere rimessa esclusivamente alla competenza di una singola azienda, ben potendo coinvolgere una pluralità di enti, ciascuno deputato a realizzare una porzione dell'IA. La questione è tutt'altro che semplice. Intanto si dovrebbe riuscire a dimostrare in giudizio quale segmento del processo produttivo sia rimasto coinvolto nell'evento lesivo cagionato dall'IA: se, dunque, vi sia stato un problema in fase di scrittura dell'algoritmo, in fase di progettazione o in fase di training del sistema (compiti, chiaramente, rimessi a soggetti diversi). Secondariamente, ove si riuscisse a individuare il soggetto competente, non potremmo considerarlo sempre e comunque responsabile del danno causato dall'IA, salvo voler ammettere la creazione di una vera e propria responsabilità

---

[405] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 10.
[406] M.B. MAGRO, *Robot*, cit., p. 1180.
[407] G. ITALIANO, *Intelligenza Artificiale*, cit., p. 224; A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 80; S. RIONDATO, *Robot: talune implicazioni di diritto penale,* cit., p. 89; C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1750; F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 24.
[408] V. MANES, *L'oracolo algoritmico,* cit., p. 549.
[409] I. SALVADORI, *Agenti artificiali*, cit., p. 107; M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 3 parla di "responsabilità distribuita"; C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 105 parla di "responsabilità diffusa"; sull'opportunità di estendere il concetto di "custodia" dell'IA a soggetti diversi M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 362.
[410] I. SALVADORI, *Agenti artificiali,* cit., p. 105.
[411] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1754.

per posizione[412] degli operatori coinvolti a vario titolo nel funzionamento del sistema intelligente.

Ove la responsabilità per l'agire illecito dell'IA si andasse a ricercare (come è probabile che avvenga) all'interno di un'organizzazione complessa, la problematica concernente la mancata individuazione dell'operatore responsabile potrebbe a prima vista risolversi facilmente. Infatti, nell'eventualità in cui non risulti possibile individuare il responsabile del funzionamento anomalo dell'IA da cui è scaturito un danno ma, al contempo, risulti indubbio che quest'ultimo sia da riferire a un difetto del sistema riconducibile alla violazione di uno standard di diligenza, sarebbe possibile sanzionare direttamente l'ente, ex art. 8 d. lgs 231/2001. Tuttavia emerge sin da subito un primo problema, probabilmente insormontabile, a meno di non voler procedere a una radicale modifica legislativa. La responsabilità da reato degli enti[413] è subordinata alla realizzazione di un lungo

---

[412] I. SALVADORI, *Agenti artificiali*, cit., p. 105.

[413] La letteratura sul tema è ormai sconfinata, pertanto, senza alcuna pretesa di esaustività, sia consentito un richiamo alla dottrina classica che si è occupata dell'argomento in prospettiva monografica: A. ALESSANDRI, *Il nuovo diritto penale delle società*, Milano, 2002; G. GARUTI, *Responsabilità degli enti per illeciti amministrativi dipendenti da reato*, Padova, 2002; G. LANCELLOTTI, *La responsabilità della società per il reato dell'amministratore*, Torino, 2003; F. SANTI, *La responsabilità delle società e degli enti. Modelli di esonero delle imprese. D.Lgs. 8/6/2001, n. 231. D.M. 26/6/2003, n. 201*, Milano, 2004; G. DE FRANCESCO, *La responsabilità degli enti: un nuovo modello di giustizia "punitiva"*, Torino, 2004; M.A. PASCULLI, *La responsabilità "da reato" degli enti collettivi nell'ordinamento italiano. Profili dogmatici ed applicativi*, Bari, 2005; R. GUERRINI, *La responsabilità da reato degli enti. Sanzioni e loro natura*, Milano, 2006; A. BASSI, T. EPIDENDIO, *Enti e responsabilità da reato. Accertamento, sanzioni e misure cautelari*, Milano, 2006; N. SELVAGGI, *L'interesse dell'ente collettivo: quale criterio di ascrizione della responsabilità da reato*, Napoli, 2006; G. RUGGIERO, *Contributo allo studio della capacità penale. Lo "statuto" della persona fisica e degli enti*, Torino, 2007; G. DE VERO, *La responsabilità penale delle persone giuridiche*, Milano, 2008; M. RIVERDITI, *La responsabilità degli enti: un crocevia tra repressione e specialprevenzione. Circolarità ed innovazione dei modelli sanzionatori*, Napoli, 2009; G. DE SIMONE, *Persone giuridiche e responsabilità da reato. Profili storici, dogmatici e comparatistici*, Pisa, 2012; G. CANZIO, L.D. CERQUA, L. LUPARIA, *Diritto penale delle società, profili sostanziali e processuali*, Padova, 2014; M. LEVIS, A. PERINI, *La responsabilità amministrativa della società e degli enti*, Bologna, 2014; F. CENTONZE, M. MANTOVANI, *La responsabilità «penale» degli enti. Dieci proposte di riforma*, Bologna, 2016; V. MONGILLO, *La responsabilità penale tra individuo ed ente collettivo*, Torino, 2018.
Non possiamo fare a meno di menzionare in tal sede i molteplici contributi, in particolare, di due autorevoli esponenti della dottrina. Si vedano, in proposito G. DE VERO, *Struttura e natura giuridica dell'illecito di ente collettivo dipendente da reato*, in *Riv. it. dir. proc. pen*., 2001, pp. 1126 ss.; G. DE VERO, *I reati societari nella dinamica evolutiva della responsabilità ex crimine degli enti collettivi*, in *Riv. it. dir. proc. pen*., 2003, pp. 720 ss.; G. DE VERO, *Il progetto di modifica della responsabilità degli enti tra originarie e nuove aporie*, in *Dir. pen. proc.*, 10/2010, pp. 1137 ss.; G. DE VERO, *Prospettive evolutive della responsabilità da reato degli enti collettivi*, in *La responsabilità amministrativa delle società e degli enti*, 2011, pp. 9 ss; G. DE VERO, *Il reo quale ente collettivo*, in G. DE VERO (a cura di), *La legge penale. Il reato. Il reo. La persona offesa*, Torino, 2015, pp. 523 ss. Sia consentito un rimando anche a C.E. PALIERO, *La responsabilità della persona giuridica nell'ordinamento italiano: profili sistematici*, in F. PALAZZO, *Societas puniri potest. La responsabilità da reato degli enti collettivi. Atti del Convegno organizzato dalla Facoltà di giurisprudenza e dal Dipartimento di diritto comparato e penale dell'Università di Firenze (15-16 marzo 2002)*, Padova, 2003, pp. 17 ss.; C.E. PALIERO, *La società punita: del* come, *del* perché, *e del* per cosa, in *Riv. it. dir. proc. pen.*, 2008, pp. 1516 ss.; C.E. PALIERO, *Responsabilità degli enti e principio di colpevolezza al vaglio della Cassazione: occasione mancata o definitivo* de profundis*?*, in *Le Società*, 2014, pp. 474 ss.

elenco di "reati presupposto" (nella loro componente tipica e antigiuridica) tra cui non rientrano attualmente i reati di omicidio o lesioni imputabili a un difetto di produzione[414]. Un secondo problema afferisce al fatto che si dovrebbe andare a dimostrare che il reato presupposto sia stato commesso nell'interesse o vantaggio dell'ente, il che sarebbe difficilmente immaginabile con riferimento ai summenzionati delitti contro la persona[415].

La difficoltà di individuare il responsabile del danno causato dall'IA all'interno di una fitta rete organizzativa, unita al grado di autonomia che tali sistemi possono raggiungere, rischiano di creare un fenomeno di *deresponsabilizzazione* degli operatori che gravitano intorno all'IA, come se il comportamento dei sistemi intelligenti fosse sconnesso da quello degli uomini che le realizzano, sollevando questi ultimi dalle loro responsabilità[416]. La ricostruzione dell'agire di tali sistemi, insomma, è in grado di mettere in difficoltà i rapporti eziologici intercorrenti tra la condotta del programmatore e l'evento lesivo cagionato dall'IA, creando conseguentemente un problema di imputazione

---

Da ultimo, sia concesso un rinvio ad J. DE FARIA COSTA, *Contributo per una legittimazione della responsabilità penale delle persone giuridiche*, in *Riv. it. dir. proc. pen.*, 1993, pp. 1238 ss.; M. ROMANO, Societas delinquere non potest *(nel ricordo di Franco Bricola)*, in *Riv. it. dir. proc. pen.*, 1995, pp. 1031 ss.; F. STELLA, *Criminalità di impresa: nuovi modelli di intervento*, in *Riv. it. dir. proc. pen.*, 1999, pp. 1254 ss.; C. DE MAGLIE, *La disciplina della responsabilità amministrativa delle persone giuridiche e delle associazioni*, in *Dir. pen. proc.*, 2001, pp. 1342 ss.; C. PIERGALLINI, *La disciplina della responsabilità amministrativa delle persone giuridiche e delle associazioni*, in *Dir. pen. proc.*, 2001, pp. 1342 ss.; D. PULITANÒ, *La responsabilità "da reato" degli enti: i criteri d'imputazione*, in *Riv. it. dir. proc. pen.*, 2002, pp. 415 ss.; A. GARGANI, *Imputazione del reato agli enti collettivi e responsabilità penale dell'intraneo: due piani irrelati?*, in *Dir. pen. proc.*, 2002, pp. 1061 ss.; ID., *Responsabilità collettiva da delitto colposo d'evento: i criteri di imputazione nel diritto vivente*, in *La Legislazione penale*, 11.1.2016; A. MANNA, *La c.d. responsabilità amministrativa delle persone giuridiche: un primo sguardo d'insieme*, in *Riv. trim. dir. pen. econ.*, 2002, pp. 501 ss.; ID., *La c.d. responsabilità amministrativa delle persone giuridiche: il punto di vista del penalista*, in *Cass. pen.*, 2003, pp. 1101 ss.; A. CARMONA, *La responsabilità degli enti: alcune note sui reati presupposto*, in *Riv. trim. dir. pen. econ.*, 2003, pp. 995 ss.; G. DE FRANCESCO, *Gli enti collettivi: soggetti dell'illecito o garanti dei precetti normativi?*, in *Dir. pen. proc.*, 2005, pp. 753 ss.; G. AMARELLI, *Profili pratici della questione sulla natura giuridica della responsabilità degli enti*, in *Riv. it. dir. proc. pen.*, 2006, pp. 151 ss.; E. AMODIO, *Rischio penale di impresa e responsabilità degli enti nei gruppi multinazionali*, in *Riv. it. dir. proc. pen.*, 2007, pp. 1287 ss.; G. MARINUCCI, *La responsabilità penale delle persone giuridiche. Uno schizzo storico-dogmatico*, in *Riv. it. dir. proc. pen.*, 2007, pp. 445 ss.; A. ROSSI, *La responsabilità degli enti: i soggetti responsabili ed i modelli organizzativi*, in R. BARTOLI, *Responsabilità penale e rischio nelle attività mediche e d'impresa: (un dialogo con la giurisprudenza)*, Firenze, 2010, pp. 393 ss.; G. DE SIMONE, *La responsabilità da reato degli enti: natura giuridica e criteri (oggettivi) d'imputazione*, in *dirittopenalecontemporaneo.it*, 28.10.2012; M. PELLISSERO, *L'estensione della responsabilità degli enti i reati colposi. Una riflessione sui rapporti tra parte generale e parte speciale del d. lgs. 231/2001*, in *Scritti in onore di Alfonso M. Stile*, Napoli, 2013, pp. 1199 ss.; M.A. BARTOLUCCI, *L'art. 8 d.lgs. 231/2001 nel triangolo di Penrose. Tra minimizzazione del rischio-reato d'impresa e "nuove forme" di colpevolezza*, in *dirittopenalecontemporaneo.it*, 9.1.2017; F. CONSULICH, *Il principio di autonomia della responsabilità dell'ente. Prospettive di riforma dell'art. 8*, in *La responsabilità amministrativa delle società e degli enti*, 2018, pp. 197 ss.

[414] Per un approfondimento sul tema v. C. PIERGALLINI, *Intelligenza artificiale*, cit., pp. 1755 ss. L'A. ricorda i tre paradigmi imputativi idonei a fondare la responsabilità ex decreto 231: la mancata auto-organizzazione, l'inadeguata auto-organizzazione e l'ipotesi in cui l'autore del reato sia rimasto ignoto.

[415] I. SALVADORI, *Agenti artificiali*, cit., p. 106.

[416] D. TAFANI, *Sulla moralità artificiale*, cit., p. 96.

della relativa responsabilità[417]. L'affievolirsi della responsabilizzazione verrebbe acuito ove si considerasse l'agire autonomo dell'IA come causa sopravvenuta da sola sufficiente a determinare l'evento ex art. 41 co. 2 c.p. e, dunque, come fattore idoneo a interrompere il nesso causale[418]. Ciò equivarrebbe a ritenere che il danno cagionato dall'IA non sia riconducibile al suo produttore bensì alla macchina medesima[419]. L'interruzione del nesso causale romperebbe la catena attributiva tra l'agire dell'operatore (ad es. il programmatore dell'IA) e il danno causato dalla macchina[420], creando così un vuoto di responsabilità[421].

Analogo discorso varrebbe ove considerassimo l'agire autonomo dannoso dell'IA come un'ipotesi di caso fortuito o forza maggiore[422], trattandosi pur sempre di rischi dotati di un certo margine di ineliminabile imponderabilità. Invero parte della dottrina ha correttamente osservato che il rischio innescato dall'agire autonomo dell'IA non è equiparabile a quelli naturalistici che normalmente integrano le ipotesi di caso fortuito o forza maggiore che non possono essere controllate dall'uomo: «il pericolo in parola, invece, ha natura tecnologica: e pertanto sempre potenzialmente soggetto all'alternativa cautelare più radicale di tutte, ovvero la vera e propria *proibizione* dell'attività»[423]. Dovremmo considerare inoltre che l'imprevedibilità dell'IA è fattore in essa insito al punto da non poter considerare il suo agire autonomo come causa esterna e, dunque, come caso fortuito[424].

Altra parte della dottrina riflette sul fatto che tra le caratteristiche di questi sistemi v'è il saper reagire a situazioni previste o, comunque, prevedibili dal programmatore. Se una data situazione non risulta prevedibile neanche per l'operatore umano, il quale non ha impartito al sistema regole e codici per poterla fronteggiare, conseguentemente non sarà possibile per l'IA reagire adeguatamente ad una situazione imprevista e imprevedibile. Precipitato logico di tale argomentare consiste nel seguente interrogativo: «la loro presunta "imprevedibilità" o causalità dell'azione – essendo macchine e non uomini – non è forse, a sua volta, prevedibile da parte del costruttore che può programmare una serie vincolata di scelte?»[425]. Insomma, l'imprevedibilità dell'IA diventa fattore prevedibile per il suo programmatore.

---

[417] U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 624.
[418] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1760.
[419] M.B. MAGRO, *Robot*, cit., p. 1211.
[420] 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*, vol. 19, n. 3, 2016, p. 432. Gli AA. evidenziano che il danno occorso potrebbe essere il risultato di una programmazione negligente e non necessariamente frutto dell'agire autonomo della macchina. Inoltre, finché tali sistemi non potranno essere considerati penalmente responsabili, le potenziali vittime degli eventi lesivi cagionati dall'IA sconterebbero un vuoto di tutela. Non sarebbe infatti possibile ritenere penalmente responsabili né la macchina né le persone dietro di essa, anche per gravi danni causati dall'IA. Alla luce di tali considerazioni gli AA. sconsigliano in generale di assolvere gli operatori dell'IA dalla responsabilità per i danni da questa causati.
[421] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 83.
[422] M.B. MAGRO, *Robot*, cit., p. 1211.
[423] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 19.
[424] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 25 svolge questa considerazione con riferimento all'agire imprevedibile degli animali ma riteniamo che le suddette riflessioni siano valevoli anche con riguardo all'intrinseca imprevedibilità dei più sviluppati sistemi di IA.
[425] M.B. MAGRO, *Biorobotica*, cit., p. 516.

Se, come abbiamo visto, in un contesto di generale imprevedibilità dell'agire dei più sviluppati sistemi di IA, occorre evitare che questo fattore diventi capro espiatorio per deresponsabilizzare i programmatori, di contro occorre evitare (come accennato pocanzi) la creazione di un inammissibile sistema di *responsabilità oggettiva* degli operatori del settore[426]. Tale scenario va tanto più evitato se prendiamo in considerazione la possibilità che gli utilizzatori di un sistema intelligente ne facciano un uso improprio: «fino a che punto gli sviluppatori di macchine intelligenti ed evolute possono essere considerati responsabili per il "duplice e malefico" uso di una tecnologia concepita a scopo benefico ma asservita a scopi bellici o persino criminali? Si potrebbe affermare una responsabilità a carico degli sviluppatori per non aver previsto i potenziali usi alternativi dannosi e persino gli illeciti commessi dalle loro creazioni? È possibile punire penalmente questa cecità, più o meno cosciente, connessa alle conseguenze più remote delle proprie azioni?»[427].

Sarebbe il caso di circoscrivere il regresso della responsabilità cui potrebbero andare incontro gli operatori dell'IA[428]. Una possibile soluzione potrebbe essere quella di modulare la responsabilità degli uomini dietro l'IA in base al diverso grado di autonomia di cui questa è dotata, bilanciandola fra i vari soggetti coinvolti[429]. Non va però mai persa di vista l'importanza di tener fermi i principi garantistici dell'imputazione penale, da cui consegue l'impossibilità di ricorrere a presunzioni e automatismi e la necessità di accertare la sussistenza degli elementi della colpevolezza: «incidendo, infatti, la pena sulla libertà personale, nessuna deroga ai principi di garanzia è tollerabile. Che, di poi, possano presentarsi

---

[426] I. SALVADORI, *Agenti artificiali*, cit., p. 108; U. RUFFOLO, Machina delinquere potest? *Responsabilità ed "illeciti" (anche penali?) della "persona elettronica" e tutele per gli agenti* software *autonomi*, in U. RUFFOLO, *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, cit., p. 305 parla di "responsabilizzazione oggettiva". Nella dottrina classica sull'argomento v. V. CAVALLO, *La responsabilità obbiettiva nel diritto penale*, Napoli, 1937; C. PATERNITI, *La responsabilità obiettiva nel diritto penale. Struttura fondamento prospettive*, Milano, 1978; F. TAGLIARINI, *I delitti aggravati dall'evento*, Padova, 1979; A.M. STILE, *Responsabilità oggettiva e giudizio di colpevolezza*, Napoli, 1989. Sul tema, più in generale, v. anche E. DOLCINI, *L'imputazione dell'evento aggravante, un contributo di diritto comparato*, in *Riv. it. dir. proc. pen.*, 1979, pp. 755 ss.; ID., *Responsabilità oggettiva e principio di colpevolezza. Qualche indicazione per l'interprete in attesa di un nuovo codice penale*, in *Riv. it. dir. proc. pen.*, 2000, pp. 863 ss.; F. MANTOVANI, *Responsabilità oggettiva espressa e responsabilità oggettiva occulta*, in *Riv. it. dir. proc. pen.*, 1981, pp. 456 ss.; A. CASTALDO, *Responsabilità oggettiva e principio di colpevolezza*, in *Riv. it. dir. proc. pen.*, 1988, pp. 1119 ss.; A. PAGLIARO, *Colpevolezza e responsabilità obiettiva: aspetti di politica criminale ed elaborazione dogmatica*, in *Riv. it. dir. proc. pen.*, 1988, pp. 387 ss.; C.F. GROSSO, *Questioni aperte in tema di imputazione del fatto*, in *Riv. it. dir. proc. pen.*, 1993, pp. 21 ss.; ID., *Il principio di colpevolezza nello schema di delega legislativa per l'emanazione di un nuovo codice penale*, in *Cass. pen.*, 1995, pp. 3125 ss. Più di recente sul tema F. BASILE, *La colpa in attività illecita: un'indagine di diritto comparato sul superamento della responsabilità oggettiva*, Milano, 2005; ID., *L'alternativa tra responsabilità oggettiva e colpa in attività illecita per l'imputazione della conseguenza ulteriore non voluta, alla luce della sentenza Ronci delle Sezioni Unite sull'art. 586 c.p.*, in *Riv. it. dir. proc. pen.*, 2011, pp. 911 ss.; ID., *La responsabilità oggettiva nella più recente giurisprudenza della cassazione relativa agli artt. 116, 584 e 586 c.p.*, in *dirittopenalecontemporaneo.it*, 22.11.2012.

[427] M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 19. Menziona il contributo causale di un uso improprio e non conforme del sistema intelligente sul verificarsi dell'evento lesivo anche M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 370.

[428] C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1915, nota 15.

[429] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 344.

obbiettive difficoltà nella individuazione del destinatario del precetto e, quindi, della persona fisica responsabile, laddove si è in presenza di strutture societarie, o più in generale, imprenditoriali complesse, non muta, a nostro avviso, i termini della questione»[430].

Il diritto penale sarà chiamato a valutare se alle nuove esigenze dell'IA possano adattarsi i tradizionali schemi di imputazione della responsabilità ancorati al principio della personalità della responsabilità colpevole, sancito dall'art. 27 co. 1 Cost., o se piuttosto vadano opzionati meccanismi di attribuzione della responsabilità (non propriamente penalistici) fondati sulla causazione oggettiva dell'evento e, pertanto, volti ad ottenere quantomeno un risarcimento del danno[431]. Il diritto penale dovrà inoltre vagliare la possibilità di ricorrere a nuove forme di imputazione della responsabilità, quale la "*colpa di programmazione*" o "*di automazione*" da imputare, principalmente, all'azienda produttrice del sistema intelligente, secondo il modello della *product liability*[432].

Quelle appena elencate sono, a ben vedere, scelte di politica legislativa. Di fronte alle nuove sfide poste dall'innovazione tecnologica, potenzialmente foriere di pericoli, spetterà al legislatore compiere una scelta di tipo "pre-penale", se cioè consentire lo svolgimento di una data attività oppure proibirlo a monte. Il diritto penale potrà, ancora, scegliere se sanzionare la violazione del divieto di svolgere una determinata attività che il legislatore ha scelto di proibire, oppure se punire l'esercizio abusivo di un'attività consentita al ricorre di certe condizioni[433]. In altre parole, si tratterà di decidere la soglia di tollerabilità di questi nuovi rischi che la società sarà disposta a sopportare a fronte dei potenziali benefici derivanti dall'uso dell'IA, così delimitando l'area del rischio consentito.

---

[430] L. STORTONI, *Angoscia tecnologica*, cit., p. 88, l'A. prosegue affermando: «che siffatto problema vada risolto, come ora accade, a livello interpretativo o che, di contro, sia auspicabile che sia il legislatore a determinare i criteri per l'individuazione della persona fisica responsabile all'interno della società e/o dell'impresa, è questione aperta ed ogni soluzione possibile, purché ciò avvenga nel rigoroso rispetto dei canoni dell'art. 27 comma 1 della Costituzione detta in riferimento ai collegamenti – oggettivo e subbiettivo – che debbono sussistere tra fatto e autore».

[431] In dottrina è stata presa in considerazione la proposta, invero avanzata dal Parlamento europeo, di predisporre un regime di assicurazione obbligatoria per i sistemi di IA più sofisticati, insieme alla costituzione di un fondo patrimoniale *ad hoc* preposto a risarcire i danni da essi causati. A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 97; C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 7; C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 105; G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 110; M.B. MAGRO, *Robot,* cit., p. 1204. Parte della dottrina parla di una «"fiscalizzazione" degli oneri risarcitori o indennitari rispetto a specifici eventi pregiudizialmente riconducibili all'operare dell'intelligenza artificiale» U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 16. La realizzazione di un sistema assicurativo *ad hoc* contempererebbe la tutela delle vittime con il non sovraccaricare i produttori di responsabilità, mantenendo vivo il loro interesse ad innovare. «L'attitudine delle imprese assicuratrici a rivalersi nei confronti dei produttori (…) [andrebbe] contenuta entro i casi in cui questi ultimi consegnino al mercato prodotti che non rispettano gli standard di sicurezza necessari». Ciò non serve a frustrare bensì a riconciliare la tutela delle vittime e la possibilità di innovare dei produttori, così M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 371.

[432] Ricostruzione proposta da V. MANES, *L'oracolo algoritmico*, cit., pp. 549-550. Sulla "*colpa di programmazione*" e "*di automazione*" v. anche N. MAZZACUVA, *Alcune riflessioni su intelligenza artificiale e diritto penale sostanziale*, in U. RUFFOLO, *XXVI lezioni di Diritto dell'Intelligenza Artificiale,* cit., p. 290.

[433] L. STORTONI, *Angoscia tecnologica*, cit., p. 83.

13. Il principio di precauzione.

Strumento necessario per effettuare le scelte di politica legislativa pocanzi menzionate – concernenti l'opportunità di consentire o vietare l'esercizio di una certa attività foriera di rischi – dovrebbe essere, a nostro avvio, il principio di precauzione.

Esso infatti, per definizione, opera in contesti di incertezza scientifica i quali, in considerazione dell'incessante progredire dell'evoluzione tecnologica, risultano in un certo senso all'ordine del giorno[434]. Ad avviso di parte della dottrina il principio di precauzione non avrebbe un'efficacia diretta nel diritto penale, essendo piuttosto volto ad ispirare le autorità pubbliche nella scelta concernente l'opportunità di vietare o regolamentare una certa attività potenzialmente pericolosa in ordine alla quale, tuttavia, non sussista ancora alcuna certezza di tipo empirico o scientifico[435]. L'applicazione del metodo precauzionale non è idonea a fondare l'incriminazione di una data condotta. Esso, piuttosto, «svolge il ruolo di criterio guida delle decisioni assunte in condizioni di incertezza, in direzione della prevenzione delle conseguenze peggiori tra tutte le opzioni disponibili e del contenimento del rischio, quando le conoscenze scientifiche non consentano di escludere, ma nemmeno provano, il carattere dannoso dell'attività svolta»[436]. Si tratta, a ben vedere, di un contesto non perfettamente coincidente con le peculiarità del diritto penale il quale, invece, deve fondarsi su dati e fatti scientificamente e giuridicamente accertabili, se non in termini di assoluta certezza, quantomeno con un'elevata probabilità razionale[437].

Occorre chiarire che ad oggi non sembra riscontrabile una definizione omnicomprensiva del concetto di "precauzione": secondo una concezione "radicale" «la regola dell'astensione scatterebbe in presenza di qualunque fattore di rischio potenziale, riguardo al quale la scienza non ha certezza delle conseguenze»; le concezioni "moderate", invece, condividono una nota comune

---

[434] G. FIANDACA, E. MUSCO, *Diritto Penale*, cit., p. 222.

[435] C. BRUSCO, *Rischio e pericolo, rischio consentito e principio di precauzione. La c.d. "flessibilizzazione delle categorie del reato"*, in *Criminalia*, 2012, p. 399. In tal senso si muove anche la dottrina spagnola: G. QUINTERO OLIVARES, *La robótica ante el derecho penal: el vacío de respuesta jurídica a las desviaciones incontroladas*, in *Revista Electronica de Estudios Penales y de la Seguridad*, 2017, pp. 20 ss., l'A. chiarisce che il principio di precauzione possa essere utilizzato per creare regole amministrative volte a proibire l'esercizio di una certa attività, ma non potrebbe del pari essere invocato per legittimare una norma incriminatrice. L'A. infatti delinea la differenza tra reati di pericolo astratto e concreto specificando che, sia nell'uno che nell'altro tipo di reato di pericolo, la possibilità di un risultato dannoso è frutto dell'esperienza, che invece per definizione manca quando riflettiamo sul principio di precauzione. Quando si parla di rischio consentito, infatti, ci si sta riferendo ad una situazione in cui si conoscono le possibili conseguenze di una certa attività, per quanto ad essa non si voglia rinunciare in quanto socialmente utile (es. non è mai stata sollevata la questione sulla sospensione della vendita delle automobili nonostante siano dimostrate le loro potenzialità inquinanti). Quando si invoca, invece, il principio di precauzione ci muoviamo in un terreno in cui non si conoscono le possibili conseguenze lesive di una data attività, in quanto ci troviamo in un contesto di incertezza. L'ignoranza di ciò che può accadere vive a fianco di un'evoluzione che genera rischi prima sconosciuti e che sono ineliminabili nelle moderne società. Sul rapporto tra principio di precauzione e responsabilità colposa v. F. GIUNTA, *Il diritto penale e le suggestioni del principio di precauzione*, in *Criminalia*, 2006, pp. 241 ss.

[436] M.B. MAGRO, *Biorobotica*, cit., pp. 507-508.

[437] C. RUGA RIVA, *Principio di precauzione e diritto penale. Genesi e contenuto della colpa in contesti di incertezza scientifica*, in *Studi in onore di G. Marinucci*, Milano, 2006, p. 1749.

che «consiste nel prevedere, di fronte a una data attività la cui pericolosità è scientificamente incerta, una alternativa al divieto tombale del suo svolgimento»[438].

Prediligere un approccio stringente del principio di precauzione comporterebbe un divieto generalizzato dell'uso dei sistemi intelligenti astrattamente in grado di generare eventi dannosi o pericolosi. Ciò comporterebbe, da un lato, rinunciare ai benefici derivanti dall'uso di questi sistemi[439] e, dall'altro, rallentare considerevolmente lo sviluppo tecnologico[440]. Ci sentiamo pertanto in tal sede di condividere l'opinione di quella parte di dottrina che ritiene che «la condizione di incertezza a riguardo dei possibili effetti negativi dell'impiego di una tecnologia (inclusa l'intelligenza artificiale) non può essere utilizzata come una ragione legittima per non regolare e limitare tale sviluppo»[441]. In altri termini, il timore dei possibili danni derivabili dall'impiego dell'IA non può esimere il legislatore, *in primis*, e i giuristi, *in secundis*, dal regolamentare e interpretare le possibili conseguenze giuridiche della materia.

Piuttosto che porre limiti penalistici alla ricerca scientifica[442] e alla libertà di iniziativa economica, occorrerebbe effettuare un bilanciamento tra pericoli e benefici delimitando l'area del rischio consentito[443] e definendo i rischi "accettabili" in base al tipo di utilità che è possibile trarre dall'impiego dell'intelligenza artificiale[444]. Ciò non escluderebbe la possibilità di attribuire agli operatori dell'IA la responsabilità per gli eventuali danni causati da quest'ultima ove essi siano riconducibili a un omesso adempimento del dovere di monitoraggio[445] oppure dipendano da un'asimmetria informativa tra il programmatore e il soggetto danneggiato. Con riferimento a tale ultimo passaggio la dottrina ha affermato che «l'interesse penalistico del principio di precauzione può risiedere principalmente nella sua attitudine a conferire rilievo al *differenziale di conoscenza*, alla *signoria esclusiva* sulle situazioni pregne di rischio di cui disponga chi intraprenda, specie se in forma organizzata, certe attività produttive aventi un potenziale impatto sui beni giuridici. *Non sarà di per sé dunque la prospettabilità di rischi che potrà generare un dovere all'astensione dalla condotta pericolosa* (e correlativamente un'imputazione di eventi non meglio descritti che se ne possano considerare, problematicamente, la «realizzazione»), *almeno laddove la condotta possa iscriversi nell'alveo di rischi "autorizzati"*, bensì un insieme di elementi più complessi scaturenti dal *confronto tra i livelli di*

---

[438] F. GIUNTA, *Il diritto penale e le suggestioni del principio di precauzione*, cit., pp. 232-233 e 236.

[439] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 19; P. SEVERINO, *Intelligenza artificiale*, cit., p. 536; R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 268; M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 353, ove si mette a confronto un approccio *paternalistico* volto a vietare ogni tecnologia potenzialmente foriera di rischi, e un approccio *liberale* propenso a consentire la diffusione dei sistemi intelligenti indipendentemente dai rischi in essi insiti. Gli stessi AA. più avanti esplicitano l'importanza che la «tutela elevata dei diritti individuali non paralizzi l'evoluzione e il miglioramento dei sistemi intelligenti, che sono forieri di grandi benefici per la collettività» p. 370.

[440] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 70.

[441] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 86.

[442] M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 20.

[443] P. SEVERINO, *Intelligenza artificiale*, cit., p. 536.

[444] G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 9.

[445] G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 9.

conoscenza di cui disponeva l'agente *al momento del fatto* e di quelli a (o messi a) disposizione delle istanze *lato sensu* pubbliche»[446].

Al fine di fronteggiare le problematiche suesposte si potrebbe meditare di imporre un generale divieto di produzione di sistemi intelligenti i cui possibili rischi superino con un certo distacco i potenziali benefici, così limitando l'accesso sul mercato ai sistemi di IA meno evoluti ma maggiormente controllabili[447]. Si potrebbe in tal senso pensare di individuare un meccanismo che imponga il rispetto di alcuni "valori soglia" di automazione, individuando un livello di autonomia oltre il quale al programmatore non è consentito spingersi, in quanto si ridurrebbe eccessivamente il controllo esercitabile dall'uomo su di esso. In questa direzione è stato osservato che «se le misure tecniche non consentono di progettare sistemi completamente sicuri, sarà sempre possibile limitarne l'autonomia e le capacità»[448].

A prescindere dalla via che si riterrà più opportuno percorrere, ciò che appare indubbio è l'indifferibilità di una ragionata regolamentazione dei più sofisticati sistemi intelligenti al fine di tutelare i consociati che con essi verranno in contatto, valutando (a nostro avviso, in *extremis*) l'opportunità di imporre limiti etici e giuridici alle diverse applicazioni della ricerca tecnologica[449]. Difatti, i due principali rischi che il progresso della ricerca scientifica può incontrare sulla sua strada sono:

- uno smodato ed irragionevole ricorso al principio di precauzione (di cui abbiamo appena parlato);
- un eccessivo richiamo al paradigma della responsabilità da prodotto (di cui ci accingiamo a parlare), il quale rischierebbe di disincentivare i produttori dal creare "prodotti intelligenti"[450].

14. La responsabilità per danno da prodotto.

Ricollegandoci al riferimento finale del paragrafo precedente, riteniamo in tal sede opportuno svolgere qualche sintetica riflessione in ordine alla responsabilità per danno da prodotto, in quanto quest'ultima è stata sovente richiamata in materia di danni cagionati dall'intelligenza artificiale.

Come è noto, la materia della responsabilità per danno da prodotto trova il suo terreno d'elezione nel diritto civile[451], ove si predilige un approccio prettamente compensativo. Ad avviso di parte della dottrina le norme attualmente applicabili ai sistemi intelligenti, intesi come "macchine", sono quelle concernenti la responsabilità da prodotto difettoso, le quali attribuiscono la responsabilità per danni al produttore al ricorrere di date circostanze che consentano di imputare in

---

[446] G. FORTI, *"Accesso" alle informazioni sul rischio e responsabilità: una lettura del principio di precauzione*, in *Criminalia*, 2006, p. 195, corsivo nostro.
[447] M.B. MAGRO, *Robot*, cit., p. 1210.
[448] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 359.
[449] P. MORO, *Macchine come noi*, cit., p. 53.
[450] Nonostante tale riflessione sia originariamente sviluppata da U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles e responsabilità nel nostro sistema ed in quello statunitense*, in *Giur. it.*, luglio 2019, p. 1705 con riferimento ai veicoli autonomi, riteniamo che sia spendibile anche in un contesto più generale, prendendo in considerazione i sistemi intelligenti genericamente intesi.
[451] Sul tema E. AL MUREDEN, *La sicurezza dei prodotti e la responsabilità del produttore*, cit.

capo a quest'ultimo la responsabilità per i danni verificatisi[452]. Altri ancora sono dell'avviso che, in realtà, la responsabilità da prodotto difettoso non sarebbe idonea a coprire i danni cagionati dai più evoluti sistemi di IA[453] in considerazione della loro costante evoluzione e della loro capacità di autoapprendimento: ciò, infatti, contribuisce a rendere ancor più difficile non solo la definizione di "difettosità" di un prodotto ma, altresì, a ricostruire il nesso causale intercorrente tra l'uso (o l'agire) del prodotto e il danno verificatosi[454]. Quando ci si riferisce alla capacità dei sistemi intelligenti di agire in modo indipendente si usa parlare dei c.d. "comportamenti emergenti", ossia della possibilità che l'IA ponga in essere condotte non prevedibili dai suoi stessi produttori[455].

La più attenta dottrina civilistica ha individuato le principali norme europee attualmente vigenti che potrebbero intervenire a regolamentare la materia *de qua*, suddividendole in tre cerchi concentrici: «Il *cerchio* che abbiamo definito *interno* è costituito dalla direttiva n. 06/42/CE[456], che disciplina la progettazione e la costruzione delle macchine e che interessa da vicino i robot considerati quali meri artefatti meccanici. Vi è, poi, un *cerchio maggiore*, nel quale si trovano le misure più generali in tema di salute, pubblica sicurezza e tutela dei consumatori: la direttiva n. 01/95/CE, la decisione n. 768/2008/CE e il reg. n. 765/2008/CE, che fissano le regole per la sicurezza dei prodotti all'interno del mercato europeo. Qui il robot è considerato, al pari di qualsiasi altro prodotto, quale possibile fonte di pericolo per la sicurezza pubblica. Infine, vi è un *cerchio esterno*, dove si collocano i diritti e le garanzie riconosciute ai consumatori dalla direttiva n. 99/44/CE sulla vendita dei beni di consumo»[457].

---

[452] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 338, gli stessi AA., poco più avanti, affermano che «in campo civilistico il riferimento più agevole, anche concettualmente, corre alle norme in tema di responsabilità per i vizi della cosa, ossia a quel plesso di regole relative alla *product liability* che distribuiscono la responsabilità tra chi produce il bene e chi ne sperimenta un utilizzo improprio» p. 340; si interroga sull'opportunità di estendere le norme sulla responsabilità da prodotto difettoso alle entità intelligenti anche U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., pp. 21-25.

[453] E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale,* cit., p. 86.

[454] U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 624.

[455] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 83; sul punto v. anche R. CALO, *Robotics and the Lessons of Cyberlaw*, in *California Law Review*, 2015, p. 539 il quale afferma che un «emergent behavior can lead to solutions no human would have come to on her own».

[456] Trattasi della c.d. "Nuova direttiva Macchine", preceduta dalla Direttiva 98/37/CE e, prima ancora, dalla Direttiva 89/392/CEE, anche nota come "Prima Direttiva Macchine", A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 71.

[457] A. SANTOSUOSSO, C. BOSCARATO, F. CAROLEO, *Robot e diritto: una prima ricognizione*, in *Nuova Giurisprudenza Civile Commentata*, 2012, p. 8. Come si evince dal corpo del testo gli AA. si riferiscono nello specifico ai robot ma riteniamo le suesposte considerazioni applicabili anche ai sistemi di IA generalmente intesi. Si concentrano invece sul ruolo della Direttiva 85/374/CEE, chiarendo come essa «copre solamente i danni causati dai difetti di fabbricazione di un robot e a condizione che la persona danneggiata sia in grado di dimostrare il danno effettivo», M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 346. Sulla necessità di adeguamento in ottica attualizzante della Direttiva 85/374/CEE v. F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 245. Necessità di adeguamento avvertita e segnalata anche dall'Unione Europea nella *Relazione della Commissione al Parlamento europeo, al Consiglio e al Comitato economico e sociale europeo sull'applicazione della direttiva del Consiglio relativa al ravvicinamento delle disposizioni legislative, regolamentari ed amministrative degli stati membri*

V'è chi ritiene che, per quanto già sufficientemente completa, la normativa in materia di responsabilità per danno da prodotto – generalmente intesa – potrebbe essere la più idonea a subire un'integrazione normativa *ad hoc* per meglio adeguarsi alle peculiarità di questi "nuovi" prodotti[458]. È stato in tal senso proposto di introdurre una norma che imponga ai produttori dei sistemi intelligenti di inserire una "scatola nera"[459] nell'hardware, al fine di accertare più facilmente la causa (o le cause) di un errore o di un malfunzionamento[460].

Il tema della responsabilità per danno da prodotto è stato esplorato anche dal punto di vista del diritto penale[461] seppur, con le sue peculiarità e i suoi "limiti"[462]. La dottrina che si è occupata del tema ha evidenziato diverse problematiche, con riguardo tanto all'*elemento oggettivo* quanto all'elemento soggettivo del reato. Sotto il primo profilo l'aspetto più spinoso che si riscontra in materia concerne l'accertamento eziologico[463]. Stante la potenziale "diffusività" del danno da prodotto, nonché la consistente difficoltà di individuare il fattore produttivo del danno[464], la giurisprudenza sembra essersi accontentata di un'"imputazione causale alternativa", fondata da un lato sulla vicinanza temporale tra l'uso del prodotto e la verificazione dell'evento dannoso e, dall'altro, sull'assenza di fattori

---

in materia di responsabilità per danno da prodotti difettosi (direttiva 85/374/CEE), Bruxelles, 7.5.2018 COM(2018) 246 final.

[458] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 25.

[459] A. TURANO, *Robotica e roboetica*, cit., p. 156 parla di "*ethical black box*".

[460] C. SALAZAR, *Umano, troppo umano*, cit., p. 261, la quale attribuisce la proposta riportata nel corpo del testo agli studiosi del progetto RoboLaw. Si tratta del progetto "*Regulating Emerging Robotic Technologies in Europe: Robotics facing Law and Ethics*", finanziato dalla Commissione Europea, avviato nel marzo 2012 e conclusosi nel maggio 2014 volto a considerare le conseguenze etiche e giuridiche delle nuove tecnologie robotiche. Sul punto v. anche A. TURANO, *Robotica e roboetica*, cit., p. 134.

[461] Oltre all'autorevole dottrina citata nel prosieguo sia consentito un rimando anche a A. DI MARTINO, *Danno e rischio da prodotti. Appunti per la rilettura critica di un'esperienza giurisprudenziale italiana*, in R. BARTOLI, *Responsabilità penale e rischio nelle attività mediche e d'impresa*, cit., pp. 437 ss.; M. DONINI, D. CASTRONUOVO (a cura di), *La riforma dei reati contro la salute pubblica. Sicurezza del lavoro, sicurezza alimentare, sicurezza dei prodotti*, cit.; A. BERNARDI, *La responsabilità da prodotto nel sistema italiano: profili sanzionatori*, in *Riv. trim. dir. pen. econ*., 2003, pp. 1 ss.; F. BRICOLA, *Responsabilità penale per il tipo e per il modo di produzione*, in AA.VV., *La responsabilità dell'impresa per i danni all'ambiente e ai consumatori*, Milano, 1978, pp. 75 ss.

[462] Invero i punti di contatto in materia tra diritto civile e diritto penale non mancano. Basti pensare al disposto dell'art. 112 del Codice del Consumo il quale prevede le seguenti fattispecie contravvenzionali: «Salvo che il fatto costituisca più grave reato, il produttore o il distributore che *immette sul mercato prodotti pericolosi in violazione del divieto di cui all'articolo 107, comma 2, lettera e)*, e punito con l'arresto da sei mesi ad un anno e con l'ammenda da 10.000 euro a 50.000 euro. 2. Salvo che il fatto costituisca più grave reato, il produttore che *immette sul mercato prodotti pericolosi*, è punito con l'arresto fino ad un anno e con l'ammenda da 10.000 euro a 50.000 euro. 3. Salvo che il fatto costituisca più grave reato, il produttore o il distributore che *non ottempera ai provvedimenti emanati a norma dell'articolo 107*, comma 2, lettere b), numeri 1) e 2), c) e d), numeri 1) e 2), e punito con l'ammenda da 10.000 euro a 25.000 euro», corsivo nostro. Sul tema v. F. CONSULICH, *Tutela del consumatore* (voce), in F. PALAZZO, C.E. PALIERO (a cura di), *Commentario breve alle leggi penali complementari*, Padova, 2007, pp. 2967 ss.

[463] Sulla difficoltà di ricostruire l'eziologia del danno causato da prodotti difettosi e sulla difficoltà di individuare i soggetti responsabili v. I. SALVADORI, *Agenti artificiali*, cit., p. 105.

[464] Riflette sull'"eziologia del difetto", distinguendo le seguenti tipologie di danni: difetti di costruzione, difetti di fabbricazione, difetti da informazione e difetti da rischio di sviluppo C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1754.

causali alternativi (c.d. "causalità negativa"). In questi contesti, infatti, l'unica certezza sembra essere la causale riconducibilità del danno al prodotto, da cui però non deriva altrettanta certezza nell'individuazione del nesso causale e, conseguentemente, del soggetto responsabile[465]. Tale autorevole dottrina usa, a tal proposito, parlare di "trame causali oscure"[466], che diventano ancor più buie quando la ricostruzione del nesso causale viene resa sostanzialmente impossibile dall'autonomia e dal *black box effect* dei sistemi intelligenti[467].

Quanto al secondo profilo, ossia quello concernente l'*elemento soggettivo*, l'illustre dottrina riscontra la seria difficoltà di formulare un giudizio di prevedibilità ed evitabilità dell'evento dannoso, proprio in ragione della complessità, a monte, di determinare la pericolosità del prodotto. Si registra in tal senso un orientamento ondivago della giurisprudenza, volto a riconoscere, ad un tempo, la sussistenza del dolo eventuale e, in altro tempo, della colpa cosciente. I contorni tra i due elementi soggettivi si fanno in tal sede ancor più sfumati, dipendendo da quanto il giudice reputi elevato il livello di cognizione della situazione da parte dei produttori[468].

Si tratta, a ben vedere, di un settore connotato da un generale stato di incertezza, che non consente di formulare regole di condotta volte ad eliminare, con probabilità vicina alla certezza, i possibili rischi derivanti dall'uso dei prodotti, specie se parliamo di sistemi intelligenti dotati di un margine di autonomia. Tale ragionamento sembra evocare una nuova forma di "rischio di sviluppo" che emerge tra la creazione dell'algoritmo e l'autoapprendimento realizzato dall'IA[469], potendo quest'ultima sfuggire al controllo del suo creatore. L'unica alternativa capace di azzerare totalmente il rischio di verificazione di eventi dannosi sarebbe (riprendendo le considerazioni svolte nel paragrafo precedente) quella di astenersi del tutto da ogni attività potenzialmente pericolosa, prediligendo il "*vivere necesse*" al "*navigare necesse*"[470].

Altro aspetto che viene in rilievo in questa materia attiene alla *struttura della condotta* in quanto si riscontra una sorta di «interscambiabilità strutturale tra *fare*

---

[465] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1755.

[466] La suesposta ricostruzione è riferibile a C. PIERGALLINI, *La responsabilità del produttore: una nuova frontiera del diritto penale?*, in *Dir. pen. e proc.*, 2007, pp. 1125 ss., in cui l'A. propone un'efficace sintesi di quanto più approfonditamente esposto in C. PIERGALLINI, *Danno da prodotto e responsabilità penale. Profili dommatici e politico-criminali*, cit. Dello stesso A. si vedano, sul medesimo tema, anche C. PIERGALLINI, *La responsabilità del produttore: avamposto o Sackgasse del diritto penale*, in *Riv. it. dir. proc. pen.*, 1996, pp. 354 ss.; C. PIERGALLINI, *Danno da prodotto e responsabilità penale*, in *Studium iuris*, 2006, pp. 299 ss.; C. PIERGALLINI, *Attività produttive, decisioni in stato di incertezza e diritto penale*, in M. DONINI, M. PAVARINI (a cura di), *Sicurezza e diritto penale*, Bologna, 2011, pp. 327 ss.

[467] Su cui ci siamo soffermati nel Cap. II, Sez. I, Par. 5.1. È suggestivo pensare come parlasse già di "effetto *black box*" a proposito della responsabilità del produttore C.E. PALIERO, *L'autunno del patriarca. Rinnovamento o trasmutazione del diritto penale dei codici*, in *Riv. it. dir. proc. pen.*, 1994, p. 1240.

[468] C. PIERGALLINI, *La responsabilità del produttore,* cit., p.1127.

[469] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 21. Tradizionalmente, quando si parla di "rischio di sviluppo" ci si riferisce alla «possibilità di esentare da responsabilità il produttore qualora lo stato delle conoscenze scientifiche e tecniche al momento della messa in commercio del prodotto non consentissero di rilevarne la difettosità» A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 92.

[470] C. PIERGALLINI, *La responsabilità del produttore,* cit., p. 1129.

e *omettere*»[471]. Sembra quasi che ci si dimentichi delle peculiarità della condotta omissiva e dell'impossibilità di assimilarla in toto a quella commissiva: basti pensare alla necessità che, nella prima ipotesi, sussista una posizione di garanzia con relativo obbligo giuridico di impedire l'evento. La più attenta dottrina riscontra tuttavia come la frammentazione dei centri decisionali tipica dei complessi produttivi, unita alla difficoltà di riconoscere la pericolosità di un prodotto e alla scarsa trasparenza delle situazioni di rischio, renda arduo individuare e delineare le suddette posizioni di garanzia: «il garante risponde non per il mancato controllo di una ben individuata fonte di pericolo, ma solo per l'*aumento* del (generale) *rischio* della produzione»[472].

Il danno da prodotto, inoltre, usa collocarsi nell'ambito di contesti produttivi che si svolgono all'interno di *organizzazioni complesse*, creando così una sorta di "spersonalizzazione" dei processi produttivi[473]. Si palesano su questo terreno le difficoltà del diritto penale ad individuare le singole responsabilità all'interno di apparati articolati rispettando il principio di personalità della responsabilità penale ex art. 27 co. 1 Cost.[474]. Si tratta di un problema di responsabilità plurisoggettiva che deriva da una «frammentazione dei centri decisionali e del processo di formazione della volontà sociale»[475]. In questo contesto distinguere i "ruoli concorsuali" diventa compito quantomai complesso, il che ha forse giustificato un ripiegamento verso un'imputazione della responsabilità che, piuttosto che andare a ricercare il soggetto più vicino al fatto tipico, si collocasse direttamente agli apici decisionali[476] (si parla a tal proposito di *channeling* della responsabilità)[477]. Parte della dottrina ha proposto, con particolare riferimento ai sistemi di IA, di individuare, in seno a un'organizzazione complessa, una rete di garanti deputati a controllare determinate fonti di pericolo[478], istituendo così vere e proprie posizioni di garanzia non solo in capo agli apici aziendali (i quali detengono il potere decisionale) ma anche ai soggetti subordinati[479] i quali, però, detengono il sapere tecnico-scientifico necessario per sviluppare l'IA[480].

Rischiamo però, anche qui, di prendere la china scivolosa di una responsabilità oggettiva "mascherata". Prediligere un meccanismo di imputazione

---

[471] C.E. PALIERO, *L'autunno del patriarca,* cit., p. 1241.
[472] C.E. PALIERO, *L'autunno del patriarca,* cit., p. 1241.
[473] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 96.
[474] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1755.
[475] C. PIERGALLINI, *La responsabilità del produttore: avamposto o Sackgasse del diritto penale*, cit., p. 361.
[476] C.E. PALIERO, *L'autunno del patriarca,* cit., p. 1242.
[477] C. PIERGALLINI, *La responsabilità del produttore: avamposto o Sackgasse del diritto penale*, cit., p. 369, ossia «l'idea di *concentrare* la responsabilità su *uno solo* dei diversi soggetti intervenuti nel processo produttivo, quale componente del vertice aziendale della ditta venditrice».
[478] «È un bisogno ineluttabile della moderna vita sociale che a certi soggetti venga affidato il compito di prevenire i pericoli che minacciano certi beni, che essi divengano "istanze di protezione" di tali beni, "garanti" dell'impedimento dei possibili eventi lesivi. Ciò appare come il correlato della – totale o parziale – incapacità dei titolari degli interessi in questione di proteggerli adeguatamente. [Si tratta di] un fenomeno destinato ad accentuarsi nella vita moderna, caratterizzata da una crescente tecnicizzazione» G. GRASSO, *Il reato omissivo improprio*, cit., p. 166.
[479] Sul ruolo del manufattore negli errori di montaggio o del programmatore per errori di programmazione che si potevano (e che si dovevano) prevedere, sempre nell'ambito della responsabilità da prodotto, v. A. CAPPELLINI, Machina delinquere non potest?, cit., p. 9.
[480] I. SALVADORI, *Agenti artificiali*, cit., p. 105.

di tipo normativo, volto non tanto alla ricostruzione del nesso causale in senso naturalistico quanto, piuttosto, ad accertare l'avvenuta trasgressione di un dovere impeditivo del fatto lesivo o, in altri termini, l'inosservanza di una regola di condotta a contenuto impeditivo, potrebbe meglio fronteggiare l'anonimizzazione del danno cui stiamo assistendo[481]. La tendenza del danno a diventare "anonimo", dunque la difficoltà di individuare il responsabile dell'evento lesivo occorso, in realtà, costituisce aspetto tipico dei danni verificatisi in contesti di progresso industriale e scientifico[482]. Abbiamo però avuto modo di evidenziare nei passaggi precedenti quanto le peculiarità dell'IA siano idonee a rendere complessa l'individuazione del soggetto deputato a ricoprire una posizione di garanzia in termini di controllo sulla fonte di pericolo e la perimetrazione delle regole cautelari da osservare, non solo in considerazione dell'elevato grado di incertezza – anche scientifica – che caratterizza tali sistemi, ma anche in ragione dell'autonomia che i prodotti intelligenti sono in grado di dimostrare, financo al punto di annullare il concreto potere impeditivo dell'evento da parte del garante[483].

Ad ogni modo, al produttore spetterà l'obbligo di servirsi soltanto di prodotti dotati di un verificato grado di sicurezza, non solo preliminarmente alla loro immissione sul mercato, ma anche successivamente, raccogliendo i *feedback* dei consumatori e monitorando i prodotti in modo da poter intervenire con le opportune cautele in caso di eventi dannosi o pericolosi. L'operatore inattivo potrebbe infatti essere chiamato a rispondere penalmente per aver immesso sul mercato un prodotto pericoloso violando il dovere di monitoraggio dei suoi possibili effetti dannosi. Questi ultimi, infatti, pur non essendo stati identificati in fase di programmazione, potrebbero emergere successivamente una volta immesso ed utilizzato il prodotto in un ambiente aperto[484]. Occorrerebbe, anche in tal sede, effettuare un'opera di bilanciamento, limitando la circolazione di alcuni dispositivi che non abbiano ancora raggiunto un accettabile standard di sicurezza, incoraggiando invece l'immissione sul mercato di altri prodotti, cercando «un equilibrio tra incentivo alla sicurezza *ex ante* e garanzia di tutela per le vittime del danno *ex post*»[485].

Se, come sostiene parte della dottrina, «il diritto penale "classico" (di evento di danno o di pericolo) non è capace di "contenere" questa fenomenologia di danno, se non al prezzo di una dissolvenza che sfocia in una inammissibile mutazione genetica»[486], tale inciso diventa ancor più vero quando il dominio

[481] G. FIANDACA, *Il reato commissivo mediante omissione*, cit., pp. 60-61.
[482] F. SGUBBI, *Responsabilità penale per omesso impedimento dell'evento*, cit., p. 73.
[483] Sull'implementazione dei danni anonimi e sulla conseguenza che l'imprenditore sia chiamato a risponderne anche in assenza di colpa, quasi come fisiologico costo della sua attività d'impresa, v. A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 96.
[484] M.B. MAGRO, *Robot*, cit., pp. 1209 ss. Per un'analisi comparatistica degli obblighi del produttore in Germania e negli Stati Uniti v. S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., pp. 426 ss.
[485] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 358.
[486] C. PIERGALLINI, *La responsabilità del produttore*, cit., p. 1128. In senso analogo si era già espresso C.E. PALIERO, *L'autunno del patriarca,* cit., pp. 1239-1240 affermando che esistono «fenomenologie criminose che appaiono – nella loro eziologia e nelle loro modalità esecutive – talmente *complesse* da non poter essere dominate dai modelli tradizionali di responsabilità penale: diciamo, per intenderci, dai *titoli* di responsabilità penale (individuale e collettiva) che sono familiari all'esperienza codicistica».

dell'uomo sulla macchina entra in crisi. Per quanto, a nostro avviso, ciò non esima i giuristi dal provare a ragionare compiutamente sulla portata operativa degli strumenti penalistici con riguardo agli eventi lesivi cagionati dall'IA[487]. Torna attuale, anche in tal sede, la necessità di interrogarsi sul "grado di rischio" che la società sarà disposta a sopportare – facendolo assurgere al rango di "rischio consentito" – trattandosi di attività socialmente utili, per quanto foriere di rischi e pericoli. Con riferimento ai sistemi intelligenti si tratterà, pertanto, di effettuare una scelta politica proattiva, volta non tanto a individuare i centri di responsabilità quanto, piuttosto, a «decidere *come governare, preventivamente, questi rischi*, interrogandosi sul livello di tollerabilità sociale»[488]. Una volta effettuata tale scelta di campo e fissate le regole cautelari da osservare, l'eventuale rischio residuale che permanga nonostante il rispetto di queste regole da parte dei produttori sarebbe a carico dei consumatori[489], purché chiaramente da essi "socialmente accettato".

SEZIONE TERZA
INTELLIGENZA ARTIFICIALE: STRUMENTO, AUTORE E VITTIMA DEL REATO

15. Premessa.

Abbiamo adesso gli strumenti per poter analizzare la materia del rapporto tra diritto penale e intelligenza artificiale sotto una doppia visuale prospettica.

Accogliendo la tesi *strutturalista*, propria dell'intelligenza artificiale "debole", dovremmo dedurre che essa non possa essere intesa se non come mero strumento di reato nelle mani dell'uomo.

Optando, invece, per la tesi *funzionalistica*, tipica dell'intelligenza artificiale "forte", i sistemi intelligenti andrebbero intesi quali autori in prima "persona" dei reati da essi commessi. A questo punto, volendo attribuire una reale soggettività all'IA – una vera e propria personalità giuridica – nulla escluderebbe che essa possa essere considerata non solo come autore ma anche come vittima di reato.

Queste "macchine" devono essere intese come oggetti o come agenti[490]? La sanzione penale sarà indirizzata agli operatori "dietro la macchina" o alla macchina stessa? Essa, ancora, sarà da considerarsi mero strumento o nuovo soggetto di diritto? Essa è agita o agisce[491]?

Ci occuperemo, preliminarmente, dell'uso dell'IA come "strumento" di diritto penale, per poi concentrarci sulla possibilità di considerare i sistemi intelligenti, in prospettiva evolutiva, come "autori" di possibili fatti criminosi.

---

[487] In particolare, in una materia come quella in esame, bisognerà prestare particolare attenzione. La summenzionata diffusività del danno potrebbe richiamare a sua volta la diffusività del pericolo la quale, per canto suo, rischia di "volatilizzare" il bene giuridico protetto, svalutandone le esigenze di garanzia. Nel ripensare i tradizionali istituti si dovrà evitare di fare un "uso estensivo del bene giuridico", così scavalcando la tipicità della norma penale volta a salvaguardarlo. «Un diritto penale "sostanziale" così "elastico" si presta ad un uso "processuale" fortemente pericoloso per i diritti di libertà dei cittadini» L. STORTONI, *Angoscia tecnologica*, cit., pp. 74 ss., per la citazione p. 81.

[488] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1750.

[489] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., p. 429.

[490] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 337.

[491] M.B. MAGRO, *Biorobotica*, cit., p. 513.

16. Intelligenza artificiale come "strumento" di reato.

Vorremmo in tal sede riferirci a quei sistemi di IA in grado di porre in essere comportamenti *a priori* predeterminati, predeterminabili e prevedibili[492]. È in tale contesto che il diritto penale si trova maggiormente a suo agio, quello cioè in cui – accogliendo una concezione "debole" di intelligenza artificiale – quest'ultima operi come "ausilio" dell'uomo, senza andare a considerare la capacità dell'IA di delinquere o di essere destinataria di un trattamento sanzionatorio[493]. Qui l'intelligenza artificiale viene intesa come prolungamento dell'uomo, come sua *longa manus*[494]. I sistemi dotati di intelligenza artificiale, dunque, possono essere programmati[495] e utilizzati per commettere fatti illeciti[496], assumendo dunque una "funzione servente" rispetto all'autore del reato[497].

La dottrina, cercando di tracciare una linea di confine tra "mezzi" e "agenti", ha affermato che «è un *mezzo* quell'elemento in un corso di azioni che contribuisce a causare un evento, ma che non esercita alcuna autonomia o discrezione. È invece un *soggetto agente* colui che avvia tale corso di azioni *causando* l'evento»[498]. La medesima dottrina ha inoltre chiarito come la tecnologia, tradizionalmente, sia stata annoverata nella categoria dei "mezzi" e che, pertanto, il responsabile di un evento è sempre stato individuato in colui il quale aveva "deciso" il corso delle azioni che, infine, avevano portato all'evento[499]. Riteniamo che tale ricostruzione possa rivelarsi valevole quando si prendono in considerazione sistemi intelligenti su cui persiste il controllo dell'uomo.

Lo strumento di un reato esiste "al di fuori" del soggetto che se ne serve. Parlare di sistemi intelligenti quali strumenti di reato presuppone riferirsi a una categoria di sistemi "dipendenti" dall'uomo, incapaci di realizzare azioni autonome. In tal senso, anche trovandoci in presenza di sistemi in grado di porre in essere un'azione (intesa da un punto di vista empirico e non giuridico)[500], dovremo sempre considerarli come meri strumenti utilizzati dall'uomo, con la conseguente imputazione a quest'ultimo delle azioni criminose realizzate dai (*rectius* mediante) suddetti sistemi[501].

Considerando i sistemi intelligenti come semplici strumenti nelle mani dell'uomo sarà conseguentemente possibile sottoporli a confisca e sequestro[502].

---

[492] I. SALVADORI, *Agenti artificiali,* cit., p. 91; R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 264; A. CAPPELLINI, Machina delinquere non potest?, cit., p. 5.
[493] M. DI FLORIO, *Il diritto penale che verrà*, cit., p. 12.
[494] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1751; M.B. MAGRO, *Biorobotica,* cit., p. 514; EAD., *Robot*, cit., p. 1207.
[495] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., p. 425.
[496] I. SALVADORI, *Agenti artificiali,* cit., p. 99.
[497] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1746.
[498] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 67.
[499] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 68, l'A. prosegue affermando efficacemente che, fino a pochi anni fa «nessuno avrebbe mai pensato di ritenere responsabile della morte di un uomo un *autoveicolo* perché "tecnicamente" è stato il mezzo che, colpendo la persona, ne ha causato l'evento morte. Ebbene oggi, dinanzi ad auto senza guidatore o ad armi a guida autonoma, questioni del genere stanno diventando molto più complesse».
[500] La specificazione è d'obbligo, come verrà meglio chiarito più avanti, Cap. II, Sez. III, Par. 17.4.
[501] S. RIONDATO, *Robotica e diritto penale*, cit., pp. 600-601.
[502] S. RIONDATO, *Robot: talune implicazioni di diritto penale*, cit., p. 89; ID. S. RIONDATO, *Robotica e diritto penale*, cit., p. 601.

Inoltre, in tal senso, l'uso dell'IA per la commissione di un fatto di reato potrebbe essere presa in considerazione da parte del giudice in fase di determinazione del trattamento edittale ex art. 133 c.p. il quale, al comma 1, n. 1, specifica che i "mezzi" del reato possono essere utilizzati per desumere la gravità dell'illecito commesso[503]. L'aspetto più interessante concerne la possibilità che l'uso dell'IA a fini criminosi vada ad integrare elemento circostanziale del reato. Se, come abbiamo accennato in precedenza[504], l'avvento delle nuove tecnologie ha portato all'introduzione di specifiche circostanze aggravanti per reati commessi con l'uso di sistemi telematici[505], è ragionevole ritenere che anche l'impiego dell'IA possa, *de iure condendo*, assurgere al rango di circostanza del reato. L'attuale aggravante degli atti persecutori commessi mediante l'uso di strumenti informatici o telematici potrebbe trovare nuova linfa nell'uso dei *social bot*[506] o degli *stalkerware* e l'aggravante del "mezzo insidioso" nel reato di omicidio potrebbe essere integrata dall'uso di droni armati a pilotaggio remoto, in quanto mezzi dai quali è sostanzialmente impossibile difendersi[507]. Occorre però tenere ben presente che, in considerazione del continuo progredire delle tecnologie intelligenti e dell'attuale assenza di definizioni unanimi nel settore, strutturare una norma penale che aggravi il trattamento sanzionatorio, ove il reato sia stato commesso mediante un sistema di IA, non sarà compito semplice. Spetterà al legislatore normare in punta di piedi, evitando di porsi in contrasto con il principio di tassatività o sufficiente determinatezza.

Alla luce delle appena svolte considerazioni ci sembra possibile affermare che, finché permanga sull'agire dell'IA un significativo controllo umano e finché essa sia in grado di porre in essere solo comportamenti predeterminati e prevedibili, possiamo affermare che sussista un'«ultrattività delle categorie penalistiche»[508]. Considerando l'IA come mero strumento nelle mani dell'uomo, in tutto e per tutto prevedibile e controllabile, le suesposte perplessità in ordine all'applicazione dei canoni classici del diritto penale tenderebbero a dissolversi, in quanto il fatto criminoso resterebbe oggettivamente e soggettivamente riconducibile al soggetto agente.

16.1. L'imputazione della responsabilità per reato commesso "mediante" l'IA.

Quindi, se quanto affermato nel paragrafo precedente è vero, quando l'IA viene in considerazione come semplice "mezzo" nelle mani dell'uomo la responsabilità per gli eventi dannosi occorsi sarà a lui causalmente riconducibile. Chiaro è che, nel contesto che stiamo cercando di delineare, il controllo dell'uomo

---

[503] S. RIONDATO, *Robotica e diritto penale*, cit., p. 601.
[504] Cap. II, Sez. II, Par. 10.
[505] L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., p. 78.
[506] I socialbot (noti anche come chatbot) sono sistemi software creati per dialogare con i frequentatori delle piattaforme social. Essi, però, possono anche diventare strumenti di molestie. Paradigmatica la vicenda del software chatbot di Twitter (denominato Tay) il quale, dopo aver interagito con gli utenti della piattaforma, ha iniziato a produrre commenti razzisti e xenofobi, R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 263; M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 4; I. SALVADORI, *Agenti artificiali*, cit., p. 109, nota 82. Tale ultimo A. si occupa anche dell'uso dei social media bots per diffondere l'hate speech e le fake news in rete, p. 87.
[507] Esempi proposti da I. SALVADORI, *Agenti artificiali*, cit., pp. 109-110.
[508] C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 135.

sul sistema intelligente rappresenta il presupposto fondamentale per potergli imputare i danni cagionati dall'IA medesima. Ciò è dimostrato, *a contrario*, dal fatto che nei sistemi maggiormente evoluti questo controllo inizia a vacillare e, con esso, i relativi meccanismi di imputazione della responsabilità penale[509]. Il controllo umano assume il ruolo non solo di "catalizzatore della responsabilità" per gli eventi dannosi cagionati dall'IA ma anche di "meccanismo di salvaguardia" volto a impedire che l'errato funzionamento del sistema intelligente possa cagionare danno a terzi[510].

Quindi, finché l'IA viene utilizzata in modo consapevole e volontario (dunque *dolosamente*) al fine di commettere un delitto *nulla quaestio*: l'agire della macchina coincide con l'agire dell'uomo che se ne serve. Semmai potrebbero sorgere dubbi afferenti all'eventuale esigenza di rimodellare fattispecie di reato già esistenti o di crearne delle nuove per renderle maggiormente confacenti con le tipicità dello strumento intelligente[511], senza andare ad intaccare i classici meccanismi di imputazione della responsabilità. Considerando dunque l'IA come strumento, nulla esclude che nel corso dell'agire criminoso del soggetto agente possa verificarsi un errore nell'uso dei mezzi d'esecuzione. La dottrina propone l'esempio del terrorista che, servendosi di un drone armato, intenda uccidere un uomo mentre si trova nel suo domicilio e finisca per uccidere erroneamente il vicino. Ci troveremmo davanti ad un'ipotesi di *aberratio ictus* ex art. 82 co. 1 c.p. Altro esempio proposto dalla medesima dottrina prende in considerazione l'eventualità che un investigatore privato si serva di un drone per riprendere illecitamente delle immagini. Nel caso in cui il drone precipitasse, ferendo una persona, saremmo in presenza di un'ipotesi di *aberratio delicti* ex art. 83 co. 1 c.p.[512].

Invero, in linea di principio, nulla osta alla possibilità di riconoscere la sussistenza anche della responsabilità *colposa* in capo all'utilizzatore dell'IA il quale ne faccia un uso improprio e negligente e, dal quale, derivi un danno prevedibile ed evitabile. Un problema si potrebbe porre, al più, nel caso in cui il sistema intelligente cagioni un danno a causa di un malfunzionamento. Attribuire la responsabilità al suo utilizzatore sarebbe più complesso, a meno di non volergli imputare una mancata manutenzione che sia stata causa dell'evento. Nel caso in cui però non possa essere attribuita alcuna responsabilità all'utente dell'IA il quale, anzi, aveva fatto affidamento sul buono stato del prodotto, si dovrebbe

[509] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., pp. 356-357, nonostante gli AA. svolgano tale considerazione con riferimento ai meccanismi di imputazione della responsabilità civile, riteniamo il ragionamento traslabile anche in materia penale. Difatti la medesima considerazione è riportata anche dalla dottrina penalistica, in particolare F. BASILE, *Intelligenza artificiale e diritto penale*, cit., pp. 27-28.

[510] Sul duplice ruolo del controllo umano o, meglio, dell'"orientabilità umana" si concentra C. PIERGALLINI, *Intelligenza artificiale,* cit., pp. 1757-1758.

[511] M.B. MAGRO, *Robot,* cit., pp. 1206-1207; G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 9 il quale a sua volta richiama F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 27; U. PAGALLO, *Saggio sui robot e il diritto penale,* cit., p. 607.

[512] Tali esempi sono formulati da I. SALVADORI, *Agenti artificiali*, cit., pp. 110-111, nota 87. Parte della dottrina propone, in prospettiva futuribile, proprio al fine di evitare ogni possibile scenario in cui l'IA venga usata per scopi malevoli, di programmare i sistemi intelligenti per non eseguire comandi provenienti dall'uomo che siano errati, A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 90.

andare a cercare il responsabile all'interno della filiera produttiva del sistema[513] secondo il meccanismo della responsabilità da prodotto difettoso sinteticamente esposta in precedenza[514].

16.2. Un problema di "autoria mediata"?

I software più tradizionali e automatici di cui ci stiamo occupando si caratterizzano per l'assenza di autonomia: il loro comportamento è in tutto e per tutto predeterminato e prevedibile da parte del programmatore. Parte della dottrina ha dunque proposto di servirsi di un classico meccanismo di imputazione penale: quello della responsabilità vicaria o indiretta. Essendo il sistema intelligente mero strumento nelle mani dell'uomo quest'ultimo verrebbe chiamato a rispondere sia a titolo doloso, in caso di uso intenzionale dell'IA per scopi criminosi, sia a titolo colposo, nel caso in cui l'evento lesivo occorso fosse prevedibile ed evitabile[515]. Si ritiene che, per quanto possa essere complessa l'attività realizzata dall'IA, essa vada imputata (secondo il meccanismo della responsabilità indiretta) sia oggettivamente che soggettivamente a colui il quale se ne sia servito direttamente o che l'abbia programmata per svolgere un certo compito. Ciò può avvenire, ad esempio, nel caso in cui un militare programmi il drone armato per sparare sui civili, sia nel caso in cui il drone uccida i civili per un difetto di programmazione che si sarebbe potuto evitare. Nel primo caso saremo davanti a un'ipotesi di imputazione per dolo, nel secondo caso invece si tratterà di un'imputazione per colpa[516].

Parte della dottrina che si è occupata del tema ha riscontrato come, in realtà, la contrapposizione dei modelli di imputazione diretta dell'IA e indiretta dell'uomo sia meramente astratta, in quanto non sarebbe ad oggi ipotizzabile una responsabilità diretta della macchina[517]. Considerando quest'ultima come strumento nelle mani dell'uomo sarebbe possibile imputargli la responsabilità per l'"agire dell'IA", senza scomodare l'autoria mediata[518].

---

[513] S. CHIARLONI, *Riflessioni minime*, cit., p. 9.
[514] Cap. II, Sez. II, Par. 14.
[515] P. SEVERINO, *Intelligenza artificiale*, cit., p. 533.
[516] R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 264.
[517] A. CAPPELLINI, Machina delinquere non potest?, cit., pp. 3-4.
[518] Nella dottrina classica sul tema delle peculiari ipotesi di concorso di persone nel reato (nel quale in genere si colloca il tema dell'autoria mediata) v. C. PEDRAZZI, *Il concorso di persone nel reato*, Palermo, 1952; per una rinnovata lettura di tale ultimo Autore, L. RISICATO, *Rileggendo Cesare Pedrazzi*, Il concorso di persone nel reato, *Palermo, 1952*, in *Criminalia*, 2020, pp. 37 ss., ora anche in *disCrimen*, 23.2.2021; R. DELL'ANDRO, *La fattispecie plurisoggettiva in diritto penale*, Milano, 1956; M. GALLO, *Lineamenti di una teoria sul concorso di persone nel reato*, Milano, 1957; A.R. LATAGLIATA, *I principi del concorso di persone nel reato*, Napoli, 1964; T. PADOVANI, *Le ipotesi speciali di concorso nel reato*, Milano, 1973; G. INSOLERA, *Problemi di struttura del concorso di persone nel reato*, Milano, 1986; S. SEMINARA, *Tecniche normative e concorso di persone nel reato*, Milano, 1987; M. PELISSERO, *Il concorso nel reato proprio*, Milano, 2004; L. RISICATO, *Il concorso colposo tra vecchie e nuove incertezze*, in *Riv. it. dir. proc. pen.*, 1998, pp. 132 ss. Sulla figura dell'autore mediato, in particolare, S. RICCIO, *L'autore mediato*, Napoli, 1939; M. SINISCALCO, voce *Autore mediato*, in *Enc. dir.*, IV, Milano, 1959, pp. 445 ss.; G. SAMMARCO, *Il concetto di autore e partecipe del reato nella più recente dottrina tedesca*, in *Riv. it. dir. proc. pen.*, 1979, pp. 1023 ss.; E. MORSELLI, *Note critiche sulla normativa del concorso di persone nel reato*, in *Riv. it. dir. proc. pen.*, 1983, pp. 422 ss.; G. CIANI, *Autore mediato e reato proprio*, in *Cass. Pen.*, 1997, pp. 1001 ss.

Nella dottrina classica si usa fare riferimento all'autore mediato in quelle ipotesi di concorso di persone in cui taluno dei concorrenti non sia punibile per questioni inerenti la sua persona[519]. In queste ipotesi, dell'illecito materialmente commesso dal soggetto indotto al reato (*deceptus*), risponde colui il quale l'ha determinato a commetterlo (*decipiens*). Volendo prescindere dalle classiche ipotesi di costringimento fisico (art. 46 c.p.), errore determinato dall'altrui inganno (art. 48 c.p.), coazione morale (art. 54 co. 3 c.p.) e determinazione in altri dello stato d'incapacità allo scopo di far commettere un reato (art. 86 c.p.), sembra che l'ipotesi di responsabilità vicaria che potrebbe meglio adeguarsi all'illecito materialmente commesso dall'IA sia quella di cui all'art. 111 c.p., che disciplina la determinazione al reato di persona non imputabile o non punibile. I sistemi intelligenti, infatti, non sarebbero (allo stato attuale) né imputabili né punibili.

Preliminarmente occorrerebbe rilevare che una rigida interpretazione del principio di tassatività non consentirebbe di applicare l'art. 111 c.p. al reato commesso dall'IA in quanto non sarebbe possibile estendere analogicamente il concetto di "persona" fino a ricomprendere sistemi non umani. Inoltre, anche a voler sperimentare un'interpretazione di tal fatta, sembra possibile riscontrare che, alla luce della penalistica classica, quando viene chiamata in campo la responsabilità vicaria si presuppone che l'autore materiale del reato sia dotato di autonomia d'azione, di cui non tutti i sistemi di IA sono forniti. Ad esempio, in tal sede, ci stiamo occupando di forme di IA in grado di porre in essere comportamenti automatici (e, dunque, prevedibili) e non autonomi (pertanto, imprevedibili). Tuttavia, anche ove prendessimo in considerazione sistemi intelligenti in grado di agire in (quasi totale) autonomia, dovremmo confrontarci con una questione probabilmente insormontabile. L'art. 111 c.p. si riferisce a persone non imputabili e non punibili: le prime potrebbero difettare della capacità d'agire, ma non anche della capacità giuridica (es. un soggetto che non ha ancora raggiunto la maggior età); le seconde, invece, potrebbero possedere entrambe tali capacità e non essere ugualmente punibili (ad esempio, perché hanno agito senza dolo). La capacità di essere titolare di diritti e di compiere atti giuridici volti ad esercitarli è caratteristica tipica dell'uomo, che se non è assimilabile alla macchina da un punto di vista empirico, non può esserlo (a nostro avviso) neanche da un punto di vista giuridico. A meno di non voler stravolgere la scelta di campo operata dal legislatore, non sembra che (allo stato attuale) sia possibile applicare l'art. 111 c.p. al reato materialmente commesso dall'IA, perché è vero che quest'ultima non è né imputabile né punibile, ma non lo è in quanto entità non umana e mero strumento nelle mani dell'uomo.

Proprio per questo ci sentiamo in tal sede di condividere l'opinione di chi ritiene che, «a differenza dei casi di c.d. autoria mediata (*mittelbare Täterschaft*), al soggetto agente non va imputata una condotta altrui, ma un fatto che gli appartiene, in quanto si è servito dolosamente di un a.a. [agente artificiale] quale *strumento* per la commissione del reato»[520].

[519] G. FIANDACA, E. MUSCO, *Diritto Penale,* cit., p. 522; G. MARINUCCI, E. DOLCINI, G.L. GATTA, *Manuale di diritto penale*, cit., p. 575.
[520] I. SALVADORI, *Agenti artificiali*, cit., p. 101.

16.3. Dall'automazione all'autonomia.

Nonostante, in linea di principio, il controllo umano andrebbe mantenuto in ogni momento[521], abbiamo già avuto modo di intuire come ciò non sia sempre possibile quando entrano in gioco i più sofisticati sistemi di IA. L'autonomia di cui essi sono dotati non è l'autonomia personale propria dei soggetti consapevoli di sé e in grado di agire in base alle proprie libere intenzioni, bensì un'"autonomia operativa", limitata dunque allo svolgimento di un dato compito che può essere svolto senza l'intervento umano[522]. Si tratta di un'autonomia "operazionale", puramente tecnologica. I problemi possono iniziare a sorgere nel momento in cui il sistema intelligente assuma una "decisione" al di fuori del programmato[523]. A prima vista, nel caso in cui l'IA sia utilizzata dall'uomo come semplice "mezzo", la responsabilità per i fatti lesivi da questa cagionati sarebbe comunque imputabile al programmatore o al produttore, anche nell'eventualità in cui le azioni realizzate dall'IA non fossero state *ex ante* programmate[524]. Tuttavia la possibilità che l'IA agisca in modo non programmato e che il danno derivato da un sistema sfugga al controllo umano è questione destinata inevitabilmente ad entrare in collisione con i principi penalistici, specie in punto di ricostruzione del nesso causale e di rispetto del principio di personalità della responsabilità penale, inteso come possibilità di muovere un rimprovero al soggetto e come divieto di responsabilità per fatto altrui[525].

L'imprevedibilità dell'agire del sistema intelligente porta con sé seri problemi di attribuzione della responsabilità penale[526]: individuare la persona fisica chiamata a rispondere di un fatto lesivo commesso dall'IA non sarà semplice, individuare l'elemento soggettivo che sorregga l'imputazione sarà ancor meno semplice. Se per le ipotesi dolose sembra possibile considerare in ogni caso l'IA come *longa manus* dell'agente, nelle ipotesi colpose la questione è destinata a complicarsi. Se il danno fosse da imputare a un uso colposo dell'IA si dovrà dimostrare la prevedibilità e l'evitabilità, da parte dell'utilizzatore, del danno verificatosi. Se il danno, invece, fosse da imputare a un malfunzionamento del sistema non solo si dovrà andare a dimostrare in cosa sia consistito il difetto dell'IA[527] ma si dovrà, inoltre, verificare se il programmatore abbia peccato di negligenza e se avrebbe potuto prevedere ed evitare il comportamento (a sua volta imprevedibile) dell'IA[528]. Siamo davanti a una *probatio diabolica* che necessiterebbe, per essere neutralizzata, di un maggiore grado di certezza, di norme volte a disciplinare l'imputazione della responsabilità in casi farraginosi come questi. Ciò sempre, chiaramente, partendo dall'assunto che ad essere considerato responsabile sia l'uomo dietro l'IA e non l'IA medesima.

---

[521] C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 7; A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 12.
[522] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., p. 36.
[523] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 12.
[524] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1760.
[525] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1758.
[526] M.B. MAGRO, *Biorobotica*, cit., p. 514; M.B. MAGRO, *Robot*, cit., p. 1208.
[527] G. CAPILLI, *I criteri di interpretazione,* cit., p. 483.
[528] M.B. MAGRO, *Biorobotica*, cit., p. 516.

Eccoci, dunque, davanti all'alternativa su cui riposano tutte le problematiche penalistiche accennate: i sistemi intelligenti devono essere considerati come strumenti per commettere un reato o come veri e propri soggetti di diritto penale?

17. Intelligenza artificiale come "autore" di reato.

A fare da apripista in questa direzione è stato il Parlamento europeo, il quale per primo si è reso conto, mettendolo nero su bianco, che «più i robot sono autonomi, meno possono essere considerati come meri strumenti nelle mani di altri attori»[529]. Ferma restando la definizione di "autonomia" fornita dal succitato Parlamento, il quale la intende come «la capacità di prendere decisioni e metterle in atto nel mondo esterno, indipendentemente da un controllo o un'influenza esterna»[530], sono stati formulati in dottrina altri tentativi definitori. Secondo alcuni con il termine "autonomia" ci si riferisce alla capacità della macchina di «svolgere i propri compiti senza l'interposizione di un soggetto esterno (umano) che impartisca comandi» o, ancora, alla capacità della macchina di «apprendere, accrescendo o modificando il bagaglio di conoscenze datole "in dotazione" al momento della sua realizzazione»[531].

La crescente autonomia di questi sistemi porta a renderli, almeno in parte, indipendenti dall'uomo e, dunque, ad offuscare la linea di confine tra soggetto e strumento[532]. In dottrina ci si è anche chiesti come sia avvenuto questo cambio di prospettiva, passando a considerare l'IA da "strumento" a "soggetto". È stata individuata a tal proposito una "mutazione diretta", dovuta al fatto che l'uomo demanda sempre più spesso l'assunzione di decisioni per suo conto a sistemi automatizzati, e una "mutazione indiretta", connessa al fatto che le decisioni ad oggi assunte dall'uomo sono fondate su conoscenze fornite, a loro volta, dai suddetti sistemi tecnologici[533].

La capacità delle più evolute forme di IA di "autodeterminarsi", di adattarsi all'ambiente che le circonda indipendentemente dal controllo umano, le rende sempre meno governabili[534], al punto da poterle definire "seagenti"[535]. Si usa parlare a tal proposito (come invero accennato in precedenza)[536] di "comportamento emergente". L'"emergenza" si riferisce, per l'appunto, alla capacità o alla tendenza di questi sistemi di comportarsi in modi complessi e imprevisti, alla loro capacità di imparare dall'esperienza e risolvere problemi in modi che i loro creatori non avrebbero mai immaginato[537]. Il livello di autonomia operativa che questi sistemi riescono a realizzare porta con sé la difficoltà di comprendere "come decide la macchina" e chi sia il reale responsabile del suo agire[538].

---

[529] *Risoluzione del Parlamento europeo*, cit., Considerando AB. Anche se il Parlamento europeo nel corpo del testo si riferisce ai robot riteniamo che queste considerazioni siano ugualmente valevoli anche per quanto concerne i sistemi intelligenti.
[530] *Risoluzione del Parlamento europeo*, cit., Considerando AA.
[531] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., pp. 83-84.
[532] R. CALO, *Robots in American Law*, cit., p. 43.
[533] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., pp. 69-70.
[534] G. CORASANITI, *Intelligenza artificiale e diritto*, cit., p. 397.
[535] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 24.
[536] Cap. II, Sez. II, Par. 14.
[537] R. CALO, *Robots in American Law*, cit., p. 40.
[538] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 13.

La vera sfida arriva quando iniziamo a considerare l'IA come «entità autonoma capace di agire indipendentemente dal dolo o dalla colpa del creatore o programmatore umano. Se presupponiamo che i soggetti di intelligenza artificiale siano capaci di determinare il proprio comportamento in completa autonomia, allora dovremmo chiederci chi debba essere ritenuto responsabile delle loro azioni criminose – dal momento che non possiamo attribuirle ad alcun essere umano»[539]. Entriamo dunque nel vivo del problema più complesso che si pone nella materia *de qua*: la possibile inclusione dell'intelligenza artificiale tra i soggetti del reato.

17.1. Nuove soluzioni per nuovi problemi?

Chiaro è che, parlando dell'IA non più come strumento bensì come autore di reato, ipotizzare una pacifica applicabilità delle categorie penalistiche sarà più complesso. Se il dominio normalmente esercitato dall'uomo sulla macchina viene a mancare, entrano conseguentemente in crisi i modelli di imputazione penale e di individuazione del soggetto responsabile degli eventuali danni causati dall'IA[540]. Il ruolo dell'intelligenza artificiale nel diritto penale è destinato a costituire terreno di sfide, a porre questioni giuridiche di rilievo e a imporre un ripensamento delle basi della teoria generale del reato: abbiamo già avuto modo di intuire come la materia *de qua* ponga problemi in punto di ricostruzione del nesso causale, di individuazione dell'elemento soggettivo e, financo, sul concetto stesso di azione[541].

Tali consolidate categorie potrebbero subire una considerevole "pressione distorsiva" che le renderebbe non sempre adattabili[542]. Per questo occorrerà verificare il possibile attrito con le norme già esistenti e prendere in considerazione, preliminarmente, una loro espansione e, secondariamente, l'introduzione di una nuova normativa[543]. Si è parlato in dottrina di "tecnologia eccezionale", annoverando in questa categoria non tanto quelle tecnologie atte a modificare il diritto in minima parte quanto, piuttosto, quelle la cui introduzione nell'ordinamento richiede un cambiamento sistematico della legge al fine di riprodurre o, financo, modificare un preesistente equilibrio di valori[544]. Il rischio, da questo punto di vista, è duplice: da un lato, quello di non colmare gli esistenti (o pretesi tali) vuoti normativi e, dall'altro, quello di dar vita a un'"iperfetazione legislativa"[545].

Riteniamo dunque che, come primo passaggio, si dovrebbe tentare un'interpretazione evolutiva delle norme attualmente vigenti, in considerazione del fatto che, come per ogni momento evolutivo, vi sarà un periodo in cui la norma "vecchia" regolerà temporalmente il fenomeno "nuovo"[546]. È però stato ritenuto in dottrina che, effettivamente, l'IA porti con sé specifici elementi che potrebbero giustificare l'introduzione di una nuova normativa, quali la sua

---

[539] S. RIONDATO, *Robotica e diritto penale*, cit., p. 601.
[540] C. PIERGALLINI, *Intelligenza artificiale*, cit., pp. 1750-1751.
[541] L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., p. 80, nonostante l'A. si riferisca in tal senso ai reati cibernetici riteniamo le Sue considerazioni perfettamente coincidenti con le problematiche che oggi pone l'IA in materia penale.
[542] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 12.
[543] Per qualche ulteriore considerazione sul punto v. Cap. II, Sez. II, Par. 10.
[544] R. CALO, *Robotics and the Lessons of Cyberlaw*, cit., p. 552.
[545] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 3.
[546] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 4.

capacità di autoapprendimento, la possibile "alterità" dell'IA rispetto al prodotto che la incorpora e la mancata coincidenza del responsabile della creazione dell'IA e del bene su cui essa si innesta[547]. In altre parole, si potrebbero rendere necessarie nuove norme e nuove categorie giuridiche per disciplinare la responsabilità derivante da condotte lesive realizzate dall'IA[548] o, in ottica penalistica, da reati commessi da, piuttosto che mediante, l'intelligenza artificiale[549].

17.2. Problemi di responsabilità: la (presunta) crisi del modello vicario.

Iniziamo a comprendere che quando ci riferiamo alle più evolute forme di intelligenza artificiale, capaci di autoapprendimento e di autodeterminazione, non si tratta più di meri strumenti, bensì di veri e propri agenti[550]. La loro presenza nel nostro ordinamento potrebbe portarci ad una rivisitazione, quantomeno parziale, dei tradizionali concetti di capacità giuridica e di agire[551]. A differenza dei sistemi su cui ci siamo concentrati in precedenza, l'IA di ultima generazione prescinde dal controllo umano, potendo realizzare condotte autonome e non predeterminabili, adattandosi all'ambiente circostante. È proprio qui che risiede il problema più complesso: il comportamento dell'IA non può essere né *ex ante* prevedibile né totalmente controllabile dall'uomo, non solo da un punto di vista soggettivo (proprio in quanto condotta imprevedibile) ma anche da un punto di vista oggettivo. Torna attuale il problema del *black box*, in quanto tra l'input inserito dall'uomo e l'output prodotto dalla macchina vi è un vuoto di comprensione che rende l'agire dell'IA caratterizzato da un «ineliminabile margine di imponderabilità. Anzi, potremmo dire che, in una certa misura, l'imprevedibilità dell'agente intelligente è "pre-programmata", e con essa i rischi associati a terzi»[552].

Se il comportamento dannoso dell'IA è imprevedibile, in quanto essa impara dall'esperienza indipendentemente dalla supervisione umana, individuare la persona cui poter muovere un rimprovero diventa tutt'altro che semplice, specie con gli attuali schemi giuridici di cui disponiamo[553]. Più tali sistemi diventano autonomi, meno il loro comportamento sarà prevedibile e il controllo dell'uomo incisivo[554]: «la mente del programma si distingue da quella del suo programmatore»[555]. Se dunque tali sistemi non possono essere puniti, si dovrà comprendere a quali condizioni gli uomini dovrebbero essere ritenuti penalmente responsabili per la produzione, la programmazione o l'utilizzo di una macchina intelligente che abbia causato un danno[556]. Non si tratta di un interrogativo di poco conto: se gli uomini non possono controllare completamente questi sistemi,

---

[547] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 17.
[548] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 82.
[549] M.B. MAGRO, *Robot,* cit., p. 1206.
[550] I. SALVADORI, *Agenti artificiali*, cit., p. 85; M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 364.
[551] G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 111.
[552] M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 5. Altri preferiscono parlare di "rappresentata imprevedibilità", A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 92.
[553] R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 265; G. CAPILLI, *I criteri di interpretazione*, cit., p. 458.
[554] G. CAPILLI, *I criteri di interpretazione*, cit., p. 463.
[555] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 11.
[556] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., p. 424.

se non possono prevedere ed evitare le loro condotte lesive, come possono essere ritenuti per esse responsabili[557]? È in questo che consiste il *responsibility gap* che caratterizza la materia *de qua*. Si tratta di un vero e proprio vuoto di responsabilità che si crea quando viene realizzato un illecito che non sia riconducibile all'azione di un soggetto ben individuato (o comunque individuabile) bensì all'agire di un sistema intelligente dotato di un certo margine di libertà decisionale[558]. Probabilmente il meccanismo d'imputazione della responsabilità oggettiva in capo al produttore dell'IA sarebbe la via migliore per colmare questo *gap* e garantire la certezza del diritto[559], ma resta una via impercorribile per i crismi del diritto penale che siamo abituati a conoscere.

Con riferimento a sistemi così evoluti da poter essere considerati autonomi – in grado, cioè, di realizzare comportamenti non prevedibili e non spiegabili – la dottrina ha riscontrato la crisi del modello di imputazione vicaria, poiché la macchina sveste i panni del mero strumento nelle mani dell'uomo[560] per assumere quelli di nuovo agente. Grazie all'impiego di algoritmi di *machine learning* (specie di *deep learning*) e di *cloud computing* l'IA è in grado di imparare dalla propria esperienza e da quella di sistemi simili ad essa connessi[561]. Ferme restando le perplessità pocanzi espresse in ordine all'opportunità di scomodare l'imputazione vicaria per ascrivere il fatto lesivo al soggetto umano[562], il ragionamento di fondo è che, finché si tratta di macchine le cui attività possano essere predefinite e, dunque, prevedibili sarebbe in astratto possibile chiamare a rispondere il programmatore almeno a titolo di colpa[563].

Parlare di imputazione vicaria in questo ambito ci sembra, invero, maggiormente confacente. Se in prima battuta, considerando l'IA come mero strumento, abbiamo ritenuto non necessario chiamare in causa i meccanismi della responsabilità indiretta, proprio perché l'IA è interamente controllabile dall'uomo, nel momento in cui quest'ultima raggiunge un livello di autonomia tale da potersi anche solo pensare di considerarla come "agente", la situazione cambia. La crisi dell'imputazione vicaria, infatti, sarebbe determinata proprio dall'imprevedibilità e incontrollabilità dell'agire dell'IA[564]. Ad avviso di alcuni sarebbe proprio l'impraticabilità del modello vicario a generare un vuoto di protezione penale per quelle offese che, in realtà, non siano attribuibili ad alcun operatore umano[565]. Tuttavia riteniamo che, se in tal sede sia superabile l'obiezione precedentemente sollevata in ordine al minimo grado di autonomia di cui dovrebbe essere dotato il *decipiens*, in quanto – a differenza dei sistemi automatici – gli esaminandi sistemi sono dotati di un certo margine di autonomia, persiste in tal sede la principale obiezione. Non siamo del tutto certi che sia possibile parlare, anche qui, di responsabilità vicaria. Intanto si potrebbe ritenere che l'imprevedibilità dell'IA sia fattore endemico, tale da dover essere previsto dal suo programmatore, il quale se ne assumerebbe, in un certo senso, la paternità: «i sistemi di Intelligenza artificiale

[557] M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 6.
[558] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 82.
[559] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 95.
[560] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 7.
[561] P. SEVERINO, *Intelligenza artificiale*, cit., p. 533.
[562] Cap. II, Sez. III, Par. 16.2.
[563] G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 9.
[564] R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 262.
[565] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 19.

non possono diventare autonomi in modo spontaneo, ma per farlo dovranno essere progettati in questo modo. I possibili eventi lesivi prodotti dall'intelligenza artificiale, pertanto, potranno dipendere sempre da come questa è stata progettata per essere autonoma e dal fatto che l'autonomia fa sì che il robot sia fuori dal controllo dell'essere umano»[566]. In quest'ottica, l'agire imprevedibile dell'IA non sarebbe fattore di crisi per la responsabilità vicaria, bensì naturale connotato del sistema intelligente. Inoltre, come approfondito pocanzi, l'ostacolo più grande da dover aggirare in materia resterebbe pur sempre la considerazione svolta a proposito dell'art. 111 c.p., il quale resta attualmente strutturato su un soggetto in carne e ossa, con i diritti e i doveri attribuitigli dall'ordinamento giuridico e, pertanto, non assimilabile alla macchina.

17.2.1. Segue: il reato diverso da quello voluto e l'interruzione del nesso causale.

Volendo dunque sintetizzare, l'autonomia di cui l'IA può essere dotata, specie grazie ai sistemi di *machine learning*, porta con sé considerevoli problemi da un punto di vista giuridico-penale:

- l'IA può compiere condotte non prevedibili da parte degli operatori che l'hanno progettata e realizzata[567], deviando dunque dal percorso per essa programmato;
- la sua imprevedibilità pone un serio problema in punto di allocazione della responsabilità[568] (il danno verificatosi sarà da imputare all'IA o all'uomo? Optando per tale ultima ipotesi, quale degli operatori dietro l'IA dovrebbe essere considerato responsabile?)[569];
- le difficoltà connesse all'individuazione del soggetto cui imputare la responsabilità per l'evento lesivo verificatosi potrebbero condurre ad inaccettabili vuoti di tutela[570].

---

[566] G. CAPILLI, *I criteri di interpretazione*, cit., p. 482.

[567] I. SALVADORI, *Agenti artificiali*, cit., p. 88; A. CAPPELLINI, Machina delinquere non potest?, cit., p. 9, il quale richiama a sua volta alla nota 22 S. BECK, *Google Cars, Software Agents, Autonomous Weapons Systems – New Challenges for Criminal Law?*, in E. HILGENDORF, U. SEIDEL, *Robotics, Autonomics and the Law*, Baden, 2017, p. 244, la quale afferma (secondo la traduzione dell'Autore) che «se l'imprevedibilità fa parte del modo in cui è concepito il robot, non ogni suo "errore" potrà essere imputato colposamente al programmatore»; M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 357.

[568] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1759; F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 10.

[569] «Tanti sono i soggetti e le componenti che interagiscono nella produzione dell'evento dannoso finale. Chi di loro può essere considerato responsabile, se come abbiamo detto, gli agenti artificiali, sebbene inizialmente programmati da agenti umani, non sono costituiti da agenti umani, né agiscono attraverso agenti umani? Una volta programmato, l'IA non fa più affidamento sul programmatore, interagisce con il mondo senza la necessità che il programmatore funga da burattinaio. (…) il comportamento di *Intelligent Agent* è imprevedibile non solo quando si trova in una situazione per la quale non è stata programma una risposta adeguata, ma anche quando, a causa di una nuova "esperienza", inizi a modellare i dati acquisiti autonomamente. Qui sta l'origine di una grande inquietudine: l'uomo non può né prevedere, né controllare totalmente il comportamento dell'agente artificiale in situazioni non pianificate» M.B. MAGRO, *Decisione umana e decisione robotica*, cit., pp. 4-5.

[570] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 10; I. SALVADORI, *Agenti artificiali*, cit., p. 107.

Vorremmo concentrarci qui sui primi dei due punti presi in esame, in quanto il terzo (ossia il vuoto di responsabilità) costituisce naturale conseguenza della mancata risoluzione degli altri due.

Quanto all'imprevedibilità dell'IA come fattore deviante dal percorso pre-programmato, tale inciso sembra ricordare, in un contesto di *versari in re illicita*, la struttura dell'art. 116 c.p., il quale disciplina il reato diverso da quello voluto da taluno dei concorrenti[571]. In quest'ottica, del reato più grave non voluto risponderebbe il reo che volle il reato meno grave, ove la fattispecie criminosa concretamente realizzatasi sia conseguenza della sua azione od omissione. Che il reato più grave originariamente non voluto sia frutto della condotta (attiva o omissiva) dell'operatore dell'IA, da un punto di vista *astratto*, è fuor di dubbio: nonostante i sistemi intelligenti possano agire autonomamente, ciò che è certo è che quantomeno l'input iniziale dovrà essere impartito dall'uomo. Pensiamo però al seguente esempio: un gruppo di malintenzionati si accorda per realizzare una rapina in banca. Dopo un'attenta osservazione i ladri riscontrano che la banca resta disabitata fino alle 6:00 del mattino e meditano, pertanto, di realizzare il colpo prima di quell'ora. I rei decidono di servirsi di una macchina a guida autonoma e di programmarla per sfondare la vetrina della banca, in modo da consentire ai ladri di introdursi nell'edificio e prelevare il denaro. La macchina, una volta avviata la sua corsa e aperto il varco nella banca, travolge un inserviente che si trovava lì (inaspettatamente) a fare le pulizie alle 5:00 del mattino. In una simile ipotesi potremmo ritenere che, in *concreto*, l'evento morte indesiderato non costituisca sviluppo logicamente prevedibile dell'azione del programmatore dell'IA, proprio in quanto la macchina non era stata programmata per cagionare l'evento morte e poi, inoltre, perché la presenza dell'inserviente all'interno della banca non era prevedibile. Anche qui, però, ci scontriamo con gli ostacoli – attualmente insormontabili – che abbiamo riscontrato a proposito dell'art. 111 c.p. Pensare ad un concorso di "persone" tra uomo e intelligenza artificiale imporrebbe una completa assimilazione delle due entità che, non essendo al momento sostenibile, ci condurrebbe a ripiegare sulla mera considerazione dell'IA come strumento programmato dall'uomo per delinquere.

---

[571] A. PAGLIARO, *La responsabilità del partecipe per il reato diverso da quello voluto*, Milano, 1966; A. PAGLIARO, *Diversi titoli di responsabilità per uno stesso fatto concorsuale*, in *Riv. it. dir. proc. pen.*, 1994, pp. 3 ss.; R. PANNAIN, *Sull'art. 116 del c.p.*, in *Archivio Penale*, 1965, pp. 439 ss.; F. BASILE, *Commento all'art. 116 – Reato diverso da quello voluto da taluno dei concorrenti*, in E. DOLCINI, G.L. GATTA (diretto da), *Codice penale commentato*, vol. I, IV ed., Milano, 2015, pp. 1839 ss.; G. INSOLERA, *Tentativo di una diversa lettura costituzionale dell'art. 116 c.p.*, in *Riv. it. dir. proc. pen.*, 1978, pp. 1489 ss.; S. CANESTRARI, *La responsabilità del partecipe per il reato diverso da quello voluto e il principio di colpevolezza*, in *Studium Iuris*, 1996, pp. 1396 ss.; A. CADOPPI, S. CANESTRARI, A. MANNA, M. PAPA, *Diritto Penale*, I, Milano, 2022, pp. 654 ss.; R. STOCCO, *Alla ricerca di una dimensione costituzionale dell'art. 116 c.p.*, in *Cass. pen.*, 1990, pp. 35 ss.; G. CIANI, *Brevi considerazioni sulla responsabilità del concorrente per reato diverso da quello voluto*, in *Cass. pen.*, 1996, pp. 3644 ss.; A. GULLO, *La responsabilità del partecipe per il reato diverso da quello voluto tra* versari in re illicita *e principio di colpevolezza*, in *Riv. it. dir. proc. pen.*, 2000, pp. 1194 ss.; C. PEDRAZZI, *Tramonto del dolo?*, in *Riv. it. dir. proc. pen.*, 4/2000, pp. 1270 ss.; A. MACCHIA, *Concorso anomalo: un tentativo (azzardato?) di ricostruzione della responsabilità per il fatto diverso da quello voluto*, in *Cass. pen.*, 2/2017, pp. 492 ss.; E. BASILE, *Condotta atipica e imputazione plurisoggettiva: alla ricerca del coefficiente di colpevolezza del concorrente "anomalo"*, in *Riv. it. dir. proc. pen.*, 2005, pp. 1336 ss.

Davanti ad un'ipotesi come quella proposta, per attribuire la responsabilità per la morte non voluta, si potrebbe al più pensare di chiamare in causa la disciplina del delitto preterintenzionale, in quanto a ben vedere il reato materialmente commesso andrebbe oltre l'intenzione, integrando un reato più grave di quello concretamente voluto. Ferme restando tutte le perplessità manifestate dalla dottrina con riguardo alla figura della preterintenzione[572], vorremmo in tal sede considerare che comunque tale elemento soggettivo risulta invocabile solo ove l'omicidio preterintenzionale (art. 584 c.p.) origini dal reato di percosse o da quello di lesioni personali. Per poter chiamare in causa, dunque, la fattispecie preterintenzionale dovremmo immaginare un sistema intelligente programmato per ledere che, andando oltre l'intenzione dell'agente, finisca per uccidere.

Quanto, invece, all'allocazione della responsabilità, l'imprevedibilità dell'IA, come accennato in precedenza[573], potrebbe essere intesa come fattore interruttivo del nesso causale[574], collocandosi tra la condotta dell'operatore e l'illecito realizzatosi. Autorevole dottrina che si è occupata del tema ha distinto un approccio restrittivo e un approccio ampio. Facendo una lettura rigida dell'art. 41 c.p. deve dedursi che il comportamento autonomo dell'IA valga a integrare un'ipotesi di "*causalità sorpassante*", atta cioè ad annullare il rischio iniziale innescato dall'uomo e a introdurne uno nuovo, ossia la decisione auto-determinata dell'IA. Adottando un approccio più morbido, invece, si potrebbe chiamare in gioco la "*causalità umana*" considerando l'evento lesivo finale come eccezionale, in quanto non governabile dall'uomo e, conseguentemente, a lui non imputabile. In tale ottica si potrebbe invocare anche la "*causalità adeguata*", ove volessimo affermare che la programmazione iniziale dell'IA non era *ex ante* idonea a produrre l'evento, a sua volta, frutto della decisione autonoma del sistema intelligente. A ben vedere, in ciascuna di queste ipotesi, non sarebbe comunque possibile imputare l'evento penalmente rilevante all'uomo dietro l'IA. Tuttavia la dottrina in esame mette in evidenza che coloro i quali realizzano un sistema intelligente sono ben consci del fatto che stanno creando un sistema in grado di realizzare condotte imprevedibili: «viene, per questa via, introdotto un rischio nella consapevolezza di non poterne governare appieno gli effetti. In presenza di un simile scenario, appellarsi all'interruzione del nesso causale non appare, quindi, giuridicamente corretto e l'adesione a questa impostazione lascerebbe le vittime in una condizione di totale, inammissibile vuoto di tutela»[575].

Sembra pertanto che, servendoci degli strumenti penalistici a nostra disposizione, imputare la responsabilità all'uomo dietro l'IA per il suo agire illecito sia compito particolarmente arduo. Ciò porta a chiederci se sia possibile riferire all'IA categorie normalmente riferibili all'agire umano, come ad esempio

---

[572] G. FIANDACA, E. MUSCO, *Diritto Penale*, cit., pp. 691 ss.
[573] Cap. II, Sez. II, Par. 12.
[574] Sul tema, tra gli altri, F. ANTOLISEI, *Il rapporto di causalità nel diritto penale*, Padova, 1934; A.A. DALIA, *Le cause sopravvenute interruttive del nesso causale*, Napoli, 1975; P. PIRAS, *Quando la causa sopravvenuta è sufficiente a determinare l'evento*, in *Dir. pen. proc.*, 1997, pp. 961 ss.; R. BLAIOTTA, *Causalità giuridica*, Torino, 2010; A. VALLINI, *"Cause sopravvenute da sole sufficienti" e nessi tra condotte. Per una collocazione dell'art. 41, comma 2, c.p. nel quadro teorico della causalità "scientifica"*, in *dirittopenalecontemporaneo.it*, 11.7.2012.
[575] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1762.

il concetto di azione o di atteggiamento psicologico[576], i quali integrerebbero l'elemento oggettivo e l'elemento soggettivo del fatto criminoso commesso dall'IA. Dovremo chiederci, in altri termini, se sia possibile applicare le tradizionali categorie del diritto a tecnologie talmente evolute da portarci a considerare "soggetto" ciò che fino a ieri avremmo considerato "oggetto", verificando «la tenuta delle norme esistenti e la loro idoneità a cogliere le problematicità derivanti dalla diffusione di intelligenze artificiali»[577].

Il margine di autonomia[578] dei più evoluti sistemi intelligenti potrebbe costituire la porta d'ingresso per le possibili teorie di imputazione della responsabilità penale ai sistemi intelligenti[579], per iniziare a parlare di responsabilità diretta dell'IA[580] o, per utilizzare un'espressione ancora più forte, di *capacità penale di agire* dell'IA[581].

17.3. Personalità giuridica elettronica: considerazioni di carattere generale.

Iniziando a pensare ai sistemi di IA non più come semplici oggetti bensì come veri e propri soggetti, da un punto di vista squisitamente giuridico e a livello teorico, non vi sarebbero troppi ostacoli al riconoscimento di una qualche forma di personalità giuridica in capo a soggetti non umani quali, ad esempio, le macchine. Nulla osterebbe in quanto il legislatore potrebbe scegliere di considerare queste macchine come soggetti[582]. Il riconoscimento della "personalità" in capo all'IA sarebbe pur sempre fittizio – in quanto anche i più sofisticati sistemi, in grado di agire in modo indipendente, sarebbero pur sempre programmati in tal senso – nonché connesso alla necessità di perseguire uno specifico obiettivo da parte dell'uomo[583]. Nonostante ad oggi si usi parlare spesso di "finzioni", ad avviso di parte della dottrina non si tratterebbe di una mera "finzione" da parte del legislatore, bensì della creazione di una "realtà giuridica" che appare ai nostri occhi come mera finzione sol perché non coincide con la "realtà naturale": «si tratta, invece, più semplicemente, della capacità, tipica del diritto, di creare soluzioni tecniche per rispondere a esigenze pratiche. Le regole giuridiche, insomma, hanno il potere di creare realtà che non necessariamente corrispondono a realtà ontologiche»[584].

---

[576] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 351.
[577] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 369.
[578] La rilevanza del concetto di "autonomia" era già stata attenzionata dalla dottrina che si è occupata dei reati commessi nel Cyberspace, dunque in rete e mediante l'uso di sistemi informatici. Quanto a questi fenomeni la dottrina si chiede «in che misura siano anche giuridicamente imputabili alla "consapevole volontà" dell'uomo che li attiva o se ne serve. L'*autonomia* (seppur relativa) delle determinazioni e "scelte" operative dei sistemi informatici non sembra poterne escludere in linea di principio l'attribuzione (anche) alla sottostante "volontà" dei soggetti umani, da cui i sistemi stessi e i loro *output* in ultima analisi necessariamente dipendono. Ma sono certamente da precisare i presupposti ed i limiti, in relazione ai quali può dirsi esercitato e mantenuto, ai fini della responsabilità penale, il "dominio" dell'uomo su tutti i risultati ed effetti che conseguono e permangono, spesso a grande distanza di tempo e di luogo» L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., p. 90.
[579] F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 30.
[580] P. SEVERINO, *Intelligenza artificiale*, cit., p. 534.
[581] I. SALVADORI, *Agenti artificiali*, cit., p. 95.
[582] A. CELOTTO, *I robot possono avere diritti?*, cit., p. 95.
[583] S. RIONDATO, *Robotica e diritto penale*, cit., p. 603; ID., *Robot: talune implicazioni di diritto penale*, cit., p. 92.
[584] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 233.

La soggettività giuridica, infatti, è stata considerata come un'invenzione dell'uomo[585], come la capacità di adattamento del diritto volta a fronteggiare le esigenze che emergono dalle diverse dinamiche della società. Il legislatore, per il tramite del diritto, infatti, ha la possibilità di attribuire soggettività giuridica a qualunque entità che egli ritenga utile dotare di soggettività, indipendentemente dal fatto che essa sia o meno dotata di una propria soggettività ontologica. «L'uso di "finzioni" per aggirare le barriere della soggettività naturalisticamente intesa, è un'operazione tutt'altro che rara, e comunque in molti casi praticata e accettata, nella dimensione giuridico-normativa»[586].

L'idea del riconoscimento di una specifica "personalità elettronica", di uno status giuridico *ad hoc*, per i sistemi di intelligenza artificiale maggiormente sviluppati, in grado di prendere decisioni autonome e di interagire col mondo esterno, è stata accarezzata anche a livello europeo, segnatamente dalla ormai più volte citata Risoluzione del Parlamento europeo contenente "Norme di diritto civile sulla robotica"[587]. A ben vedere la distinzione tra intelligenza artificiale debole e forte torna anche in tal sede, in quanto è chiaro che discutendo di "soggettività" dei sistemi intelligenti ci si sta riferendo alla versione "forte" dell'IA, quella cioè in grado di essere considerata come soggetto, anche di diritto[588]. Occorre capire se le peculiarità dell'IA giustifichino la creazione di nuove categorie giuridiche[589] o se le categorie tradizionali del diritto possano adeguarsi alle evolute tecnologie di cui ci stiamo occupando, tali cioè da apparire più come "soggetti" che come "oggetti"[590]. Lo sviluppo dell'intelligenza artificiale è tale da farla apparire ai nostri occhi non più come mera macchina ma come entità animata, forse anche meritevole di "personalizzazione".

Prendendo le mosse dall'autorevole proposta europea – volta a riconoscere la personalità elettronica al fine di rendere i sistemi intelligenti responsabili per i danni da loro causati, imputandogli il relativo risarcimento – si dovrà riflettere con mente aperta sulla questione filosofico/giuridica dello status della "persona elettronica", rimeditando i classici concetti di capacità giuridica e capacità d'agire, ma anche chiedendoci se ci troveremo innanzi a una nuova generazione di diritti e doveri[591]. Andrebbero in tal senso evitati condizionamenti antropocentrici e abbandonati pregiudizi come la comparabilità con le fattezze umane o con la capacità di "sentire"[592].

Le macchine, finora, sono sempre state considerate solo come "oggetto" di diritti ma, come abbiamo accennato, questo stato di cose potrebbe essere destinato

---

[585] La dottrina infatti considera la "capacità di diritto" come mero prodotto dell'ordinamento F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 240.

[586] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 25.

[587] *Risoluzione del Parlamento europeo*, cit., Punto 59, lett. f). In dottrina, tra gli altri, ne parlano F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 227; A. CELOTTO, *I robot possono avere diritti?*, cit., p. 96.

[588] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 221.

[589] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 237; C. CASONATO, *Potenzialità e sfide dell'intelligenza artificiale*, cit., p. 182.

[590] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 369, gli AA. parlando a tal proposito di "diritto 2.0".

[591] P. MORO, *Macchine come noi*, cit., p. 61; G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., pp. 110-111.

[592] U. RUFFOLO, *Il problema della "personalità elettronica"*, in *Journal of Ethics and Legal Technologies*, 2020, p. 79.

a mutare[593]. Intanto occorrerebbe stabilire chi possa essere considerato titolare di diritti i quali, al pari di ogni altra posizione giuridica soggettiva, vengono attribuiti dall'ordinamento: «in altri termini, i diritti (soggettivi) sono attribuiti dal diritto (oggettivo). Con una precisazione importante. Il diritto oggettivo può attribuire diritti anche a non persone umane, come accade per le persone giuridiche e, in alcuni ordinamenti, agli animali»[594]. L'attribuzione alle macchine di una nuova forma di soggettività porta con sé considerevoli implicazioni. Intanto, la difficoltà di definire compiutamente cosa debba intendersi per "personalità elettronica"[595]. Inoltre il riconoscimento di una personalità di diritto in capo all'IA la renderebbe centro autonomo di imputazione di interessi e rapporti giuridici, direttamente responsabile per i danni da essa prodotti, nonché destinataria di diritti e doveri[596].

La dottrina ha riscontrato il vero problema del processo di "personificazione" dell'IA, non tanto nell'imputazione della responsabilità, quanto nel riconoscimento di diritti assimilabili a quelli umani[597]. È stato condivisibilmente osservato che la disponibilità della società a riconoscere ed accordare ai sistemi intelligenti «la titolarità di diritti dipende molto più dalle nostre proiezioni emotive, che non da loro qualità o capacità intrinseche»[598], un po' come è già accaduto in Nuova Zelanda con l'attribuzione della personalità giuridica a un fiume al fine di poterlo giuridicamente tutelare[599] o con l'introduzione nel nostro ordinamento dei delitti contro il sentimento per gli animali. Non si tratta di discorsi troppo lontani dalla realtà. Il caso più eclatante, ma invero non il solo[600], è quello del robot Sophia al quale l'Arabia Saudita nel 2017 ha riconosciuto la cittadinanza[601].

---

[593] A. CELOTTO, *I robot possono avere diritti?*, cit., p. 91.
[594] A. CELOTTO, *I robot possono avere diritti?*, cit., p. 93.
[595] La dottrina civilistica evidenzia che i concetti di "personalità" e di "soggettività del diritto" sono ormai da tempo scissi, F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 229.
[596] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 229-230; P. MORO, *Libertà del robot?*, cit., p. 535; U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 20, chiarisce che la "personificazione" dell'IA attiene principalmente al riconoscimento di diritto e doveri.
[597] U. RUFFOLO, *Il problema della "personalità elettronica"*, cit., p. 78.
[598] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 231.
[599] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 27.
[600] Nel 2017 a Tokyo è stata riconosciuta la residenza a Shibuya Mirai, un *chatbot* in grado di comunicare con i cittadini che lo interpellano, F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 29; C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 5. Quest'ultima A. riporta anche l'esperienza di Fabio, un robot licenziato nel 2018 dal suo impiego in un supermercato di Edimburgo, e di JIBO, il primo robot destinato alle famiglie, esempio di come l'evoluzione ci condurrà a ripensare al modo in cui interagiamo con le macchine. Ancora, nel 2014 il robot Vital è entrato a far parte del consiglio di amministrazione della compagnia giapponese Deep Knowledge grazie alla sua capacità di prevedere le tendenze del mercato, U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 622. Su questi temi v. U. PAGALLO, *Vital, Sophia, and Co. – The Quest for the Legal Personhood of Robots*, in *Information*, 10.9.2018.
[601] Menzionano il caso di Sophia C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 101; ID., *Potenzialità e sfide dell'intelligenza artificiale*, cit., p. 182; M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 11 nota 22; A. TURANO, *Robotica e roboetica*, cit., p. 152 nota 95. Più approfonditamente sul tema B. WATERS, *Citizen Sophia: It's (Past) Time to Legislate Robotics Regulation*, in *Georgetown Law Technology Review*, 2017; F.G. PIZZETTI, *The Robot Sophia as a "New Citizen" of Saudi Arabia: what about granting legal personhood, "citizenship" and eventually dignity to non-human entities with artificial intelligence?*, in *Notizie di Politeia*, 2019, pp. 63 ss.

Venendo però all'aspetto che desta maggiormente il nostro interesse, ossia quello dell'imputazione della responsabilità, è stato riscontrato in dottrina che considerare l'IA come soggetto di diritto implicherebbe attribuire a quest'ultima la capacità giuridica e d'agire[602]. Il riconoscimento in capo a questi sistemi della personalità giuridica consentirebbe di dotarli di un autonomo patrimonio, separato da quello del suo produttore, cui poter attingere per risarcire i danni causati dall'IA, limitando conseguentemente la responsabilità dei soggetti dietro la macchina[603]. «Responsabilizzare la macchina, ed essa sola, comporterebbe, di fatto, la limitazione della responsabilità patrimoniale ad uno specifico (e più ridotto) patrimonio di rischio»[604]. Tale considerazione assume significato solo ove se ne condivida il punto di partenza, ossia «*l'alterità del sistema autonomo rispetto all'utilizzatore*»[605]. Si tratterebbe, a ben vedere, di una considerevole novità: le entità personificate che siamo stati abituati a conoscere, finora, si sono sempre poggiate su un substrato umano, basti pensare alle classiche persone giuridiche. La "persona elettronica", invece, sarebbe sprovvista di questa immedesimazione, essendo per l'appunto "altra" rispetto all'ipotetico uomo dietro la macchina, nonché in grado di agire senza l'intermediazione umana[606].

Ci sentiamo di condividere in tal sede l'opinione di chi ha ritenuto che il modo in cui il diritto considererà i sistemi intelligenti dipenderà in larga parte dalle interazioni che l'uomo avrà con questi ultimi e dalle implicazioni relazionali che potrebbero derivarne (basti pensare, a titolo meramente esemplificativo, ai sistemi intelligenti predisposti all'assistenza alle persone vulnerabili)[607]. Come evidenziato nell'incipit del presente paragrafo, in astratto, non vi sarebbero consistenti ostacoli al riconoscimento della personalità giuridica in capo all'IA, trattandosi in fondo di una mera scelta del legislatore. Il fatto però che sia una via praticabile non implica anche che sia una soluzione opportuna[608].

17.3.1. L'assimilazione alle persone giuridiche: una falsa pista.

Attribuire la personalità giuridica ad un'entità non umana non sarebbe una novità per il nostro ordinamento, posto che essa viene pacificamente riconosciuta in capo agli enti[609]. Superando il pregiudizio antropocentrico che, in un primo momento, si era opposto al riconoscimento di questa forma di personalità[610], ad

---

[602] C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 134.
[603] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., pp. 242-243.
[604] U. RUFFOLO, *Il problema della "personalità elettronica"*, cit., p. 76.
[605] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 244, corsivo nostro.
[606] U. RUFFOLO, *Il problema della "personalità elettronica"*, cit., p. 78.
[607] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 28. Se da un lato C. SALAZAR, *Umano, troppo umano,* cit., p. 262 evidenzia come l'iterazione emotiva con sistemi antropomorfi potrebbe migliorare i risultati dei pazienti con disabilità che si interfacciano con l'IA, di contro A. TURANO, *Robotica e roboetica*, cit., p. 140 parla di "inganno robotico", riferendosi alle doti della macchina di simulare le capacità affettive dell'uomo. La degenerazione di questo stato di cose potrebbe condurre alla verificazione di possibili danni, come evidenziano R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 55 i quali parlano di *robot deception.*
[608] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 240: «le soluzioni giuridiche hanno un senso se *servono* a raggiungere finalità date nell'ambito di un certo contesto ordinamentale; tra più soluzioni giuridiche dirette a raggiungere le medesime finalità, deve privilegiarsi quella che minimizzi i costi/rischi per il sistema».
[609] I. SALVADORI, *Agenti artificiali,* cit., p. 97.
[610] R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 267.

oggi la nozione giuridica di "persona" si è andata estendendo, fino ad includere al suo interno non soltanto le persone fisiche ma anche le persone giuridiche[611]. Del pari non costituisce *novum* per l'ordinamento la responsabilizzazione, anche penale, di entità diverse dalle persone fisiche, grazie al d.lgs. 231/2001: «storicamente l'incriminazione delle persone giuridiche era reputata una *fictio iuris*, poiché la sanzione penale si riteneva riservata ai soggetti individuali autori del reato; tuttavia, con il passare del tempo, si è avuto modo di constatare come possa ammettersi una capacità criminale in capo all'ente, imboccandosi pertanto la via di una corresponsabilizzazione dei soggetti collettivi»[612].

Volendo dunque provare a immaginare più concretamente un'ipotetica struttura della personalità elettronica, dovremmo intanto riscontrare che sarebbe indubbiamente più facile costruirla sulla falsa riga delle persone giuridiche, piuttosto che su quella delle persone fisiche[613]. Secondo questo classico schema «sono gli uomini che in realtà fanno muovere nel traffico giuridico le "persone" (maschere) giuridiche»[614]. Tuttavia è proprio in questa preliminare considerazione che risiede la prima obiezione (come invero già accennato) all'assimilazione tra personalità giuridica ed elettronica.

Nel modello organicistico tipico delle persone giuridiche il fulcro centrale resta sempre l'uomo. Esse sono costruite partendo da un presupposto antropocentrico, dunque sull'idea di un insieme di soggetti in carne e ossa che sono, in realtà, i veri autori degli atti giuridici (anche penalmente rilevanti) posti in essere dall'ente. L'immedesimazione che connota il rapporto tra la persona giuridica e le persone fisiche consente di imputar loro gli atti dell'ente anche da un punto di vista psicologico, rispettando dunque il principio di colpevolezza. Per tale ragione la struttura della personalità elettronica non potrebbe essere costruita secondo il modello della personalità giuridica, proprio in quanto nei sistemi intelligenti mancherebbe questa pervasiva presenza umana[615]. Per quanto essa possa essere presente, sarebbe indubbiamente molto più remota, al pari dell'elemento soggettivo sotteso all'illecito commesso[616]: nonostante sia pur vero che comunque alla base dell'agire dell'IA vi sia sempre l'input umano, non possiamo dire che si tratti di una presenza costante, specie se ci riferiamo a sistemi intelligenti in grado di agire con un certo grado di autonomia.

Altra differenza di non poco conto risiede nel fatto che le persone giuridiche, per loro natura, sono enti immateriali. Sono costituiti da uomini, dotati dell'imputabilità necessaria per potergli muovere un rimprovero (anche penale) e in capo ai quali è possibile riscontrare gli elementi psicologici necessari per imputare un dato fatto lesivo. I sistemi intelligenti, invece, sono spesso enti materiali, tangibili, capaci di autodeterminarsi ma sprovvisti di coscienza e non

---

[611] S. RIONDATO, *Robot: talune implicazioni di diritto penale*, cit., p. 90.
[612] P. SEVERINO, *Intelligenza artificiale*, cit., p. 535. Sul tema v. anche F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 29.
[613] G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 111. Su questa distinzione v. anche F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 227.
[614] S. RIONDATO, *Robot: talune implicazioni di diritto penale*, cit., p. 91; A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 27 parla di "maschere della soggettività".
[615] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., pp. 237-238.
[616] U. RUFFOLO, *Il problema della "personalità elettronica"*, cit., p. 83.

necessariamente riconducibili ad uno specifico e individuabile soggetto umano[617]. Esse sono «libere della libertà che è stata loro elargita dal proprio creatore»[618].

È proprio sulla possibilità di muovere un rimprovero che si appunta la terza obiezione all'assimilabilità tra personalità giuridica ed elettronica. Il processo di imputazione della responsabilità in capo alla persona giuridica passa comunque da una persona fisica, che sia in posizione apicale o subordinata, ma pur sempre da un soggetto che può essere destinatario di un rimprovero. L'agire autonomo e imprevedibile dell'IA, invece, renderebbe più difficoltoso ammonire l'uomo che l'abbia *ab origine* programmata, «privando della sua ragion d'essere il rimprovero della macchina in quanto tale»[619].

Una quarta obiezione, invero già accennata[620], si riferisce all'attuale struttura della legge sulla responsabilità amministrativa degli enti. L'art. 5 del d.lgs. 231/2001 chiarisce che l'ente è responsabile per i reati commessi nel suo interesse o a suo vantaggio. Tale principio appare difficilmente estendibile ai reati commessi dall'IA, proprio in ragione del fatto che quest'ultima è in grado di determinarsi autonomamente, secondo processi e logiche che spesso sfuggono ai suoi stessi programmatori[621].

Alla luce delle suesposte considerazioni, non sembrerebbe possibile costruire la personalità elettronica sul modello della personalità giuridica. Volendo insistere su questa via si dovrebbe forse optare per una terza categoria, diversa dalla persona fisica e da quella giuridica, ripensando radicalmente il concetto stesso di soggettività giuridica[622]. Anche qui, come accennato in precedenza, tale via sarebbe anche percorribile, dipendendo da una scelta esclusiva del legislatore. Dovremmo però chiederci a cosa servirebbe riconoscere all'IA una sua personalità giuridica: servirebbe a «riconoscere gli algoritmi sì "come attori capaci di agire in quanto provvisti di capacità giuridica", ma di capacità giuridica limitata, o parziale, "accuratamente calibrat[a] sul ruolo che essi effettivamente svolgono"»[623].

17.4. La diretta responsabilizzazione penale dell'IA (?)

Questa corposa premessa ci conduce al cuore della questione di nostro interesse: posto che «ogni responsabilizzazione diretta, penale come civile o amministrativa, dell'A.I. resta subordinata al riconoscimento d'una "personalità elettronica"»[624], dobbiamo chiederci se e come l'IA possa essere destinataria di

[617] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 241.
[618] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 18, il quale a sua volta si dedica alla differenza tra persone giuridiche come enti "naturalisticamente fittizi" e intelligenze artificiali come entità materialmente e fisicamente esistenti, pp. 17-18.
[619] P. SEVERINO, *Intelligenza artificiale*, cit., p. 535.
[620] Cap. II, Sez. II, Par. 12.
[621] C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 134 nota 77.
[622] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 238. Sull'opportunità di considerare i soggetti intelligenti come "*tertium genus*", dunque come categoria giuridica autonoma, v. C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 5.
[623] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 249 la quale, a sua volta, richiama G. TEUBNER, *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi*, Napoli, 2019, p. 125.
[624] U. RUFFOLO, Machina delinquere potest?, cit., p. 301. Invero occorre evidenziare che, riferendosi ad un contesto più generale, lo stesso A. afferma in altro scritto che «il riconoscimento dello *status* di persona è presupposto indefettibile per il riconoscimento soltanto di (particolari)

divieti penali e, conseguentemente, responsabile per la loro violazione[625]. Dobbiamo chiederci, ancora, se l'IA possa aspirare ad assurgere al rango di "agente morale"[626] ed essere considerata non come mero prodotto bensì come soggetto di diritto, al fine di essere ritenuta responsabile per i danni da questa arrecati[627]. Si tratterà, dunque, di valutare l'opportunità di pensare ad una capacità giuridica e ad una capacità d'agire di "nuova generazione"[628]. I problemi non si arresterebbero qui. Considerare i sistemi intelligenti come muniti di una soggettività autonoma e, dunque, potenzialmente imputabili, porrebbe una serie di problemi: come dovremmo intendere il concetto di capacità soggettiva, di azione e di colpevolezza? Potremmo annoverare l'IA tra i possibili autori di reato, considerandola capace di agire in modo cosciente e volontario? Nel caso rispondessimo negativamente a questa domanda, resterebbe il problema di comprendere chi debba essere considerato responsabile delle azioni commesse dall'IA[629].

In ottica penalistica la personificazione dell'intelligenza artificiale comporterebbe, da un lato, la possibilità di responsabilizzarla, imputandole anche sanzioni tipicamente penali per gli illeciti da essa commessi, e dall'altro, la possibilità di tutelarla dai possibili abusi realizzati ai suoi danni[630]. Con il riconoscimento di una personalità giuridica elettronica (*e-personhood*)[631] l'IA diventerebbe autonomo centro d'imputazione, ben potendo essere chiamata a rispondere degli eventi lesivi da essa cagionati, ciò chiaramente anche in ragione della sua crescente autonomia[632]. Dobbiamo dunque chiederci se i sistemi intelligenti possano essere equiparati all'uomo in punto di attribuzione della responsabilità penale[633]. Per certi versi potrebbe sembrare che ci si stia già muovendo in questa direzione: si parla già di soggettività elementale, di agentività e di nuove attorialità[634]. A tal proposito, per distinguere i "soggetti artificiali" dal classico "attore" umano si usa riferirsi ad essi con il termine "attanti"[635].

---

diritti, e non anche di responsabilità; ed ancor meno di responsabilità patrimoniali. Per le quali, dunque, sono ipotizzabili, ove utili, molteplici forme di (totale o parziale) personificazione giuridica, anche in assenza dello *status* di "persona"» U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 30.

[625] C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 134. Sull'IA come "soggetto di diritto penale" M. SIMMLER, N. MARKWALDER, *Guilty robots – Rethinking the nature of culpability and legal personhood in an age of artificial intelligence*, in *Criminal Law Forum*, 4.12.2018, pp. 16 ss.

[626] M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 1.

[627] G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 109.

[628] C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 7.

[629] M.B. MAGRO, *Biorobotica*, cit., p. 513.

[630] U. RUFFOLO, Machina delinquere potest?, cit., p. 301.

[631] Parla di *e-personhood* anche V. MANES, *L'oracolo algoritmico,* cit., p. 550.

[632] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1764; I. SALVADORI, *Agenti artificiali*, cit., p. 96; P. MORO, *Macchine come noi*, cit., p. 54, il quale aggiunge che il lessico che ci vediamo costretti ad usare in questo contesto, essendo il solo a nostra disposizione, si manifesta del tutto inadeguato a spiegare la "libertà di agire" di una macchina autonoma.

[633] F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 29.

[634] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 25.

[635] «La realtà futura si fonderà infatti sempre di più sull'interazione in *cloud* di oggetti e soggetti, nel vasto mondo dell'*Internet of things*. *Networks* complessi di enti umani e non umani, così, nei quali i singoli attanti si porranno come "nodi" di un più ampio sistema di *distributed intelligence*, annebbieranno ancora di più quegli individuali percorsi causali e di imputazione soggettiva che è

Non stiamo parlando di fantascienza. Nel 2014 è stato messo in funzione il "Random Darknet Shopper", un bot per lo shopping online programmato per effettuare acquisti casuali nel Deep Web una volta a settimana, pagando in Bitcoin. Gli acquisti erano destinati ad una mostra d'arte a Zurigo intitolata "The Darknet - From Memes to Onionland". Tra i vari prodotti acquistati figuravano anche delle pillole di ecstasy, confiscate dalla procura, insieme al Random Darknet Shopper. Dopo pochi mesi il pubblico ministero fece cadere le accuse ai danni dei creatori del bot, giustificando le finalità artistiche da loro perseguite e restituendogli il sistema intelligente precedentemente confiscato[636]. Tralasciando le finalità artistiche che hanno portato al ritiro delle accuse, tale esempio è sintomatico del processo di deresponsabilizzazione[637] che può riguardare i creatori di forme di IA autonome e indipendenti come il RDS[638].

Si tratta di situazioni in cui viene a mancare una figura umana cui poter imputare la responsabilità per l'evento lesivo causato dall'IA proprio perché, all'iniziale programmazione effettuata *ex ante* dal soggetto umano, si aggiunge la capacità dell'IA di imparare dall'esperienza e di *agire* autonomamente in modo incomprensibile *ex post* dall'osservatore umano. È proprio sul concetto di "azione" che vorremmo in tal sede soffermarci. L'IA si caratterizza per la sua interattività, autonomia e adattività, caratteristiche che la rendono in grado di reagire agli stimoli esterni, cambiando e migliorando le sue reazioni. L'insieme di questi connotati produce, conseguentemente, l'imprevedibilità delle azioni dell'IA[639]. La questione della capacità penale d'agire dell'intelligenza artificiale dipende, a ben vedere, dal concetto di azione cui si decide di far riferimento. Accogliendo una nozione "causale" di azione – con essa intendendo ogni «"movimento volontario[640] del corpo"» – sarebbe possibile considerare i sistemi intelligenti come veri e propri agenti. Assumendo, invece, una nozione

---

necessario tracciare con nettezza, collegando un agente determinato a un certo evento, per poterglielo correttamente attribuire a titolo di responsabilità penale» A. CAPPELLINI, Machina delinquere non potest?, cit., p. 10. U. RUFFOLO, Machina delinquere potest?, cit., p. 300 li definisce come «"*agenti software autonomi*", cui attribuire idoneo "*status privatistico*" quali "*soggetti giuridici digitali*"». Sul tema, per tutti, G. TEUBNER, *Ibridi ed attanti. Attori collettivi ed enti non umani nella società e nel diritto*, Milano, 2015.

[636] F. LAGIOIA, G. SARTOR, *AI Systems Under Criminal Law: a Legal Analysis and a Regulatory Perspective*, in *Philos. Technol*., 2020, p. 452.

[637] Cap. II, Sez. II, Par. 12.

[638] Qualche altro esempio di sistemi di IA dotati di una consistente autonomia e utilizzati per delinquere (specialmente in abito finanziario) sono rinvenibili in R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 263. L'A. propone l'esempio dei social bot usati per il *pump and dump*, ossia una «frode che consiste nel fare lievitare artificialmente il prezzo di un titolo, mediante dichiarazioni false, fuorvianti o esagerate, con l'obiettivo di vendere titoli acquistati a buon mercato ad un prezzo superiore» e degli agenti commerciali artificiali usati per lo *spoofing* finanziario, ossia «piazzare ordini, in modo continuativo per un certo periodo di tempo, senza avere l'intenzione di eseguirli, al fine di manipolare i prezzi di mercato».

[639] U. PAGALLO, *Saggio sui robot e il diritto penale*, cit., p. 602.

[640] In realtà, a nostro avviso, il requisito della volontarietà sarebbe già di per sé idoneo a risolvere a monte il problema, in quanto non si può (ancora) dire che l'IA sia dotata di una propria volontà. Ogni azione posta in essere dall'IA si muove pur sempre (quantomeno in una fase iniziale) nel solco tracciato dal programmatore. Né, sempre a nostro avviso, varrebbe obiettare che essa è in grado di imparare dall'esperienza in quanto non sembra possibile considerare l'azione autonoma dell'IA come vera e propria manifestazione di volontà proprio perché, in ogni caso, essa si muove pur sempre entro programmi a monte predeterminati, "limitandosi" (grazie all'autoapprendimento) a migliorarli.

"finalistica" di azione, essa esprimerebbe una «volontà intenzionale e cosciente dell'attore» che non sembrerebbe possibile ricondurre all'IA[641].

Dedurre l'intelligenza della macchina soltanto dalla sua condotta esterna e reattiva costituisce metodo d'indagine che trae le sue fondamenta dalla teoria psicologica del comportamentismo, la quale riposa sull'assunto per cui «il comportamento esplicito è l'unica unità di analisi scientificamente studiabile, in quanto direttamente osservabile, mentre l'introspezione non può fornire alcun dato affidabile»[642]. Tale considerazione, tuttavia, non appare compatibile con il concetto di azione accolto dal diritto penale il quale, secondo una concezione classica del reato, considera l'azione come «un movimento corporeo *cosciente e volontario*»[643] dell'uomo, da attribuirsi ad esso oggettivamente e soggettivamente. Questa considerazione comincia a palesare un primo indizio in ordine alle difficoltà di attribuire direttamente l'azione criminosa ad un sistema intelligente, pre-programmato per svolgere un compito (per quanto autonomo che sia) e a cui non sembra che si possano riferire i concetti di coscienza e volontà richiesti dal nostro codice penale (ex art. 42 co. 1 c.p.).

Altro indizio delle summenzionate difficoltà si rinviene dal punto di vista del trattamento sanzionatorio. La via della punizione diretta dell'IA è stata in astratto ritenuta percorribile, pensando a sanzioni consistenti in «(a) monitoraggio e modifica (cioè "manutenzione"); (b) rimozione di una componente disconnessa del Cyberspazio; (c) annientamento dal Cyberspazio (cancellazione senza backup)»[644]. Tale tipo di sanzioni potrebbe in astratto ricadere su un sistema intelligente, essendo possibile che quest'ultimo sia dotato di un "corpo" che potrebbe essere disattivato o financo distrutto[645]. La funzione punitiva tipica del diritto penale, inoltre, può essere raggiunta non soltanto mediante un effetto di deterrenza o di neutralizzazione, bensì anche attraverso la sanzione pecuniaria. A tal proposito, infatti, argomentare che l'IA non possa essere ritenuta responsabile in quanto non possiede "*a body to kick, and a soul to be damned*" sarebbe, a maggior ragione, affermazione contraddittoria, ove si considerasse l'idea di responsabilizzare la macchina da un punto di vista risarcitorio[646].

Vorremmo però brevemente far notare che entrambe le vie non sembrerebbero percorribili. Infliggere una sanzione "corporea" all'IA in realtà non la affliggerebbe direttamente, non essendo quest'ultima in grado di "sentire" la pena e, men che meno, di comprenderne il significato. La sanzione sarebbe piuttosto rivolta all'uomo dietro l'IA, il quale subirebbe il pregiudizio economico connesso alla distruzione del suo "bene intelligente"[647]. Quanto all'aspetto risarcitorio è stato fatto notare che l'IA sarebbe di per sé priva di una fonte patrimoniale cui poter attingere per risarcire i danni causati, il che renderebbe a

---

[641] M.B. MAGRO, *Robot*, cit., p. 1203. Tale impostazione sembra latamente ricordare la distinzione tra diritto penale "oggettivo puro" e diritto penale "soggettivo puro" riproposta da N. MAZZACUVA, *Alcune riflessioni su intelligenza artificiale e diritto penale sostanziale,* cit., pp. 287 ss.

[642] P. MORO, *Biorobotica e diritti fondamentali*, cit., p. 540.

[643] G. FIANDACA, E. MUSCO, *Diritto Penale*, cit., p. 232.

[644] L. FLORIDI, J.W. SANDERS, *On the morality of artificial agents*, in *Minds and Machine*, 2004, p. 368.

[645] M.B. MAGRO, *Biorobotica*, cit., p. 514.

[646] U. RUFFOLO, Machina delinquere potest?, cit., p. 295.

[647] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 18.

questo punto superfluo attribuirle una personalità giuridica. Il punto è che, anche se l'IA fosse dotata di un suo patrimonio, tale fondo pecuniario sarebbe pur sempre alimentato da soggetti umani, i quali, in conclusione, sarebbero i veri destinatari della sanzione formalmente inflitta all'IA[648].

Prescindendo dalle malcelate perplessità in ordine al riconoscimento di una personalità elettronica volta a rendere l'IA direttamente e penalmente responsabile dei danni da essa cagionati, ci sentiamo in astratto di poter ritener vero l'assunto secondo cui il riconoscimento dell'IA come centro di imputazione giuridica sarà strettamente connesso alla pregnanza della visione antropocentrica dell'uomo rispetto alle sue interazioni con la macchina[649]. «Solo se si continua ad essere permeati di cultura antropocentrica, per attribuire la "personalità elettronica" potrebbe non bastare il pensiero artificiale, anche *self-learning*, ritenendosi invece necessaria l'autocoscienza (il *cogito ergo sum* cartesiano) o i requisiti di soggetto pensante»[650].

17.4.1. Una via difficilmente percorribile.

Non dovrebbe forse sorprendere che l'accenno alla possibilità di riconoscere una personalità elettronica ai sistemi intelligenti fatto nel 2017 dalla Risoluzione del Parlamento europeo sembri essere rimasto tentativo isolato[651]. E questo, riteniamo, non solo perché l'Unione Europea non ha il potere di intervenire in merito, essendo rimesso agli ordinamenti dei singoli Stati membri il compito di determinare il concetto giuridico di "persona" che essi intendono adottare[652]. Il vero motivo per il quale, a nostro avviso, questa proposta non ha trovato un seguito è da imputare alla reazione generale che ne è derivata. Un gruppo di quasi 300 esperti di intelligenza artificiale ha, ad oggi, sottoscritto una lettera aperta indirizzata alla Commissione Europea in cui vengono chiariti i motivi per i quali attribuire una personalità giuridica elettronica ai sistemi intelligenti sarebbe inappropriato[653]. Il Comitato economico e sociale europeo, dal canto proprio, ha

---

[648] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., pp. 347-348.

[649] S. RIONDATO, *Robotica e diritto penale*, cit., p. 602; ID., *Robot: talune implicazioni di diritto penale*, cit., p. 90.

[650] U. RUFFOLO, *Il problema della "personalità elettronica"*, cit., pp. 78-79.

[651] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 246. M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 30 ci ricorda che tale proposta era invero presente nel Rapporto Delvaux il quale ha preceduto l'entrata in vigore della Risoluzione del Parlamento europeo del 2017 riguardante le norme di diritto civile sulla robotica.

[652] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 228.

[653] «Da un punto di vista etico e legale, creare una personalità giuridica per un robot è inappropriato qualunque sia il modello di status giuridico: a. Uno status giuridico per un robot non può derivare dal modello delle persone fisiche, poiché il robot sarebbe quindi titolare di diritti umani, come il diritto alla dignità, il diritto alla sua integrità, il diritto alla remunerazione o il diritto alla cittadinanza, confrontandosi così direttamente con i diritti umani. Ciò sarebbe in contraddizione con la Carta dei diritti fondamentali dell'Unione europea e la Convenzione per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali. b. Lo status giuridico di un robot non può derivare dal modello delle persone giuridiche, poiché implica l'esistenza di persone umane dietro la persona giuridica che lo rappresenti e lo diriga. E questo non è il caso di un robot. c. Lo status giuridico di un robot non può derivare dal modello del Trust anglosassone chiamato anche *Fiducie* o *Treuhand* in Germania. In effetti, questo regime è estremamente complesso, richiede competenze molto specializzate e non risolverebbe il problema della responsabilità. Ancora più importante, implicherebbe comunque l'esistenza di un essere umano come ultima

definito l'introduzione di una forma di personalità giuridica per l'intelligenza artificiale come «un rischio inaccettabile di azzardo morale»[654]. Insomma, sembra che tale proposta abbia riscosso più critiche che elogi.

In molti hanno ravvisato perplessità in ordine al riconoscimento di questa specifica personalità giuridica elettronica[655], fino ad affermare che essa «introdurrebbe, se attuata, una mera *fictio* del tutto priva di alcuna sottostante sostanza giuridica, rischiando di costituire una soluzione a un falso problema e, in quanto tale, di divenire a sua volta fonte di equivoci interpretativi e di applicazioni abnormi, nonché di cagionare essa stessa quei "vuoti di responsabilità" che la proposta dovrebbe essere volta a colmare»[656].

Parlare di "soggettività non umana" secondo una tesi funzionalista non tiene conto del fatto che, come abbiamo già evidenziato, la capacità di un sistema intelligente di interagire con l'esterno e con l'uomo non implica che esso sia cosciente di ciò che sta facendo, trattandosi di una mera simulazione non equiparabile all'agire umano[657]. Un simile paragone tra uomo e macchina sarebbe fondato su una visione molto riduttiva della soggettività umana[658], spogliata di tutte le sue peculiarità esclusive e non meccanizzabili. Su questa via si sono incardinate le prime obiezioni al riconoscimento di una personalità giuridica in capo all'IA[659] e si sono sviluppate le principali critiche intorno al paragone tra l'intelligenza artificiale e la coscienza umana. In dottrina è stato fatto notare che equiparare l'intelligenza della macchina a quella dell'uomo implicherebbe una svalutazione delle capacità umane. Se l'uomo e il sistema intelligente hanno in

---

risorsa – il fiduciario o il fiduciario – responsabile della gestione del robot concesso con un Trust o un *Fiducie*», lettera consultabile al sito http://www.robotics-openletter.eu/.

[654] *Parere del Comitato economico e sociale europeo su «L'intelligenza artificiale — Le ricadute dell'intelligenza artificiale sul mercato unico (digitale), sulla produzione, sul consumo, sull'occupazione e sulla società»*, (2017/C 288/01), Punto 3.33.

[655] G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 9; C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 135 chiarisce che siamo ancora distanti dal riconoscimento di una "coscienza artificiale"; A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 27. V'è anche chi, pur ravvisando delle perplessità, mostra un margine di apertura, considerando il riconoscimento di tale personalità, semplicemente, come "non necessario": «Non è necessario, per responsabilizzarla, che la macchina quale "corpo" abbia un "cervello" e "senta" come quello umano. Non è, in altri termini, indispensabile la transizione da bene a persona (non lo era necessariamente neppure per lo schiavo quando dotato di un qualche "peculio"). Si può essere "responsabili", e titolari di risorse patrimoniali, anche senza avere personalità giuridica, comunque senza dover necessariamente ricevere la equiparazione allo *status* della persona umana» U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 28.

[656] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 97. La "proposta" cui l'A. fa riferimento nel corpo del testo è la *Risoluzione del Parlamento europeo*, cit., Punto 59, lett. f).

[657] M.B. MAGRO, *Robot*, cit., pp. 1196-1201.

[658] P. MORO, *Macchine come noi*, cit., p. 47.

[659] Le prime perplessità in tal senso risalgono agli anni 90 del secolo scorso e si appuntavano essenzialmente su tre considerazioni: a) "*AIs are not humans*": si rivendica un approccio antropocentrico, principalmente connesso alle difficoltà di attribuire all'IA i diritti costituzionali tipici dell'uomo; b) "*The Missing-Something Argument*": l'IA difetterebbe di alcuni elementi fondamentali della persona umana, quali l'anima, la coscienza, l'intenzionalità, le emozioni, gli interessi propri e il libero arbitrio. Inoltre si evidenzia come l'output prodotto dall'IA si limiti ad imitare il comportamento umano, essendo frutto di una mera simulazione; c) "*AIs ought to be property*": l'IA è un artefatto e non dovrebbe essere considerato come qualcosa di più di una mera proprietà del suo produttore, così L.B. SOLUM, *Legal Personhood for Artificial Intelligences*, cit., pp. 1258 ss.

comune una spiccata capacità logico computazionale (in cui invero è la macchina a superare l'uomo), di contro l'IA non è (al momento) dotata di connotati tipicamente ed esclusivamente umani. L'intelligenza artificiale, infatti, non sarebbe in grado di riprodurre un procedimento mentale simile alla nostra "intuizione", né sarebbe dotata di autocontrollo e autocoscienza, proprio in quanto si tratta di processi intellettuali che non sono formalizzabili o codificabili. Il pensiero artificiale, inoltre, a differenza del pensiero umano, avrebbe pur sempre un'origine "derivata", in quanto prodotto dell'uomo. L'IA, ancora, difetterebbe di una dimensione dialogico-relazionale tipica delle interazioni tra le persone, per questo ci si interroga sul futuro delle interazioni uomo-macchina, specie considerando che l'introduzione di quest'ultima nel nostro mondo esterno rappresenta pur sempre un fattore di rischio potenzialmente foriero di danni[660].

L'IA non ha la consapevolezza del proprio agire e non è conscia di ciò che fa, per questo motivo i suoi processi logico-computazionali, per quanto possano apparire dall'esterno simili al pensiero umano, in realtà non sono ad esso assimilabili[661]. Tale assenza di consapevolezza costituisce, ad avviso di alcuni, il principale ostacolo al riconoscimento di diritti e doveri in capo all'IA. La questione potrebbe essere riconsiderata solo nel momento in cui l'IA dovesse riuscire a sviluppare "dimensioni tipicamente umane"[662] che potrebbero (a quel punto, sì) giustificare un'assimilazione tra uomo e macchina. Per il momento, non resta che prendere atto anche dell'opinione di chi ritiene che non sia possibile attribuire alle macchine diritti e doveri, per quanto intelligenti e autonome che siano[663].

In punto di responsabilizzazione diretta dell'IA, a nostro avviso, l'obiezione risolutiva potrà sembrare tanto efficace quanto banale: la responsabilità penale si fonda sull'imputabilità del soggetto agente, intesa come capacità di intendere e di volere ex art. 85 c.p., capacità di cui l'IA ad oggi non sembra (poter essere) provvista[664]. Parte della dottrina ha, a tal proposito, evidenziato che discutere di una diretta imputabilità dell'IA non avrebbe senso, essendo quest'ultima sprovvista della coscienza e dell'intenzionalità delle proprie azioni, requisiti imprescindibili per poter imputare ad un soggetto la responsabilità penale per il suo comportamento[665].

Col trascorrere del tempo l'uomo sembra esser diventato unico possibile attore[666], individualmente o in forma associata, dietro lo schermo della "persona giuridica". Né varrebbe obiettare che nel secolo corrente stiamo assistendo a una

---

[660] P. MORO, *Biorobotica e diritti fondamentali*, cit., pp. 541 ss. Tra le ragioni di dubbio concernenti l'assimilazione dell'intelligenza umana e artificiale l'A. riporta anche la pretesa neutralità della scienza ed il fatto che «la cibernetica continua ad essere in contrasto con la natura mai completamente prevedibile e programmabile dell'attività sviluppata dall'intelligenza e dalla volontà dell'uomo».

[661] M.B. MAGRO, *Decisione umana e decisione robotica*, cit., pp. 18-19. Per la distinzione tra coscienza di accesso o cognitiva e coscienza fenomenica p. 16. L'A. evidenzia altresì che i sistemi intelligenti sarebbero dotati di una «solidità esistenziale [che] li allontana da tutto ciò che è vivente, a prescindere dalle loro *performance* più o meno equivalenti a quelle umane» p. 15.

[662] M.B. MAGRO, *Robot*, cit., pp. 1202-1203.

[663] P. MORO, *Macchine come noi*, cit., p. 53.

[664] A. CELOTTO, *I robot possono avere diritti?*, cit., p. 97.

[665] U. PAGALLO, *Saggio sui robot e il diritto penale*, cit., p. 600.

[666] Sono lontani i tempi in cui venivano celebrati processi a carico degli animali, U. PAGALLO, *Saggio sui robot e il diritto penale*, cit., p. 604.

sempre maggiore personificazione di entità non umane, quali gli animali, in quanto il riconoscimento di diritti in capo a questi ultimi – al fine di predisporre un meccanismo di tutela contro gli abusi che la società umana può porre in essere ai loro danni – non equivale ad una personificazione dell'IA volta a creare (a contrario) nuovi e autonomi "centri aggressivi d'azione"[667]. Il riconoscimento di una personalità giuridica elettronica in capo alle più evolute forme di intelligenza artificiale avrebbe come effetto la creazione di uno "schermo" limitativo della responsabilità delle persone dietro l'IA, cioè coloro i quali la programmano o se ne servono per i propri scopi[668].

La prospettiva di una diretta responsabilizzazione penale dell'IA trova, infine, un ultimo ostacolo nelle teorie della pena. Anche prescindendo dalla precedente obiezione e volendo ritenere il sistema intelligente colpevole del suo agire, come potrebbe quest'ultima «mai pagare il suo debito nei confronti della società? In che modo la sua punizione potrebbe riformarlo al fine di non compiere nuovamente simili delitti? Ovvero, secondo i dettami della pena in funzione della prevenzione generale, in che modo la condanna inflitta al robot potrebbe rappresentare un avvertimento, o una minaccia, per la generalità dei consociati?». Si tratterebbe sempre di una responsabilità tra uomini: la "non responsabilizzazione" dell'IA dipenderebbe proprio dall'incoscienza delle proprie azioni, dalla sua incapacità di "volere" un certo comportamento. L'assenza di consapevolezza non consente, ad oggi, di considerare l'IA come soggetto giuridicamente imputabile e dunque, per quanto essa possa essere foriera di danni, difficilmente potrà essere riconosciuta come "soggetto colpevole"[669].

Dunque, se assumiamo per vero che l'intelligenza artificiale non possa essere direttamente e penalmente responsabile dei danni da essa commessi, non resterà altra via che chiamare a rispondere le persone dietro l'IA. Muovere un rimprovero penale a questi ultimi non sarà del pari semplice, specie in considerazione dell'agire autonomo e imprevedibile dell'IA, difficilmente riconducibile alla colpevolezza di colui il quale l'ha programmata o se ne è servito. Per tale ragione c'è da aspettarsi che, negli scenari futuribili che potrebbero venire a crearsi, si prediligerà la via del rimedio risarcitorio il quale, purché adeguato, potrebbe evitare la creazione di zone franche dietro le quali possa celarsi un inammissibile vuoto di tutela, adducendo come giustificazione l'imprevedibilità del comportamento dell'IA[670].

Tornano in tal sede utili le proposte avanzate dalla Risoluzione del Parlamento europeo recante "Norme di diritto civile sulla robotica", la quale suggerisce l'istituzione di un regime assicurativo *ad hoc* e di un fondo patrimoniale volto a garantire quantomeno il risarcimento per i danni causati dall'IA[671]. Conseguentemente, le sanzioni formalmente dirette all'intelligenza artificiale, in realtà, resterebbero concretamente dirette all'uomo dietro l'IA il

---

[667] U. PAGALLO, *Saggio sui robot e il diritto penale*, cit., pp. 605-606.
[668] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 226.
[669] La presente e, a nostro avviso, ineccepibile ricostruzione è rinvenibile in U. PAGALLO, *Saggio sui robot e il diritto penale*, cit., p. 600.
[670] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 355.
[671] *Risoluzione del Parlamento europeo*, cit., Punti 57-58-59. Tali proposte vengono recepite in dottrina da C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 7; G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 110; A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 97.

quale alimenta questo fondo[672]. È stato però correttamente osservato in dottrina che affidare le esigenze riparatorie ad un fondo alimentato dai proprietari dell'IA rischierebbe di rendere l'obbligo risarcitorio non più connesso alla responsabilità colpevole bensì al solo verificarsi dell'evento causalmente riconducibile all'agire dell'IA, finendo per chiamare nuovamente in causa forme (non troppo occulte) di responsabilità oggettiva[673]. Nelle appena svolte considerazioni trova conferma quanto abbiamo già avuto modo di evidenziare nelle pagine precedenti: il diritto civile lambisce il diritto penale fino quasi a confondersi con esso il quale, in una materia così nuova e peculiare, mostra tutti i suoi limiti, rischiando di apparire come "strumento vecchio"[674].

Volendo sintetizzare. Potremmo anche in astratto riconoscere una personalità giuridica all'IA al fine di imputarle la responsabilità penale delle sue azioni ma, prima, dovremmo chiederci a cosa servirebbe. Un sistema intelligente non può essere considerato in grado di porre in essere un'"azione" intesa in senso penalistico. Né da un punto di vista oggettivo, in quanto il suo agire dipende comunque da una precedente creazione e programmazione diretta ad un fine stabilito dall'uomo. Men che meno da un punto di vista soggettivo, in quanto non ci sentiamo di affermare che (ad oggi) i sistemi intelligenti siano in grado di integrare la coscienza e volontà richiesta dal codice penale ai fini dell'imputazione dell'azione criminosa ex art. 42 co. 1 c.p. Dobbiamo conseguentemente dedurne che la responsabilità penale non può essere imputata all'IA in quanto l'azione che essa è in grado di realizzare non rientra nel concetto di "azione" tipico del diritto penale, e questo non solo in quanto la sua condotta non potrebbe essere considerata colpevole, ma anche perché l'IA non percepirebbe il disvalore della sua azione e non comprenderebbe il significato della relativa sanzione. Si potrebbe al più produrre un ristoro economico per i danni da essa causati attingendo ad un fondo riferito all'IA medesima (magari con capienza limitata) ma quest'ultimo sarebbe pur sempre alimentato dal suo produttore o dal suo proprietario, quindi in ogni caso finiremmo sempre per punire l'uomo, peraltro limitandone la responsabilità. Se da un lato tale limitazione potrebbe essere positivamente accolta nei casi in cui l'IA abbia agito in quasi totale autonomia, in modo dunque non riconducibile ad una pregressa e ben individuabile azione umana, di contro potrebbe costituire un vuoto (seppur parziale) di responsabilità nel caso in cui venga fatto un uso strumentale dell'IA a fini delittuosi che venga dal reo giustificato dietro lo schermo fittizio dell'"incidente" causato dall'agire imprevedibile dell'IA, limitando dunque la sua responsabilità ad un mero ristoro pecuniario connesso al fondo stanziato per il sistema intelligente.

[672] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1765; M.B. MAGRO, *Robot,* cit., p. 1205.
[673] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 16. Più avanti l'A. chiarisce che, peraltro, il fondo patrimoniale potrebbe servire a ristorare il pregiudizio della vittima nel caso in cui il responsabile sia incapiente o ove sia impossibile individuarlo, magari imponendo un prelievo fiscale connesso alla proprietà dell'IA o al suo utilizzo, volto proprio ad alimentare un fondo con il quale risarcire o indennizzare i soggetti lesi, p. 27.
[674] Cap. II, Sez. II, Parr. 10-11.

17.5. Il pensiero di Gabriel Hallevy.

Che non si tratti di astratte congetture è dimostrato dal fatto che in dottrina vi è stato chi ha fermamente sostenuto la possibilità di una diretta responsabilizzazione delle macchine. Ci stiamo evidentemente riferendo all'opinione di Gabriel Hallevy[675]. Vorremmo qui dedicare uno spazio autonomo alle teorie del noto Autore, anche a costo di riproporre per cenni alcuni passaggi evidenziati nelle pagine precedenti, in quanto esse si fondano sul presupposto del riconoscimento di una personalità giuridica all'IA, dato del quale abbiamo voluto dar conto prima di approdare al pensiero di Hallevy. Crediamo infatti che, occupandosi del rapporto tra intelligenza artificiale e diritto penale, non si possa prescindere da una seppur breve disamina del pensiero dell'Hallevy e delle specifiche obiezioni mosse alle sue teorie.

Ad avviso dell'Autore sarebbe possibile distinguere tre modelli di imputazione penale dell'intelligenza artificiale[676]:

- *Perpetration-via-Another*: tale modello imputativo considera l'IA come un agente innocente, non riconducendo ad essa alcun attributo umano. I sistemi intelligenti vengono considerati come meri strumenti nelle mani del reale autore del reato che sarà individuato nel programmatore o nell'utilizzatore dell'IA. Dunque, pur essendo l'offesa materialmente realizzata dall'IA, il modello della perpetrazione del reato attraverso un altro soggetto consente di ricondurre la responsabilità all'uomo che programma o si serve dell'IA. L'Autore chiarisce cha tale modello imputativo può venire in gioco quando ci si confronta con le forme meno sviluppate di intelligenza artificiale, sprovviste delle avanzate capacità autonome dell'IA e, pertanto, considerabili come meri strumenti nelle mani dell'uomo. Tale meccanismo, infatti, entra in crisi quando ci si riferisce ad evolute forme di IA che siano in grado di autodeterminare il proprio agire grazie alla propria esperienza. Sembra pertanto che tale modello d'imputazione sia riconducibile alla responsabilità indiretta dell'uomo dietro la macchina e alle considerazioni svolte nelle pagine precedenti a proposito dell'art. 111 c.p.[677].

- *Natural-Probable-Consequence Liability Model*: tale meccanismo imputativo viene in gioco quando il programmatore o l'utilizzatore non intendono servirsi dell'IA per cagionare un'offesa che, però, si realizza nel corso dell'agire dell'intelligenza artificiale. Tale modello si fonda sulla negligenza dell'uomo dietro l'IA il quale, con la dovuta diligenza, avrebbe potuto prevedere l'evento in

---

[675] Per la bibliografia essenziale dell'Autore si rinvia a G. HALLEVY., *Dangerous Robots – Artificial Intelligence vs. Human Intelligence*, in *SSRN*, 2018; G. HALLEVY, *Liability for Crimes Involving Artificial Intelligence Systems*, Svizzera, 2015; G. HALLEVY, *AI v. IP. Criminal Liability for IP Offences of AI Entities*, in *SSRN*, 2015; G. HALLEVY, *Unmanned Vehicles: Subordination to Criminal Law under the Modern Concept of Criminal Liability*, in *SSRN*, 2012; G. HALLEVY, *Virtual Criminal Responsibility*, in *SSRN*, 2011, pp. 1 ss.; G. HALLEVY, *I, Robot – I, Criminal — When Science Fiction Becomes Reality: Legal Liability of AI Robots committing Criminal Offenses*, in *Syracuse Journal of Science and Technology Law*, 2010, pp. 1 ss.; G. HALLEVY, *The Criminal Liability of Artificial Intelligence Entities – From Science Fiction to Legal Social Control*, cit., pp. 171 ss.

[676] La ricostruzione che segue è rinvenibile in G. HALLEVY, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, cit., pp. 179 ss. Si dedicano all'analisi dei modelli imputativi di Hallevy anche M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., pp. 365 ss. e R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 267.

[677] Cap. II, Sez. III, Par. 16.2.

quanto possibile conseguenza di una certa condotta. L'Autore distingue a tal proposito un uso lecito ed un uso illecito del sistema intelligente. Nel primo caso ci troveremo davanti ad un caso di "negligenza pura", in cui si riscontra un errore in fase di programmazione o di uso del sistema intelligente, al quale non sarà possibile muovere alcun rimprovero. Il secondo caso si verifica, invece, quando l'IA viene programmata o utilizzata per cagionare un danno ma essa, in fase d'esecuzione, devia dal piano pre-programmato e commette un'offesa diversa o un'offesa ulteriore rispetto a quella programmata. Tale ragionamento sembra evocare il su esaminato modello dell'*aberratio delicti*[678], con la sola differenza che, in tal caso, l'Autore ritiene possibile muovere un rimprovero tanto all'uomo dietro l'IA quanto all'IA medesima. In ragione del cumulo delle responsabilità che si verrebbe a creare e della deviazione posta in essere dall'IA, l'uomo potrebbe chiamare in causa il reato diverso da quello voluto da taluno dei concorrenti ex art. 116 c.p. sul quale abbiamo già svolto qualche considerazione (invero preclusiva) nelle pagine precedenti[679].

- *Direct Liability Model*: l'Autore individua gli elementi costitutivi del reato nell'*actus reus* (elemento esterno) e nella *mens rea* (elemento interno) e ritiene che l'IA sia in grado di integrarli entrambi al fine si risultare diretta destinataria di un'imputazione penale. Pochi problemi porrebbe il requisito dell'*actus reus*, in quanto l'IA dispone di componenti mobili in grado di muoversi nello spazio: il semplice azionarsi di un braccio robotico varrebbe ad integrare un'ipotesi di condotta attiva. Ancor più semplice sarebbe l'ipotesi della condotta omissiva, in quanto non verrebbe richiesto all'IA (sulla quale dovrebbe gravare un dovere d'agire) alcun tipo di atto[680]. L'Autore individua la vera sfida nell'attribuire al sistema intelligente l'elemento della *mens rea*. Egli usa a tal proposito riferirsi al *general intent*, inteso come l'elemento soggettivo fondamentale, il quale incarna l'idea stessa della colpevolezza. Il *general intent* si compone, a sua volta, degli elementi della *cognition* e della *volition* (i quali ricordano tanto l'elemento intellettivo e volitivo del dolo). L'elemento *cognitivo* ingloba l'elemento della consapevolezza, intesa come la percezione sensoriale dei dati fattuali e la loro comprensione. La consapevolezza, a sua volta, si dividerebbe in due fasi: l'assorbimento dei dati fattuali tramite i sensi (*rectius* sensori) e l'elaborazione dei dati acquisiti, creando una generale rappresentazione della realtà. Per quanto concerne l'elemento volitivo l'Autore si concentra in particolare sul requisito dell'*intent*, inteso come il più intenso livello di volontà. Si dà a tal proposito conto delle difficoltà che spesso si incontrano nel dimostrare il requisito dell'*intent* e del fatto che, per tale ragione, molti ordinamenti si servono della presunzione legale della *foreseeability rule*: si presume, cioè, che un soggetto tenda al risultato lesivo se, durante la consapevole realizzazione della condotta, è possibile affermare che egli abbia previsto il verificarsi dell'evento con una probabilità particolarmente elevata. La *foreseeability rule* funzionerebbe nello stesso modo per l'IA, il cui

[678] Cap. II, Sez. III, Par. 16.1.
[679] Cap. II, Sez. III, Par. 17.2.1.
[680] L'A. accoglie una definizione di "atto" che si concentra sugli aspetti fattuali dell'atto stesso e non coinvolge, invece, gli aspetti mentali, G. HALLEVY, *Liability for Crimes Involving Artificial Intelligence Systems*, cit., pp. 60 ss.

comportamento sarebbe considerato (presuntivamente) intenzionale in quanto orientato verso un obiettivo[681].

Nel contesto che si è cercato di delineare, l'IA viene ritenuta in grado di integrare tanto l'elemento esterno quanto quello interno del reato e la responsabilità penale di quest'ultima andrebbe scissa da quella dell'uomo, ben potendo, piuttosto, combinarsi con essa. Precipitato logico-giuridico di tale ricostruzione sarebbe il naturale riconoscimento della responsabilità criminale in capo all'IA[682].

L'Autore chiarisce l'importanza di non considerare tali modelli in via alternativa, ben potendo essere applicati in combinazione tra di loro al fine di delineare un quadro completo delle forme di responsabilità penale che possono venire in gioco quando vi sia un diretto coinvolgimento dei sistemi intelligenti. Da ultimo viene inoltre chiarito che, con gli opportuni correttivi, sarebbe perfino possibile sanzionare direttamente il sistema intelligente con le medesime pene previste per l'uomo[683]. Ad avviso di Hallevy, in conclusione, la responsabilità penale potrebbe essere riconosciuta senza troppi ostacoli all'intelligenza artificiale. Nell'affermare ciò si adduce che, dopo tutto, non poche riserve erano state mosse, in passato, nei confronti della penalizzazione delle persone giuridiche e che, nonostante ciò, sono state sviluppate soluzioni legali per riconoscere la responsabilità penale ad entità senza corpo e senza anima. Ad avviso dell'Autore, infatti, non vi sarebbe una sostanziale differenza tra la criminalizzazione degli enti e quella dell'intelligenza artificiale, specie considerando che «tempi moderni

---

[681] La ricostruzione sul *general intent* è rinvenibile in G. HALLEVY, *Liability for Crimes Involving Artificial Intelligence Systems*, cit., pp. 82 ss. L'A. riporta in tal sede l'esempio del computer programmato per giocare a scacchi. Questi computer «hanno la capacità di analizzare lo stato attuale del gioco in base alla posizione dei pezzi sulla scacchiera. Valutano tutte le opzioni possibili per decidere la mossa successiva. Per ogni opzione vengono vagliate le possibili reazioni dell'altro giocatore. Per ogni reazione si valutano tutte le reazioni possibili, e così via fino all'eventuale mossa finale che termina con la vittoria di un giocatore. Ciascuna delle opzioni viene valutata in base alla probabilità della sua realizzazione e, di conseguenza, il computer decide la sua prossima mossa. Se si trattasse di un umano diremmo che agisce con l'intenzione di vincere la partita. Non sapremmo con certezza se egli avesse tale intenzione ma la sua condotta corrisponderebbe alla presunzione della regola di prevedibilità. La tecnologia di intelligenza artificiale, che è programmata per giocare a scacchi, ha un comportamento orientato all'obiettivo di vincere le partite di scacchi. I giocatori di scacchi umani hanno anche un comportamento orientato all'obiettivo di vincere tali partite. Per i giocatori umani si può dire che hanno l'intento di vincere le suddette partite. Sembra che sia possibile affermare ciò non solo con riferimento ai giocatori umani ma anche con riguardo ai giocatori dotati di intelligenza artificiale. L'analisi del loro comportamento nelle situazioni prese in esame corrisponde esattamente alla presunzione della regola di prevedibilità» (traduzione nostra).

[682] L'A. chiarisce inoltre che sarebbe possibile ricondurre ai sistemi intelligenti non solo gli elementi positivi del reato ma anche quelli negativi, ossia le cause di giustificazione. Sul punto v. G. HALLEVY, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, cit., p. 192.

[683] L'A. riporta l'esempio della pena capitale, del carcere, del servizio civile e perfino del pagamento di una multa: la pena capitale si realizzerebbe con l'eliminazione del software; l'incarcerazione mettendo fuori uso l'IA per un determinato periodo di tempo; il servizio civile sarebbe facilmente realizzabile visto che un'entità intelligente può essere impegnata come "lavoratore" in molti ambiti; infine, dotando l'IA di un proprio patrimonio, sarebbe possibile imporle una sanzione pecuniaria nello stesso modo in cui essa viene inflitta alle persone o alle società, così G. HALLEVY, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, cit., pp. 196 ss.

giustificano misure legali moderne al fine di risolvere gli odierni problemi legali»[684].

17.5.1. Le obiezioni alla teoria di Hallevy.

Partendo dal presupposto comune che l'IA sia dotata di autonoma personalità giuridica[685], ad avviso di Hallevy, non sussisterebbero validi argomenti per opporsi alla perseguibilità e alla punibilità delle entità dotate di intelligenza artificiale[686]. L'Autore parte dall'assunto secondo cui, come è stato possibile riconoscere la responsabilità penale in capo agli enti, seppur con tutte le perplessità del caso, del pari dovrebbe essere possibile attribuire la responsabilità penale all'IA per i fatti dannosi o pericolosi da essa commessi[687]. Ciò sarebbe realizzabile esportando il modello di responsabilità delle persone giuridiche ed applicandolo anche ad entità legali non umane[688]. Inoltre i sistemi intelligenti, a differenza delle persone giuridiche, seppur sprovvisti di una "*soul to be damned*", potrebbero quantomeno essere dotati di un "*body to kick*" sul quale far ricadere la sanzione[689]. In tale ottica le pene inflitte agli esseri umani sarebbero equivalenti a quelle inflitte all'IA, non modificandosi neanche il significato della pena in base al suo destinatario[690].

È possibile notare che questo modo di argomentare presuppone l'attribuzione in capo all'intelligenza artificiale di un'autonoma personalità, senza la quale non sarebbe possibile disquisire di una diretta imputazione penale. Nonostante sembri che l'obiettivo dell'Autore (invero nobile) fosse quello di evitare la creazione di zone franche di responsabilità rispetto ad eventi integranti specifiche ipotesi di reato commesse dall'IA, quantomeno dal punto di vista oggettivo[691], il rischio sotteso a tali teorie è quello di un'eccessiva forzatura degli strumenti penalistici, fino a creare una sorta di panpenalizzazione, in realtà neanche troppo giustificata.

---

[684] G. HALLEVY, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, cit., p. 199.

[685] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 340.

[686] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 11.

[687] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 364; R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 267.

[688] «Il tema quindi presenta forti analogie con quello dell'affermazione di un'etica aziendale e di una conseguente responsabilità da reato degli enti collettivi, sistema con cui potremmo confrontarci nell'ipotesi di danni arrecati dall'agire autonomo di IA. Le analogie sono evidenti: gli enti collettivi non hanno né corpo né anima, ma sono comunque "soggetti giuridici", cioè autori di reati (per il tramite delle persone fisiche incardinate in essi) per la legge penale; i robot invece hanno un "corpo" fisico che interagisce con l'ambiente tramite sensori, una materia su cui far ricadere la sanzione penale (ad esempio, la disattivazione o riprogrammazione della macchina o la sua distruzione) e sono dotati di autonomia decisionale. Ciò consente di ipotizzare un sistema di responsabilità dell'agente artificiale che, sulla falsariga della responsabilità amministrativa da reato degli enti, nel capovolgerne i presupposti, renderebbe responsabile l'agente artificiale soggettivizzato per il suo operare e per quello dell'uomo *frontman*, ovvero l'utilizzatore, il programmatore, il designer, il produttore, etc.» M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 8.

[689] M.B. MAGRO, *Biorobotica*, cit., p. 514; EAD., *Robot*, cit., p. 1204; R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 267; U. RUFFOLO, Machina delinquere potest?, cit., p. 297; P. ASARO, *A Body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics*, in P. LIN, K. ABNEY, G.A. BEKEY, *Robot Ethics: The Ethical and Social Implications of Robotics*, Cambridge: MIT Press, 2011, p. 182.

[690] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 368.

[691] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 368.

Parte della dottrina ritiene che siffatto modo di ragionare si fondi su due esemplificazioni: la prima attiene alla completa assimilazione tra le caratteristiche dei sistemi intelligenti e quelle tipiche degli esseri umani; la seconda riguarda invece l'erronea convinzione che, se la responsabilità penale è stata riconosciuta in capo alle persone giuridiche, conseguentemente sia imputabile ad ogni entità diversa dall'uomo[692]. Per quanto concerne la *societas*, i reali destinatari dei precetti penali sono colori i quali, non solo l'hanno creata, ma altresì la compongono, determinandone i comportamenti. In altri termini, le persone giuridiche esistono tanto da un punto di vista sociale quanto giuridico ma sono "animate" dagli uomini[693], il che spiega l'impossibilità di assimilarle ai sistemi intelligenti. Lo schema della responsabilità degli enti non sarebbe dunque replicabile con riferimento ai sistemi intelligenti in quanto questi ultimi non potrebbero essere destinatari di un rimprovero penale, a differenza invece delle persone giuridiche, in quanto il giudizio di rimproverabilità ad esse rivolto sarebbe pur sempre indirizzato agli individui che compongono la *societas*[694].

Nonostante la suggestività delle teorie suesposte, le critiche della dottrina penalistica non si sono certo fatte attendere. Preliminarmente viene fatto notare come il concetto di *actus reus* (e, dunque, di azione) accolto dal giurista israeliano sia meramente materialistico, tipico – invero – degli ordinamenti di *common law*[695]. Tale idea di "azione", inoltre, si fonderebbe sull'errata assimilazione tra robot e intelligenza artificiale. Come evidenziato in precedenza[696], i robot sono necessariamente dotati di una componente hardware, magari anche in grado di porre in essere movimenti visibili. Non altrettanto però sembra possibile dire con riferimento ai sistemi intelligenti, i quali potrebbero essere integrati esclusivamente da un articolato software. Limitarsi ad un'azione intesa solo come movimento corporeo (*rectius* meccanico) vorrebbe dire ignorare che esistono crimini sprovvisti di una componente motoria diretta, come ad esempio i crimini informatici[697].

L'IA, pertanto, non potrebbe manifestare una reale capacità d'azione in quanto sprovvista della *suitas*. Il sistema intelligente non è in grado di integrare gli estremi della volontarietà: è agito e non agisce. A ben vedere, in realtà, anche le "decisioni" (*rectius* output) assunte dall'IA non sarebbero autonome, essendo pur sempre riferibili a colui il quale ha disegnato l'algoritmo: «quel che accade "dopo" si espone ad una *imprevedibilità necessitata*, non altrimenti governabile, che non può in alcun modo essere incasellata come manifestazione di un agire intelligentemente gestito, perché è *carente* del requisito costitutivo della *libertà di autodeterminazione*. E, quindi, da ritenere che il riconoscimento di una capacità di azione in capo alla macchina artificiale si fondi sopra una frettolosa assimilazione

---

[692] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 369.
[693] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1768.
[694] P. SEVERINO, *Intelligenza artificiale*, cit., p. 535. Gli enti, a differenza dell'IA, posseggono un "substrato umano", U. RUFFOLO, Machina delinquere potest?, cit., p. 296.
[695] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1767; A. CAPPELLINI, Machina delinquere non potest?, cit., p. 12; R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 266.
[696] Cap. II, Sez. I, Par. 7.
[697] P.M. FREITAS, F. ANDRADE, P. NOVAIS, *Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible*, in U. PAGALLO, M. PALMIRANI, P. CASANOVAS, G. SARTOR, S. VILLATA, *AI Approaches to the Complexity of Legal Systems*, Berlin-Heidelberg, 2014, pp. 152-153.

metaforica tra intelligenza umana e artificiale»[698]. Parlare di responsabilità giuridica della macchina sarebbe dunque un fuor d'opera, non essendo quest'ultima libera, bensì determinata. Essa è sprovvista della *coscienza* e dell'intenzionalità del proprio agire, ergo, della possibilità di determinarsi in modo diverso[699]. Essa difetterebbe, in sintesi, del fondamentale elemento della colpevolezza[700]. Essendo pertanto sprovvista del libero arbitrio[701], l'IA non potrebbe decidere di comportarsi diversamente e, conseguentemente, non potrebbe essere destinataria di una sanzione penale[702] (della quale, comunque, non comprenderebbe il significato): «chi è privo di "buona o cattiva coscienza", anche intesa come entità legale, non può essere considerato soggetto responsabile, perché non è possibile con lui instaurare alcun dialogo etico, non è possibile muovere alcun rimprovero»[703].

Per quanto ad un osservatore esterno potrebbe sembrare che l'agire dell'IA sia sorretto dall'intenzionalità tipica del dolo, a ben vedere si tratta di una mera apparenza[704]. Il comportamento dei sistemi intelligenti non integra una vera e propria volontà, essendo piuttosto qualificabile come una «reazione automatica del dispositivo agli stimoli fisici che riceve dall'ambiente»[705]. L'IA sarebbe pertanto pur sempre programmata per reagire in un certo modo, anche nel caso in cui agisca in modo imprevedibile. Conseguentemente non sarebbe possibile affermare che essa abbia la possibilità di agire diversamente, il che renderebbe insensato infliggere una sanzione volta a compensare il danno cagionato dalla commissione di un reato[706]. In sintesi «l'inflizione della pena presuppone la *attribuibilità psicologica* e la *rimproverabilità* del fatto di reato al soggetto che lo ha posto in essere e la sanzione penale inflittagli deve tendere alla sua

---

[698] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1769.

[699] Nell'ordinamento tedesco si parla del c.d. *principle of blameworthiness*, il quale presuppone la capacità del soggetto agente di scegliere tra il bene e il male S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., p. 420.

[700] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 4.

[701] Il libero arbitrio sarebbe la risultante del combinarsi di tre elementi strutturali: il corredo genetico, il suo sviluppo psico-fsico e l'influsso ambientale, M.B. MAGRO, *Biorobotica*, cit., pp. 500-501. Si comincia a prospettare l'idea di un "libero arbitrio artificiale" in M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 6. Sul ruolo del *free will* nel dibattito concernente l'imputazione penale di un fatto criminoso all'IA v. M. SIMMLER, N. MARKWALDER, *Guilty robots,* cit., pp. 12 ss.

[702] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1770.

[703] M.B. MAGRO, *Robot,* cit., p. 1204.

[704] «Usando un linguaggio dogmatico continentale di stampo post-finalista, si potrebbe dire che i soggetti artificiali intelligenti di tal tipo possano agire con *tipicità dolosa*. Ma questa nulla ancora ci dice riguardo alla *colpevolezza dolosa*. O meglio: è ben vero che nell'imputazione di un fatto ad agenti umani di solito la seconda – la colpevolezza – è considerata presupposta ove sussista chiaramente la prima – la tipicità –, salvo si provi la sua eccezionale insussistenza nel caso concreto (ad esempio, nel caso dell'inimputabilità o dell'*ignorantia legis* inevitabile). Ma questo meccanismo presuntivo – è evidente – si regge sul presupposto tacito per cui l'uomo, in situazioni normali, non eccezionali, sia dotato della capacità di autodeterminarsi, *scegliendo* un comportamento antigiuridico, che con tale scelta diviene anche *colpevole*» A. CAPPELLINI, Machina delinquere non potest?, cit., pp. 14-15.

[705] P. MORO, *Macchine come noi*, cit., p. 55. La "facoltà di volere" può essere intesa come una «tensione verso uno scopo, compiuto dall'uomo responsabile attraverso una scelta deliberata» p. 56.

[706] R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 267.

rieducazione (art. 27, co. 3, Cost.)», il che non sembra concretamente sostenibile per l'IA[707].

Nonostante, per certi versi, possa sembrare che con l'introduzione della responsabilità da reato degli enti collettivi si sia superata la "concezione antropomorfa della pena"[708], in realtà dobbiamo riscontrare che il nostro ordinamento resta ancora saldamente ancorato ad un concetto di *pena* antropocentrico[709]. In quest'ottica una sanzione penale inflitta ad un sistema intelligente finirebbe inevitabilmente per porsi in contrasto con le funzioni della pena che caratterizzano il nostro ordinamento[710]:

- La funzione *retributiva* non potrebbe essere integrata in quanto presupporrebbe la possibilità di muovere un "rimprovero colpevole" il che, alla luce di quanto affermato nelle battute precedenti, non sembra possibile
- La funzione di *prevenzione speciale*, intesa in un'ottica "curativa" tipica della scuola positivista, potrebbe forse realizzarsi mediante una riprogrammazione dell'IA, perseguendo una sorta di intento riabilitativo[711]. Se guardiamo, in particolar modo, ai sistemi dotati di *machine learning* possiamo riscontrare come questi ultimi riconoscano i propri errori e imparino da essi. La "penalizzazione" di questo errore servirebbe a "educare" la macchina e a far sì che lo corregga senza ripeterlo[712]. A questa prospettiva futuribile sembra però possibile obiettare, in prima istanza, che l'"ideale rieducativo" perseguito dal nostro ordinamento presupporrebbe pur sempre il consenso del condannato. Inoltre la mera disattivazione del sistema non sarebbe confacente con l'istanza di risocializzazione proposta dall'art. 27 co. 3 Cost., in quanto quest'ultima non potrebbe essere realizzata mediante la semplice neutralizzazione[713]. Da ultimo (ma non per importanza) l'IA non potrebbe comprendere il significato della sanzione che le viene inflitta in quanto essa non è in grado di percepire la riprovevolezza della propria condotta[714].
- La funzione di *prevenzione generale* verrebbe, del pari, frustrata in quanto non sembra che la sanzione inflitta all'IA possa sortire alcun effetto deterrente nei confronti di altre entità intelligenti a loro volta prive di coscienza[715]. Stiamo parlando di macchine incapaci di provare il timore – fattore che sta alla base della

---

[707] I. SALVADORI, *Agenti artificiali*, cit., p. 98. S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., pp. 422-424 affermano che «Avrebbe poco senso (sociale) attribuire la colpa a un essere incapace di riconoscere il proprio passato e di valutare le proprie azioni passate secondo un sistema di riferimento morale. Un'entità che non ha una coscienza non può partecipare a un dialogo su questioni etiche e non può rispondere al rimprovero. Un tale agente (…) può essere trattenuto fisicamente se rappresenta una minaccia per se stesso o per gli altri, ma non ha senso trattarlo come colpevole. (…) I robot, in breve, sono incapaci di comprendere il significato della punizione e quindi non possono tracciare un collegamento tra tutto ciò che è stato loro fatto e la loro precedente colpa».

[708] M.B. MAGRO, *Robot*, cit., p. 1204.

[709] U. RUFFOLO, Machina delinquere potest?, cit., p. 296. Nello stesso senso S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., p. 423.

[710] I. SALVADORI, *Agenti artificiali*, cit., p. 99; F. BASILE, *Intelligenza artificiale e diritto penale*, cit., pp. 31-32; G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 9.

[711] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1770.

[712] U. RUFFOLO, Machina delinquere potest?, cit., p. 301.

[713] R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., pp. 267-268.

[714] M.B. MAGRO, *Robot*, cit., p. 1204.

[715] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1770.

funzione di prevenzione generale[716] – e, del pari, di cogliere l'effetto pedagogico della norma penale[717].

Ogni misura sanzionatoria astrattamente diretta nei confronti dell'IA (in termini di distruzione o riprogrammazione) finirebbe per essere concretamente diretta all'uomo dietro l'IA[718], il quale si vedrebbe privato del suo sistema intelligente o sarebbe costretto a sopportarne la modifica. Ciò varrebbe, a maggior ragione, se si trattasse di una sanzione economica che verrebbe inevitabilmente pagata dal programmatore o dall'utilizzatore dell'IA[719].

17.6. Come declinare l'elemento soggettivo: a) il dolo.

Se (pur con tutte le riserve mosse in precedenza con riguardo al concetto di azione in diritto penale) potremmo in astratto affermare che l'agire dell'IA possa integrare gli elementi oggettivi del reato (in termini di fatto tipico e antigiuridico), il reale ed insormontabile ostacolo risiede nell'elemento soggettivo. L'assenza della "libertà del volere" in capo agli agenti artificiali si traduce in una «carenza ineliminabile di colpevolezza»[720].

Se l'incipiente "spersonalizzazione dell'elemento soggettivo"[721] può non essere considerata un serio problema nel diritto civile, ove il centro dell'imputazione aquiliana non è l'elemento soggettivo che sorregge il fatto illecito quanto, piuttosto, il danno ingiusto in sé considerato[722], in diritto penale la situazione non può che essere più complessa. In tal sede non è sufficiente che sia stato realizzato un fatto illecito tipico e antigiuridico, occorre anche che il fatto sia psicologicamente attribuibile all'agente e che quest'ultimo sia rimproverabile[723].

Non essendo dunque possibile ricondurre la colpevolezza in capo all'intelligenza artificiale, non resta che accertarla con riferimento all'uomo dietro l'IA. Il principale ostacolo che si incontra in tal sede afferisce alla prevedibilità, da parte del produttore o dell'utilizzatore, dei possibili eventi lesivi che possono essere cagionati dall'IA, da ciò derivando che l'attenzione del giurista dovrà appuntarsi sul tipo di rapporto che intercorre tra uomo e intelligenza artificiale[724].

Minori problemi pone l'accertamento del dolo[725]. Per quanto al progredire dell'autonomia dell'intelligenza artificiale si accompagni una sempre più

---

[716] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 16; R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 34.

[717] R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 267; M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 368.

[718] M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 9.

[719] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 98; S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., p. 424.

[720] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 4. Per un approfondimento sul tema della colpevolezza v., per tutti, R. BARTOLI, *Colpevolezza: tra personalismo e prevenzione*, Torino, 2005.

[721] U. RUFFOLO, *Il problema della "personalità elettronica"*, cit., p. 83.

[722] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 20.

[723] I. SALVADORI, *Agenti artificiali*, cit., p. 98.

[724] U. PAGALLO, *Saggio sui robot e il diritto penale*, cit., p. 607.

[725] Sul tema del dolo, in generale, A. DE MARISCO, *Coscienza e volontà nella nozione del dolo*, Napoli, 1930; M. GALLO, *Il dolo oggetto e accertamento*, Milano, 1953; A. PECORARO ALBANI, *Il dolo*, Napoli, 1955; L. EUSEBI, *Il dolo come volontà*, Brescia, 1993; C. PEDRAZZI, *Tramonto del dolo?*, in *Riv. it. dir. proc. pen.*, 2000, pp. 1265 ss.; G. MARINUCCI, *Finalismo, responsabilità obiettiva, oggetto e struttura del dolo*, in *Riv. it. dir. proc. pen.*, 2003, pp. 363 ss.; M. MASUCCI, *"Fatto" e "valore" nella definizione del dolo*, Torino, 2004; G. CERQUETTI, *La rappresentazione*

profonda scissione tra la condotta dell'IA e l'elemento soggettivo da ricondurre all'uomo che se ne serve, non sembra potersi concretamente mettere in dubbio che l'azione vada soggettivamente ricondotta alla persona umana: «l'azione, non importa quanto complessa, appartiene in definitiva all'agente umano che la realizza con coscienza e volontà»[726]. Come abbiamo già evidenziato, i sistemi intelligenti possono essere dolosamente utilizzati per delinquere, sia che vengano programmati per realizzare "di proprio pugno" un dato illecito, sia che l'agente si limiti a servirsene per commettere un qualsiasi delitto doloso: «In tali casi, la presenza di quella tensione verso il risultato che è la volontà, tipica del dolo, fa prescindere in larga misura dal concreto svolgimento del percorso causale attivato, purché l'evento voluto si sia poi effettivamente verificato»[727].

Il percorso causale assume in tal sede ancor meno importanza ove il fatto di reato si realizzi comunque, ma non secondo le modalità previste e programmate dall'utilizzatore dell'IA. Ci stiamo riferendo ad un'ipotesi di *aberratio causae*, ossia una «divergenza del decorso causale prefigurato da quello effettivo»[728]. Nel caso in cui l'IA abbia deviato in modo imprevedibile dal comportamento programmato, non viene meno l'elemento soggettivo del dolo rispetto al fatto concretamente realizzatosi, nonché voluto dal soggetto agente[729]. Quando, invece, la deviazione imprevedibile dell'agire dell'IA conduca alla mancata realizzazione del fatto di reato potrebbe configurarsi, in capo all'utilizzatore o programmatore umano, un'ipotesi di delitto tentato, sempre purché sussistano i fondamentali elementi della idoneità e univocità degli atti concretamente realizzati[730].

Dunque, l'azione apparentemente volontaria dell'agente non umano riconduce la responsabilità all'agente umano in quanto il primo si palesa come mera *longa manus* del secondo[731]. Nel momento in cui l'uomo programma il sistema intelligente con il precipuo scopo di commettere un reato «l'agire intenzionale della macchina coincide perfettamente con la volontà dell'utilizzatore o del programmatore». Il soggetto agente (a questo punto da considerarsi indubbiamente tale) non potrà incolpare la macchina, anche per un suo comportamento deviante, in quanto «se la usa intenzionalmente come strumento o se conosce e comunque accetta per sé i rischi penali che possono derivare dal suo agire»[732] si tratterà sempre di una condotta dolosa, al più declinabile nella sua forma eventuale.

---

e la volontà dell'evento nel dolo, Torino, 2004; A. MADEO, *Il dolo nella concezione "caleidoscopica" della giurisprudenza*, in *Riv. it. dir. proc. pen.*, 2010, pp. 837 ss.; G.P. DEMURO, *Il dolo*, II, *L'accertamento*, Milano, 2010; M. PIERDONATI, *Dolo e accertamento nelle fattispecie penali c.d. «pregnanti»*, Napoli, 2012; D. PULITANÒ, *I confini del dolo. Una riflessione sulla moralità del diritto penale*, in *Riv. it. dir. proc. pen.*, 2013, pp. 22 ss.; S. RAFFAELE, *Essenza e confini del dolo*, Milano, 2018.

[726] S. RIONDATO, *Robotica e diritto penale*, cit., p. 601.

[727] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 7.

[728] R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 265.

[729] I. SALVADORI, *Agenti artificiali*, cit., p. 101; A. CAPPELLINI, Machina delinquere non potest?, cit., p. 8.

[730] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 8.

[731] M.B. MAGRO, *Biorobotica*, cit., p. 514.

[732] Per entrambe le citazioni M.B. MAGRO, *Robot*, cit., p. 1207.

17.6.1. (Segue) b) la colpa.

Una volta affermato l'assunto per cui, a nostro avviso, parlare di *knowledge* ed *intention* con riferimento all'IA parrebbe (al momento) forzato, riteniamo del pari difficile poter giungere a diversa conclusione con riferimento all'elemento della *negligence*. Non crediamo che i tempi siano ancora maturi per poter parlare di "agente modello artificiale"[733], una sorta di *homo eiusdem condicionis et professionis* di nuova generazione[734]. Ciò non sarebbe possibile con riferimento alle forme di intelligenza artificiale debole, poiché il vero soggetto agente, nonché parametro di riferimento nel giudizio colposo, sarebbe pur sempre il soggetto umano. Del pari non sarebbe possibile neanche con riferimento alle ipotesi di IA forte in quanto, come ormai abbiamo imparato a comprendere, essa non può essere destinataria di un giudizio di colpevolezza in quanto priva di una "realtà interiore" assimilabile a quella umana[735].

Conseguentemente non resta, anche qui, che ragionare sull'imputazione dell'elemento soggettivo colposo[736] in capo all'uomo dietro l'IA. È chiaro che il terreno che ci accingiamo a percorrere è il più accidentato, trattandosi di ipotesi di danno non desiderato dall'utilizzatore[737], nonché quello in cui le ipotesi dei delitti commessi dall'IA sono destinate ad avere maggior incidenza[738]. La necessaria attribuibilità dell'azione al soggetto umano vale anche per i reati colposi in considerazione del fatto che, per quanto l'evento dannoso non sia voluto, l'azione colposa è sempre volontaria e consapevole, nonostante l'evento lesivo potesse essere preveduto ed evitato[739]. La colpa è stata considerata in dottrina come «criterio di attribuzione della responsabilità al tempo stesso socialmente equo ed oggettivamente idoneo sia a minimizzare il rischio che a responsabilizzare chi lo gestisce»[740].

---

[733] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1767; M. DI FLORIO, *Il diritto penale che verrà*, cit., p. 7.

[734] I parametri di riferimento sarebbero diversi rispetto a quelli dell'uomo, in quanto muterebbe «la valutazione sull'applicabilità o meno degli stessi *standard of care* da applicare al *reasonable man* in quanto, si presuppone, ci si trova in presenza di un'intelligenza superiore in grado di operare scelte come frutto della padronanza di migliaia di dati» G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 109.

[735] M. DI FLORIO, *Il diritto penale che verrà*, cit., p. 14. L'A. chiarisce che l'IA non sarebbe in grado di comprendere «il "disvalore" di una condotta illecita, che costituisce una "costruzione" sociale, ancor prima che giuridica» p. 11.

[736] La letteratura sul tema è sconfinata. Pertanto, senza alcuna pretesa di esaustività, sia consentito un rinvio – nella dottrina classica – a S. RICCIO, *Il reato colposo*, Milano, 1952; E. ALTAVILLA, *La colpa*, Torino, 1957; G. MARINUCCI, *La colpa per inosservanza di leggi*, Milano, 1965; A.R. CASTALDO, *L'imputazione oggettiva nel delitto colposo d'evento*, Napoli, 1989; U. PIOLETTI, *Contributo allo studio del delitto colposo*, Padova, 1990; F. GIUNTA, *Illiceità e colpevolezza nella responsabilità colposa*, I, *La fattispecie*, Padova, 1993; ID., *La legalità della colpa*, in *Criminalia*, 2008, pp. 149 ss., ora anche in *disCrimen*, 28.11.2018; D. CASTRONUOVO, *La colpa penale*, Milano, 2009; ID., *L'evoluzione teorica della colpa penale tra dottrina e giurisprudenza*, in *Riv. it. dir. proc. pen.*, 2011, pp. 1594 ss.; A. CANEPA, *L'imputazione soggettiva della colpa. Il reato colposo come punto cruciale nel rapporto tra illecito e colpevolezza*, Torino, 2011; A. CADOPPI, S. CANESTRARI, A. MANNA, M. PAPA, *Diritto Penale*, I, cit., pp. 446 ss.

[737] M.B. MAGRO, *Robot*, cit., p. 1207.

[738] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 8.

[739] S. RIONDATO, *Robotica e diritto penale*, cit., p. 601.

[740] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 14.

I principali problemi di attribuzione dell'elemento soggettivo colposo riguardano soprattutto le più evolute forme di IA, ossia quelle potenzialmente in grado di porre in essere comportamenti dall'evoluzione imprevedibile. È proprio questa *imprevedibilità* a porre i più consistenti problemi di attribuzione della responsabilità penale in capo agli operatori dell'IA, programmatori o utilizzatori che siano. Tuttavia occorre rilevare come l'imprevedibilità del sistema intelligente sia nota al suo stesso produttore, rendendola pertanto, in un certo senso, "preprogrammata"[741]. È stato sostenuto in dottrina che l'imprevedibilità dei risultati dell'apprendimento autonomo renderebbe prevedibile, a sua volta, la possibilità che l'IA ponga in essere attività dannose[742]. Nello stesso senso è stata paventata la possibilità di ritenere che chiunque metta in circolazione un sistema intelligente debba, per ciò stesso, prendere in considerazione ogni evento lesivo astrattamente possibile, anche in assenza di leggi scientifiche certe. Sarebbe dunque sufficiente la mera prevedibilità astratta del possibile evento lesivo a incardinare la responsabilità per colpa, anche in assenza di una sicura ricostruzione di tutti i passaggi del nesso di causalità. Secondo questo schema sarebbe «colposo il comportamento del progettista, programmatore o utilizzatore che *non preveda l'imprevedibilità del robot intelligente*»[743].

Occorre tuttavia osservare che svalutare la prevedibilità dell'evento lesivo verificatosi come conseguenza dell'azione del soggetto agente si scontrerebbe col principio di colpevolezza[744]: sarebbe infatti iniquo sostenere che per «integrare la colpa non è necessaria la prevedibilità dell'evento specifico e della sua derivazione causale dalla propria condotta (attiva o omissiva) bastando una generica e non meglio specificata *prevedibilità* della *pericolosità* del proprio agire»[745]. Diversamente, ossia pretendendo una generale ed universale prevedibilità di qualsivoglia possibile evento lesivo, si verrebbe a creare una sorta di malcelata responsabilità per posizione, scenario dal quale – come abbiamo già evidenziato – occorre rifuggire[746].

Tuttavia bisogna cercare una via mediana che eviti pericolosi vuoti di tutela. È stato a tal proposito evidenziato che l'imprevedibilità dei sistemi intelligenti non può di per sé sola esonerare i loro operatori da ogni responsabilità[747], proprio in considerazione del fatto che la loro imprevedibilità dà origine a doveri di cura. Una volta immesso il prodotto intelligente sul mercato, infatti, in capo al produttore residuano doveri di monitoraggio e di intervento che potrebbero valergli un'imputazione penale in caso di morte e lesioni – derivate dalla mancata

---

[741] M.B. MAGRO, *Robot*, cit., p. 1208.

[742] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., p. 430.

[743] M.B. MAGRO, *Robot*, cit., p. 1209; concetto ripreso dalla stessa A. in EAD., *Biorobotica*, cit., p. 516.

[744] «L'*imprevedibilità* degli eventi dovuti all'auto-apprendimento del sistema di IA (…) paralizza il giudizio di imputazione per colpa (…) nell'ambito dell'IA la colpa potrebbe operare solo al prezzo della sua definitiva consustanziazione con la responsabilità oggettiva» C. PIERGALLINI, *Intelligenza artificiale*, cit., pp. 1762-1763.

[745] L. STORTONI, *Angoscia tecnologica*, cit., p. 79.

[746] Cap. II, Sez. II, Par. 12.

[747] «L'argomento diffusamente esposto della mancanza di prevedibilità (e spiegabilità scientifica) dell'agire robotico non solleva il progettatore dalla responsabilità penale colposa poiché, così come la categoria della colpa è elaborata nel diritto vivente, la prevedibilità astratta non richiede alcuna previsione dettagliata e specifica dell'eventuale evento dannoso» M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 19.

osservanza di questi standard di diligenza – che, dunque, si sarebbero potuti evitare[748]. Certo è che appare difficoltoso, *in primis*, individuare tali standard e le relative *regole di diligenza*. Anche a seguito dell'individuazione e della relativa osservanza di regole di fonte giuridica atte ad integrare ipotesi di colpa specifica, non verrebbe comunque meno la possibilità di muovere un rimprovero per colpa generica, ove ne ricorrano gli estremi. Tale vaglio andrà fatto con capillare attenzione in quanto punire chi ha prestato fede alle regole cautelari vigenti al momento della verificazione del danno non perseguirebbe alcun fine di prevenzione generale e speciale. Inoltre, punire colui il quale non sa che la sua attività sarà idonea a provocare un evento dannoso non perseguirebbe alcun intento disincentivante[749].

Minuzioso vaglio andrà svolto non soltanto in quanto, di norma, il nesso di causalità in materia colposa tende a fondersi con la violazione delle regole cautelari[750] (le quali, come accennato, andrebbero chiaramente identificate), assottigliando così un rimprovero soggettivamente riconducibile al soggetto agente, ma anche in considerazione di un altrettanto complesso fattore: la realizzazione di un evento lesivo colposo potrebbe altresì dipendere non solo dal concorso di condizioni colpose indipendenti (art. 41 co. 3 c.p.) ma anche dalla cooperazione colposa di più persone (art. 113 c.p.) e dall'accavallarsi di più concause[751]. La colpa viene qui in rilievo nel suo aspetto relazionale, non soltanto tra gli operatori dietro l'IA ma anche tra gli uomini e l'IA medesima, parlandosi a tal riguardo del c.d. rischio di interconnessione[752].

È stato altresì sostenuto in dottrina come l'agire imprevedibile dell'IA non impedisca «l'attribuzione di una responsabilità per negligenza dell'operatore per qualsiasi danno (anche imprevedibile) causato dal comportamento dell'agente intelligente, ma che ne costituisca il *logico sviluppo*»[753]. Il richiamo al logico sviluppo non può che ricordarci la struttura dell'art. 116 c.p., il quale imputa l'evento lesivo al concorrente che volle il reato meno grave quando esso sia conseguenza della sua azione od omissione. Intanto è stato preliminarmente fatto osservare che il disposto dell'art. 116 c.p. si colloca in un contesto di *versari in re illicita*, pertanto non assimilabile al danno causato dall'IA che il soggetto agente avrebbe potuto evitare con la dovuta diligenza, muovendosi pur sempre in un contesto di rischio lecito[754]. Inoltre resterebbe da chiarire in cosa possa consistere, nel contesto che stiamo cercando di delineare, il logico sviluppo prevedibile. Tale passaggio è sostanzialmente ciò che rende costituzionalmente legittima una delle residue ipotesi di responsabilità oggettiva contenute nel nostro codice penale. Ora,

---

[748] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., p. 428.
[749] L. STORTONI, *Angoscia tecnologica*, cit., pp. 79 ss.
[750] C. PIERGALLINI, *Attività produttive e imputazione per colpa: prove tecniche di «diritto penale del rischio»*, in *Riv. it. dir. proc. pen*., 1997, p. 1480.
[751] I. SALVADORI, *Agenti artificiali*, cit., p. 103. Con riferimento alle concause l'A. riporta l'esempio delle auto a guida autonoma, affermando che «la mancata adozione, da parte dei programmatori e produttori del veicolo *smart*, di adeguate misure di sicurezza per prevenire (prevedibili) interferenze nel sistema di guida costituisce *una* delle "concause", meglio delle *condizioni* necessarie per il verificarsi dell'evento lesivo».
[752] V. MANES, *L'oracolo algoritmico,* cit., p. 550 parla di *interconnectivity risk* e di *contributory neglicence*.
[753] M.B. MAGRO, *Robot*, cit., p. 1209, corsivo nostro.
[754] M.B. MAGRO, *Biorobotica*, cit., p. 516.

mentre può essere semplice considerare, ad esempio, il reato di rapina come sviluppo del reato di furto, più complesso sarebbe comprendere il funzionamento del "logico sviluppo" rispetto all'agire (*ab initio* lecito) dell'IA. Ci si chiede quale dovrebbe essere il parametro di riferimento: se possa considerarsi sufficiente fornire l'input alla macchina per considerare ogni sua possibile condotta come logico sviluppo dell'ordine impartito dall'uomo (il quale verrebbe conseguentemente considerato sempre responsabile), o se sarebbe più opportuno andare a ricercare altrove l'innesco del "logico sviluppo". In realtà però occorrerebbe a monte chiederci se, a fronte di un comportamento imprevedibile dell'IA – pertanto non evitabile – si possa davvero parlare di uno sviluppo "logico". La mancata individuazione di questo logico sviluppo renderebbe sostanzialmente impossibile ricondurre l'evento lesivo dell'IA ad un soggetto in carne ed ossa, riprendendo così la china scivolosa del "vuoto di responsabilità".

Vuoto di tutela che si verrebbe a creare anche accogliendo un'impostazione diversa ed opposta, ossia ritenendo che l'operatore di un sistema autonomo non possa mai essere responsabile per negligenza essendo impossibilitato a prevederne il comportamento[755]. Anche tale impostazione presenta delle criticità, in quanto finirebbe per dar vita ad una sorta di speciale causa di non punibilità potenzialmente foriera di pericolosi precedenti. Tale assunto, infatti, andrebbe a scontrarsi con l'opinione di quanti ritengono che «i possibili eventi lesivi prodotti dall'intelligenza artificiale, pertanto, potranno dipendere sempre da come questa è stata progettata per essere autonoma e dal fatto che l'autonomia fa sì che il robot sia fuori dal controllo dell'essere umano»[756]. Secondo tale orientamento un controllo dell'uomo (seppur in fase preliminare) sembrerebbe residuare sempre, financo nel momento in cui l'IA raggiunga una sua autonomia. Per far ciò andrebbe preliminarmente compreso quale nozione di "controllo" dovremmo accogliere in tale contesto, proprio in quanto esso potrebbe anche essere inteso in modo "dinamico" includendo «quelle forme indirette di sorveglianza qualificate dalla capacità di un soggetto di evitare il prodursi di conseguenze maggiormente invasive o, per così dire, di "limitare i danni"»[757].

A nostro avviso, l'idea di "*controllo*" che meriterebbe di essere accolta in tal sede dovrebbe fondarsi su un residuo margine di potere impeditivo dell'evento in capo all'uomo dietro l'IA: tale dovrebbe essere il perno fondamentale per poter muovere un rimprovero per colpa all'operatore del sistema intelligente[758]. Sussisterebbe, ad esempio, la colpa del programmatore nel caso in cui quest'ultimo avrebbe potuto e dovuto prevedere l'evento lesivo, nonché adottare tutte le cautele idonee ad evitarlo[759]. Non sarà certo semplice comprendere se una certa condotta sarebbe stata prevedibile ed evitabile, specie in considerazione del fatto che l'uomo che utilizza o progetta l'IA potrebbe fare affidamento sul suo

[755] M.B. MAGRO, *Robot*, cit., p. 1208.
[756] G. CAPILLI, *I criteri di interpretazione*, cit., p. 482.
[757] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 357, gli AA. riportano l'esempio della *culpa in eligendo* per spiegare come non sia infrequente che gli ordinamenti prevedano obblighi risarcitori in capo a soggetti impossibilitati a svolgere un reale controllo sull'agire dei loro sottoposti.
[758] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 8.
[759] M.B. MAGRO, *Robot*, cit., p. 1208.

corretto funzionamento. Il principio di affidamento[760], nella penalistica classica, costituisce uno dei limiti al dovere obiettivo di diligenza, il quale persegue la funzione di riaffermare la personalità della responsabilità penale circoscrivendola all'autodeterminazione di ognuno. Vorremmo però in tal sede far notare come tale teoria si riferisca, *in primis*, al comportamento negligente di terze *persone*, concetto che (abbiamo imparato a comprendere) non è del tutto assimilabile a quello dei sistemi intelligenti. A voler però superare tale preliminare obiezione, ci scontreremmo con la figura – attualmente non immaginabile – di agente modello artificiale. Il principio di affidamento, infatti, si fonda sull'assunto per il quale ogni cittadino «può confidare che ciascuno si comporti adottando le regole precauzionali normalmente riferibili al modello di agente proprio dell'attività che di volta in volta viene in questione»[761]. In dottrina è stato proposto di dosare l'intensità della colpa tenendo conto del fattore dell'affidamento, qualificando la colpa come "lieve" a fronte del riscontro dell'esistenza un serio affidamento da parte dell'uomo e come "grave" nell'ipotesi in cui il comportamento deviante fosse stato causato dall'imperizia umana[762].

Prescindendo dalla questione concernente l'affidamento, il terreno della colpa si candida ad accogliere una moltitudine di nuove tipologie di errore: basti pensare alle ipotesi di errori di funzionamento, di montaggio ma, soprattutto, di programmazione[763]. La dottrina si è particolarmente concentrata sull'opportunità di pensare ad un nuovo concetto normativo di colpevolezza volto a fronteggiare questa nuova "colpa da programmazione" e sulla possibilità di prevedere delle relative cause di esclusione della colpevolezza consistenti nella predisposizione di misure di sicurezza volte a prevenire la consumazione di fattispecie delittuose commesse da sistemi intelligenti[764]. Procedendo verso un impoverimento dell'elemento soggettivo della colpa, in direzione dunque di un rafforzamento della colpa come criterio di imputazione di tipo normativo, connessa cioè alla mera inosservanza di regole cautelari, rischiamo di spostarci (ancora una volta) verso malcelate forme di responsabilità oggettiva. Non a caso nella sistematica dello studio sul rapporto tra colpa penale e intelligenza artificiale ci si è anche mossi verso la prospettazione della "colpa eventuale", in quanto tipologia di colpa che non richiederebbe la prevedibilità dell'evento concretamente verificatosi,

---

[760] Sul tema, in generale, v. M. MANTOVANI, *Il principio di affidamento nella teoria del reato colposo*, Milano, 1997; ID., *Alcune puntualizzazioni sul principio di affidamento*, in *Riv. it. dir. proc. pen.*, 1997, pp. 1051 ss.; F. MANTOVANI, *Il principio di affidamento nel diritto penale*, in *Riv. it. dir. proc. pen.*, 2002, pp. 536 ss. Per un approfondimento del tema in materia di responsabilità medica v., tra gli altri, P. PIRAS, G.P. LUBINU, *L'attività medica plurisoggettiva fra affidamento e controllo reciproco*, in S. CANESTRARI, F. GIUNTA, R. GUERRINI, T. PADOVANI (a cura di), *Medicina e diritto penale*, Pisa, 2009, pp. 301 ss.; A. MASSARO, *Principio di affidamento e "obbligo di vigilanza" sull'operato altrui: riflessioni in tema di attività medico-chirurgica in* équipe, in *Cass. pen.*, 2011, pp. 3857 ss.; L. RISICATO, *L'attività medica di équipe tra affidamento ed obblighi di controllo reciproco. L'obbligo di vigilare come regola cautelare*, Torino, 2013; G. FORTUNATO, *Ancora sui rapporti tra il principio di affidamento ed équipe medica*, in *dirittopenalecontemporaneo.it*, 4.5.2017, pp. 31 ss.; F. LOMBARDI, *Il principio di affidamento nel trattamento sanitario d'équipe*, in *Giurisprudenza penale*, 2.7.2018.

[761] G. FIANDACA, E. MUSCO, *Diritto Penale*, cit., p. 590.

[762] M. DI FLORIO, *Il diritto penale che verrà*, cit., pp. 13-14.

[763] I. SALVADORI, *Agenti artificiali*, cit., p. 102; A. CAPPELLINI, Machina delinquere non potest?, cit., p. 9.

[764] M.B. MAGRO, *Biorobotica*, cit., p. 516.

essendo sufficiente la mera prevedibilità astratta del rischio o, meglio, «l'impossibilità di escludere la verificazione di un qualsiasi evento di danno, anche concretamente imprevedibile»[765].

Al fine di rifuggire dalla pericolosa deriva della responsabilità oggettiva è stato proposto in dottrina di introdurre una sorta di *margine di tolleranza* con riferimento ad alcuni errori commessi dai programmatori dei sistemi intelligenti[766]. Una sistematizzazione di tale orientamento potrebbe, da un lato, delimitare l'ambito di responsabilità dei programmatori dei sistemi intelligenti e, dall'altro, evitare di disincentivare l'iniziativa economica tutelata dall'art. 41 Cost. Da regole giuridiche incerte, infatti, non può che derivare una conseguente incertezza, per gli operatori del settore, in ordine alle condotte che gli è consentito compiere. L'estremizzazione di tale incertezza potrebbe facilmente condurre ad una rinuncia *a priori* ad approcciarsi ad un settore produttivo potenzialmente pericoloso e non sufficientemente regolamentato ma anche – e soprattutto – foriero di considerevoli benefici evolutivi. Altra proposta che merita di essere presa in considerazione nel presente contesto afferisce alla previsione di un "livello di cautele differenziato" in base al tipo di intelligenza artificiale che viene immessa sul mercato e all'uso cui quest'ultima viene destinata. Proprio in base a questi fattori si determinerebbe non solo un diverso livello di rischio ma anche – nell'eventualità in cui questo rischio dovesse essere "accettato" dalla collettività – un diverso livello di misure cautelari, anche parametrato al tipo di benefici connessi all'uso dell'IA[767]. È proprio nel contesto dell'accettazione del rischio che si muove la c.d. "scelta morale" del sistema penale[768].

Alla luce degli spunti che abbiamo cercato di fornire nelle pagine precedenti, ci sentiamo di poter affermare che sarà proprio sul terreno colposo che si registreranno maggiormente le riflessioni della dottrina. Nel riflettere compiutamente su questi temi occorrerà mantenere «ben fermi i requisiti ed i criteri dell'imputazione penale senza ricorso ad inaccettabili presunzioni o ad automatismi di accertamento; ciò in particolare in tema di colpevolezza. Incidendo, infatti, la pena sulla libertà personale, nessuna deroga ai principi di garanzia è tollerabile. Che, di poi, possano presentarsi obbiettive difficoltà nella individuazione del destinatario del precetto e, quindi, della persona fisica responsabile (…) non muta, a nostro avviso, i termini della questione (…) purché ciò avvenga nel rigoroso rispetto dei canoni che l'art. 27, comma 1 della Costituzione detta in riferimento ai collegamenti – oggettivo e subbiettivo – che debbono sussistere tra fatto ed autore»[769].

17.7. Due brevi intermezzi: a) la prospettabilità di una posizione di garanzia.

Alla realizzazione di un fatto di reato da parte del sistema intelligente potrebbe ricollegarsi il mancato impedimento dell'evento lesivo medesimo da parte dell'uomo dietro l'IA, ove venisse riconosciuto in capo a quest'ultimo un

[765] M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 19. Più in generale sul tema G. CIVELLO, *La "colpa eventuale" nella società del rischio. Epistemologia dell'incertezza e "verità soggettiva" della colpa*, Torino, 2013.
[766] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., p. 431.
[767] M.B. MAGRO, *Robot,* cit., p. 1210.
[768] L. STORTONI, *Angoscia tecnologica*, cit., p. 74.
[769] L. STORTONI, *Angoscia tecnologica*, cit., pp. 87-88.

obbligo giuridico di impedire l'evento, a sua volta connesso all'esistenza di una posizione di garanzia[770]. In capo a costui ricadrebbe l'obbligo di prevenire ed evitare il concretizzarsi dei rischi insiti nell'IA. Conseguentemente, «il mancato rispetto degli obblighi di protezione (*Versicherungspflichten*) potrà essere fonte di una responsabilità (omissiva) in relazione ad eventi lesivi concreti, verificatisi in conseguenza dell'utilizzo di un a.a. [agente artificiale] o del suo "autonomo" operare, *purché rappresentino lo sviluppo del pericolo insito nella fonte da controllare e siano obiettivamente prevedibili ed evitabili*»[771].

L'introduzione di un sistema intelligente sul mercato o il suo uso nel mondo esterno potrebbe, a prima vista, integrare una "posizione di garanzia da ingerenza" nei confronti dell'uomo dietro l'IA, proprio in quanto quest'ultimo introduce una fonte di pericolo prima inesistente. L'IA in questo senso andrebbe intesa come una "fonte di pericolo", definita dalla dottrina tradizionale come «ogni realtà

---

[770] Sia consentito un rinvio alla dottrina classica in tema di reato omissivo improprio e di posizione di garanzia, O. VANNINI, *I reati commissivi mediante omissione*, Roma, 1916; M. SPASARI, *L'omissione nella teoria della fattispecie penale*, Milano, 1957; F. SGUBBI, *Responsabilità penale per omesso impedimento dell'evento,* cit.; G. FIANDACA, *Il reato commissivo mediante omissione*, cit.; ID., *Reati omissivi e responsabilità penale per omissione*, in *Il Foro italiano*, 1983, pp. 27 ss.; G. GRASSO, *Il reato omissivo improprio*, cit.; C.E. PALIERO, *La causalità dell'omissione: formule concettuali e paradigmi prasseologici*, in *Riv. it. med. leg.*, 1992, pp. 821 ss.; V. MILITELLO, *La colpevolezza nell'omissione: il dolo e la colpa del fatto omissivo*, in *Cass. pen.*, 1998, pp. 979 ss.; I. LEONCINI, *Obbligo di attivarsi, obbligo di garanzia, obbligo di sorveglianza*, Torino, 1999; M. DONINI, *La causalità omissiva e l'imputazione "per l'aumento del rischio". Significato teorico e pratico delle tendenze attuali in tema di accertamenti eziologici probabilistici e decorsi causali ipotetici*, in *Riv. it. dir. proc. pen.*, 1999, pp. 32 ss.; L. BISORI, *L'omesso impedimento del reato altrui nella dottrina e giurisprudenza italiane*, in *Riv. it. dir. proc. pen.*, 1997, pp. 1339 ss.; F. GIUNTA, *La posizione di garanzia nel contesto della fattispecie omissiva impropria*, in *Dir. Pen. Proc.*, 1999, pp. 620 ss.; A. GARGANI, *Ubi culpa, ibi omissio. La successione di garanti in attività inosservanti*, in *Ind. Pen.*, 2000, pp. 581 ss.; ID., *Le posizioni di garanzia*, in *Giur. it.*, 2016, pp. 214 ss.; ID., *Posizioni di garanzia nelle organizzazioni complesse: problemi e prospettive*, in *Riv. trim. dir. pen. econ.*, 2017, pp. 508 ss.; ID., *La responsabilità omissiva dei titolari di funzioni di protezione civile tra passato e futuro*, in *disCrimen*, 24.6.2019; F. MANTOVANI, *L'obbligo di garanzia ricostruito alla luce dei principi di legalità, di solidarietà, di libertà e di responsabilità personale*, in *Riv. it. dir. proc. pen.*, 2001, pp. 337 ss.; ID., *Causalità, obbligo di garanzia e dolo nei reati omissivi*, in *Riv. it. dir. proc. pen.*, 2004, pp. 984 ss.; L. RISICATO, *Combinazione e interferenza di forme di manifestazione del reato*, Milano, 2001; EAD., *La partecipazione mediante omissione a reato commissivo*, in *Riv. it. dir. proc. pen.*, 1995, pp. 1267 ss.; G. CRICENTI, *Il problema della colpa omissiva*, Padova, 2002; A. NAPPI, *Condotta omissiva e colpa per omissione: la causalità tra diritto e processo*, in *Cass. Pen.*, 2004, pp. 4296 ss.; G. AMARA, *Fra condotta attiva e condotta omissiva: nuovi criteri distintivi e reali conseguenze sul piano dell'imputazione dell'evento*, in *Cass. Pen.*, 2007, pp. 2795 ss.; L. RAMPONI, *Concause antecedenti e principio di affidamento: fra causalità attiva ed omissiva*, in *Cass. pen.*, 2008, pp. 566 ss.; G. MARINUCCI, *Causalità reale e causalità ipotetica nell'omissione impropria*, in *Riv. it. dir. proc. pen.*, 2009, pp. 523 ss.; G. BARBIERI, *Reato colposo: confini sostanziali tra azione ed omissione e obbligazione giuridica di prevenire l'evento*, in *Cass. pen.*, 2010, pp. 4329 ss.; S. CAMAIONI, *Trasferimento e successione di posizioni di garanzia fra riserva di legge e autonomia privata*, in *Riv. it. dir. proc. pen.*, 2010, pp. 1628 ss.; C. PAONESSA, *Obbligo di impedire l'evento e fisiognomica del potere impeditivo*, in *Criminalia*, 2012, pp. 641 ss., ora anche in *disCrimen*, 4.2.2019; A. MASSARO, *La colpa nei reati omissivi impropri*, Roma, 2011; EAD., *La responsabilità colposa per omesso impedimento di un fatto illecito altrui*, Napoli, 2013; R. CALCAGNO, *Reato omissivo improprio e responsabilità contrattuale, tra "contatto sociale" e contratto: riflessioni sul principio di legalità*, in *Cass. pen.*, 2014, pp. 3559 ss.

[771] I. SALVADORI, *Agenti artificiali,* cit., p. 106, corsivo nostro.

fattuale che si presenta *di per sé sola* (…) dotata della capacità di creazione potenziale di danni per certi beni od interessi»[772].

La fonte di questa posizione di garanzia sarebbe da rinvenire nel fare pericoloso precedente, cui però non sembra essere pacificamente riconosciuta cittadinanza nel nostro ordinamento. Dottrina più rigida ha ritenuto che il reale significato di tale posizione di garanzia sarebbe da individuare in un «distillato del principio di "*versari in re illicita*" nella forma del "*dolus subsequens*". All'attiva causazione di un evento si equipara la "volontà cattiva", cioè la decisione, successiva rispetto al comportamento illecito, di lasciare le cose come stanno, rinunziando a qualunque intervento»[773]. Dottrina (forse un po') più cauta ha ritenuto "artificioso" far discendere una posizione di garanzia dall'agire pericoloso precedente, affermando che in tal caso la "posizione di controllo" non deriverebbe dalla pregressa azione ormai giunta a compimento, bensì dalla "signoria attuale" sulla fonte di pericolo[774].

Stringendo le maglie dell'indagine l'esito non sarebbe diverso: mettere in circolazione un prodotto sarebbe un «*comportamento preliminare (prodromico) a rischio crescente (gesteigertes risikantes Vorverhalten)*, la cui riprovevolezza, come espressione di contrarietà al diritto, non deriva dalla consapevolezza della pericolosità del prodotto, ma, a ben vedere, dalla *verificazione dei primi casi di danno*: proprio la loro insorgenza legittimerebbe *ex post* la prognosi di pericolosità del prodotto *genericamente* formulabile *ex ante*»[775]. Tale scenario si porrebbe in contrasto con l'esigenza di delineare a priori ed in modo chiaro e preciso le regole cautelari che viene richiesto al produttore di osservare e, conseguentemente, l'area del rischio consentito. Accogliendo il modello della posizione di garanzia *ex Ingerenz* finiremmo per richiedere quell'astratta prevedibilità, cui sarebbe connesso un generico potere d'impedimento, di ogni possibile evento lesivo ascrivibile al prodotto, specie se intelligente.

Ci sembrerebbe più confacente ipotizzare, al più, una più generale "posizione di controllo"[776] in capo all'uomo dietro l'IA, ferma restando la difficoltà di individuare la fonte della suddetta posizione e sempre purché sussista una "effettiva possibilità di controllo"[777]. In tal senso sembra orientata quella parte di dottrina la quale ravvisa l'esistenza di una posizione di garanzia, in termini di posizione di controllo, per quanto riguarda i conducenti delle auto a guida autonoma, sui quali ricadrebbe, per l'appunto, il compito di governare una fonte di pericolo[778].

---

[772] F. SGUBBI, *Responsabilità penale per omesso impedimento dell'evento,* cit., p. 230.
[773] G. GRASSO, *Il reato omissivo improprio,* cit., p. 290.
[774] G. FIANDACA, *Il reato commissivo mediante omissione*, cit., p. 209.
[775] C. PIERGALLINI, *Danno da prodotto e responsabilità penale,* cit., p. 242.
[776] Sulla posizione di garanzia derivante da cose in custodia e sull'estendibilità di tale ultimo concetto v. M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., pp. 340-362.
[777] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 25.
[778] A. CAPPELLINI, *Profili penalistici delle self-driving cars*, in *Dir. pen. cont. - Riv. Trim.* 2/2019, p. 334; C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1753. Sulla possibilità, invece, di identificare una responsabilità per omissione per la mancata predisposizione di cautele volte ad evitare attacchi hacker v. G. CAPILLI, *I criteri di interpretazione*, cit., p. 485; sul ruolo dell'*internet service provider* L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., p. 53; A. INGRASSIA, *Il ruolo dell'ISP nel ciberspazio: cittadino, controllore o tutore dell'ordine? Risposte attuali e*

17.8. (Segue) b) il rischio consentito.

«L'irruzione dell'IA si salda pienamente con il risaputo *ruolo ancipite della scienza* nella Società del rischio»[779] nella quale viviamo. In dottrina il rischio è stato definito come il pericolo (inteso come possibilità o probabilità) di verificazione di un danno, generalmente connesso ad attività potenzialmente utili (scientifiche o produttive che siano), foriere di benefici ma anche di possibili effetti negativi, seppur collocati nel contesto di attività lecite[780]. La definizione del concetto di rischio, unita all'individuazione del margine di sicurezza necessario per governare il grado di rischio che sarà considerato tollerabile[781], chiama in causa, ancora una volta, il principio di precauzione. Come abbiamo avuto modo di accennare[782], la responsabilità per colpa dell'operatore del sistema intelligente non dovrebbe condurre ad un totale divieto di una certa attività per motivi precauzionali[783], proprio al fine di evitare di confondere la precauzione con il "prudenzialismo"[784].

Lo sviluppo dell'intelligenza artificiale è in grado di fomentare pericoli che, spesso, non è in condizione di controllare[785], per tale motivo occorrerebbe delineare con chiarezza l'area del rischio consentito[786]. Il compito di quest'ultimo è quello di raggruppare «quella sfera di attività ammesse o tollerate dall'ordinamento pur nella prevedibilità che dalle stesse possano scaturire eventi dannosi»[787]. Il problema però in questo specifico caso è duplice. *In primis*, per delimitare l'area del rischio consentito, occorrerebbe positivizzare le regole cautelari da osservare: se non si conosce la regola di diligenza da osservare non sarà conseguentemente possibile neanche sapere se, violandola, si è oltrepassato il limite del rischio consentito. *In secundis*, occorre ricordare che, per determinare l'area del rischio consentito è necessario operare un complesso bilanciamento tra l'utilità collettiva da un lato e rischi imponderabili (insiti nell'uso dei sistemi intelligenti) dall'altro[788]. La questione è tutt'altro che secondaria, posto che, al di

---

scenari futuribili di una responsabilità penale dei provider nell'ordinamento italiano, in *dirittopenalecontemporaneo.it*, 8.11.2012.

[779] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1749.

[780] L. STORTONI, *Angoscia tecnologica*, cit., pp. 72-73.

[781] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1750.

[782] Cap. II, Sez. II, Par. 13.

[783] M.B. MAGRO, *Robot*, cit., p. 1210.

[784] M.B. MAGRO, *Biorobotica*, cit., p. 516.

[785] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1749.

[786] I. SALVADORI, *Agenti artificiali*, cit., p. 108. In generale sul tema v. V. MILITELLO, *Rischio e responsabilità penale*, cit.; G. FORTI, *Colpa ed evento nel diritto penale*, Milano, 1990; F. GIUNTA, *Illiceità e colpevolezza nella responsabilità colposa*, cit.; A. CASTALDO, *La concretizzazione del «rischio giuridicamente rilevante»*, in *Riv. it. dir. proc. pen.*, 1995, pp. 1096 ss.; F. BRICOLA, *Aspetti problematici del cd. rischio consentito nei reati colposi*, in S. CANESTRARI, A. MELCHIONDA (a cura di), *Scritti di diritto penale*, Milano, 1997, vol. I, tomo I, pp. 67 ss.; P. VENEZIANI, *Regole cautelari "proprie" e "improprie" nella prospettiva delle fattispecie colpose causalmente orientate*, Padova, 2003; C. PERINI, *Il concetto di rischio nel diritto penale moderno*, Milano, 2010; G. DE VERO, *Il nesso causale e il diritto penale del rischio*, in *Riv. it. dir. e proc. pen.*, 2016, pp. 670 ss.; S. ZIRULIA, *Esposizione a sostanze tossiche e responsabilità penale*, Milano, 2018.

[787] C. PIERGALLINI, *Danno da prodotto e responsabilità penale*, cit., p. 242.

[788] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 19; R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 268.

fuori dell'area del rischio consentito, «eventuali deviazioni da parte del *software* dalle funzioni programmate daranno luogo a responsabilità del produttore a titolo di colpa con previsione o, nei casi più gravi, di dolo eventuale»[789]. Si tratterà, insomma, di definire quale sarà il livello di rischio accettato dalla società[790].

Volendo tirare le somme, il principio di precauzione e quello del rischio consentito appaiono inestricabilmente connessi, al punto da poter ritenere che il fondamento di quest'ultimo vada rinvenuto nella tutela della libertà d'agire dell'uomo, evitando una tutela assoluta dei beni giuridici che si realizzerebbe facendo un'applicazione preclusiva del principio di precauzione. Occorrerà, anche qui, compiere un bilanciamento di interessi, stante l'impossibilità di vietare del tutto ogni attività potenzialmente pericolosa, essendo queste ultime all'ordine del giorno nella società del rischio. La liceità di quest'ultimo dipenderà da una moltitudine di fattori tra cui «l'importanza del bene in gioco, il grado di probabilità dell'evento dannoso, il numero di soggetti che potrebbero essere colpiti dall'offesa, [e, da ultimo] l'utilità sociale dell'attività svolta»[791].

17.9. De iure condito *vs.* De iure condendo.

Al momento non sembra vi siano alternative percorribili: anche le più evolute forme di intelligenza artificiale sono sprovviste dell'autocoscienza necessaria al fine di comprendere il disvalore delle loro azioni e di cogliere il significato della relativa sanzione[792]. Essendo l'IA "incapace di colpevolezza" non potrebbe percepire il senso di un rimprovero[793] né si potrebbe avviare nei suoi riguardi alcun percorso di rieducazione[794]. Non sembra che questo stato di cose sia destinato a mutare nel breve termine, quantomeno «finché l'Agente intelligente non diventi anche un Agente morale»[795]. Come è possibile dedurre dal tenore delle pagine precedenti, alla possibilità di muovere un rimprovero penale all'IA si oppone non solo il principio di personalità della responsabilità penale[796] ma anche, e soprattutto, quello di colpevolezza[797]. «Per concepire una sanzione punitiva "personalmente" a carico di un agente artificiale, dovremmo immaginare che questo sia in grado di *percepire* o *comprendere* la sua condotta; dovremmo cioè ravvisare un requisito *minimo ontologico* che consenta di attribuire la *soggettività giuridica* (compresa la titolarità di diritti) e quindi anche la *responsabilità penale* per il suo comportamento»[798].

Non possiamo che confermare, per il momento, che la macchina non può delinquere, non può essere considerata colpevole e, conseguentemente, non può

---

[789] P. SEVERINO, *Intelligenza artificiale*, cit., p. 536.
[790] G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 9.
[791] C. PIERGALLINI, *Danno da prodotto e responsabilità penale,* cit., p. 243.
[792] I. SALVADORI, *Agenti artificiali,* cit., p. 96.
[793] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1765.
[794] M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 9.
[795] M.B. MAGRO, *Robot,* cit., p. 1205. L'A. si chiede, in altro scritto, se sia «possibile che, in un ipotetico futuro, gli agenti intelligenti riescano a sviluppare coscienza, sentimenti, empatia e moralità, così da poter infliggere loro una punizione» EAD., *Decisione umana e decisione robotica,* cit., p. 10.
[796] I. SALVADORI, *Agenti artificiali,* cit., p. 98.
[797] A meno che lo scopo della pena non sia esclusivamente quello di recare un "beneficio psicologico" alla vittima della condotta lesiva causata dall'IA, A. CAPPELLINI, Machina delinquere non potest?, cit., p. 21; R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 268.
[798] M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 10, corsivo nostro.

essere punita[799]. Ad oggi, dunque, la responsabilità deve sempre essere imputata a un soggetto umano[800], essendo nel dominio di quest'ultimo l'input che viene dato alla macchina[801].

Non ci resta che ragionare in prospettiva futuribile e chiederci se i sistemi intelligenti potranno essere chiamati a rispondere per le condotte penalmente rilevanti da essi stessi realizzate[802] e come, sempre *de iure condendo*, reagiranno a ciò i corollari del diritto penale. Un domani l'IA potrebbe essere destinataria di un armamentario sanzionatorio *ad hoc* ma, a ben vedere, ciò dipenderà dal riconoscimento sociale che si deciderà di riservarle[803]. Potremmo arrivare al punto in cui un giorno i sistemi intelligenti saranno così capillarmente diffusi nelle nostre vite da considerare i danni da essi cagionati alla stregua di un evento naturale non controllabile dall'uomo. Si tratterebbe di un rischio "normale" del quale nessun operatore umano potrebbe essere chiamato a rispondere penalmente – specie ove dovesse trattarsi di un malfunzionamento solo astrattamente prevedibile – venendo egli considerato responsabile soltanto per le conseguenze dannose derivate da errori commessi in fase di programmazione e, pertanto, evitabili[804].

Ci si chiede se sia possibile costruire, in prospettiva *de iure condendo*, una penalità delle cose o, per meglio dire, un diverso diritto penale per le "soggettività minori", che abbiamo imparato a conoscere come "attanti"[805]. Per il momento, non ci resta che prendere atto che anche le più evolute forme di IA, ossia quelle in grado di agire in modo indipendente, sono programmate per farlo, in quanto si muovono nel solco tracciato dall'algoritmo per esse prestabilito dal loro creatore umano. Da ciò deriverebbe che, in astratto, in capo all'uomo residuerebbe sempre il potere di evitare l'evento dannoso cagionato dall'IA[806]. In quest'ottica, anche una responsabilità penale direttamente indirizzata a un sistema intelligente sarebbe pur sempre «al servizio degli obiettivi umani (come il controllo sociale) e fondata sul senso di giustizia umano»[807]. Meditando sulla possibilità di considerare in futuro l'IA come un'entità direttamente responsabile da un punto di vista penale, in dottrina è stata paventata la proposta di adottare una concezione normativa della responsabilità secondo cui «la libertà di azione non è un fenomeno naturale di cui occorre individuare i referenti ontologici, ma un mero

---

[799] A. CAPPELLINI, Machina delinquere non potest?, cit., pp. 15-19; I. SALVADORI, *Agenti artificiali*, cit., p. 97; C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1771. *Contra* U. RUFFOLO, Machina delinquere potest?, cit., p. 302 ad avviso del quale *machina delinquere potest*, al pari della *societas*. La penalizzazione dell'IA si fonderebbe sulla stigmatizzazione degli errori da questa commessi: la sua capacità di assumere decisioni testimonierebbe la sua consapevolezza e conoscenza, nonché la possibilità di considerare il suo agire sorretto da *intent* o *negligence*. La mancanza dell'autocoscienza, infine, non ne impedirebbe la responsabilizzazione.

[800] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 100 il quale, a sua volta, richiama la Risoluzione del Parlamento europeo recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica.

[801] C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 1.

[802] I. SALVADORI, *Agenti artificiali*, cit., p. 98.

[803] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1766.

[804] M.B. MAGRO, *Robot*, cit., p. 1211.

[805] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 20. Abbiamo già parlato degli attanti nel Cap. II, Sez. III, Par. 17.4.

[806] S. RIONDATO, *Robot: talune implicazioni di diritto penale*, cit., p. 92.

[807] S. RIONDATO, *Robotica e diritto penale*, cit., p. 603.

attributo normativo dell'agire del responsabile autore del reato, funzionale agli scopi dell'organizzazione sociale. Questo concetto di libertà e di colpevolezza potrebbe fornire una buona base di partenza per progettare sistemi normativi di attribuzione della responsabilità giuridica delle IA»[808].

Ci sentiamo di condividere in tal sede l'opinione di quella parte di dottrina ad avviso della quale «la questione di cosa sia la responsabilità penale e di chi possa essere penalmente responsabile dipende dalla società»[809]. Sarà proprio su questo terreno che si giocherà la partita del riconoscimento di una responsabilità penale in capo ai sistemi intelligenti, quando cioè arriveranno al punto di essere talmente tanto umanizzati da poter deludere le aspettative che la società ripone in loro.

In dottrina è stato ritenuto che, in considerazione della "relatività sociale" del contenuto e del concetto stesso di responsabilità penale, spetterà alla società del futuro (ed al modo in cui essa opererà) determinare se i sistemi intelligenti saranno riconosciuti come persone e se le loro azioni avranno il potenziale per destabilizzare il vigente assetto normativo. In tal caso, sarà sempre compito della società sviluppare meccanismi volti a prevenire la succitata destabilizzazione normativa al fine di garantire la stabilità. Tale dottrina evidenzia come sarà altamente possibile che lo strumento utilizzato dalla società per svolgere questo compito sarà il diritto penale e che, conseguentemente, sarà possibile che l'idea di un'IA colpevole diventi parte della nostra vita quotidiana. La questione decisiva riguarderà, dunque, non tanto ciò di cui l'IA è capace, quanto piuttosto quale sarà il ruolo che noi, come società, decideremo di attribuirle[810].

18. Conclusioni (dalle quali siamo ancora lontani).

Sembra dunque che, in fin dei conti, quello che è stato da più parti definito come un "pregiudizio antropocentrico"[811] in realtà si riveli più un giudizio ponderato: la responsabilità (almeno allo stato) non può che essere dell'uomo[812]. La responsabilità penale, in particolare, rimane ad esclusivo appannaggio della persona in carne ed ossa, ferma restando l'esigenza di continuare a studiare il nuovo volto che possono assumere l'elemento oggettivo e l'elemento soggettivo del reato materialmente commesso dall'IA. L'imprevedibilità di quest'ultima è destinata ad incidere sul concetto di "colpevolezza"[813] e il vaglio di nuove forme di responsabilità penale è destinato, a sua volta, a mettere sotto sforzo gli elementi costitutivi del fatto di reato[814]. Volendo fare un primo passo in tal senso potremmo affermare che, per garantire la personalità della responsabilità penale, il controllo

---

[808] M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 7.
[809] M. SIMMLER, N. MARKWALDER, *Guilty robots,* cit., p. 23.
[810] M. SIMMLER, N. MARKWALDER, *Guilty robots,* cit., pp. 25-26.
[811] I. SALVADORI, *Agenti artificiali,* cit., p. 96; A. CAPPELLINI, Machina delinquere non potest?, cit., p. 13; U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 20.
[812] C. CASONATO, *Potenzialità e sfide dell'intelligenza artificiale*, cit., p. 181.
[813] Sul difficile rapporto tra IA e colpevolezza M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 10; F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 30.
[814] U. PAGALLO, *Saggio sui robot e il diritto penale*, cit., pp. 603-604. L'A. individua i tre punti nodali della questione concernente la responsabilità penale dell'IA: «a) l'imputabilità e/o responsabilità personale dell'agente artificiale; b) i temi della colpevolezza umana relativi all'imprevedibilità e autonomia dei robot; c) la possibilità di concepire nuovi reati in rapporto a nuovi (presunti) "diritti dei robot"» p. 606.

dell'uomo sul sistema intelligente non dovrebbe mai venire meno[815]. Abbiamo però imparato a comprendere che, specialmente con riguardo ai più sofisticati sistemi di IA, tale controllo non è sempre completamente esercitabile.

All'esito della disamina appena svolta possiamo dare per acquisito un altro dato, ossia la difficoltà di parlare di una piena assimilazione giuridica, ai fini del diritto penale, tra l'uomo e l'intelligenza artificiale[816]. Si sono in tal senso contrapposti due orientamenti: secondo l'opinione dogmatica l'IA sarebbe in grado di imitare, con una riproduzione fedele, il comportamento umano; secondo l'opinione scettica, invece, l'IA sarebbe comunque priva di una mente assimilabile a quella dell'uomo, non potendo pertanto considerarla intelligente. «Entrambe le opinioni sono ingenue, perché riducono l'uomo e la macchina ad un'unica dimensione. Nella prima concezione si presume che la mente sia soltanto un algoritmo, che si può implementare in un sistema esperto di intelligenza artificiale e che è elaborato dalla ragione che calcola e dalla volontà che reagisce agli stimoli esterni; nella seconda concezione si suppone che la macchina cibernetica non possa svolgere ragionamenti, né imitare comportamenti né provare emozioni tipiche soltanto della natura umana»[817].

È stato da alcuni sostenuto che l'intelligenza artificiale possa effettivamente sostituirsi all'uomo nello svolgimento di attività lecite e illecite[818], financo al punto di superarlo[819]. Si tratta però, a ben vedere, di una prospettiva puramente tecnologica, in quanto è possibile riscontrare come i sistemi intelligenti siano in grado di svolgere determinate funzioni con un'efficienza inarrivabile per l'uomo ma, al contempo, non riescano a realizzare condotte che per l'uomo sono quasi banali. L'IA è in grado di risolvere problemi computazionali di particolare complessità con rapidità impressionante[820], ma avrebbe difficoltà a svolgere

---

[815] G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 111, l'A. approfondisce il suo pensiero affermando che «tale, ultima postilla appare l'evidenza *de facto* della volontà di non lanciare nel mondo un nuovo tipo di soggetto giuridico del tutto autonomo e indipendente dal proprio realizzatore».

[816] S. RIONDATO, *Robotica e diritto penale*, cit., p. 603.

[817] P. MORO, *Biorobotica e diritti fondamentali*, cit., p. 538.

[818] I. SALVADORI, *Agenti artificiali*, cit., pp. 84-87. Sulla sostituzione dell'uomo in estesi aspetti del suo agire ad opera dell'"informatizzazione" v. L. PICOTTI, *Diritto penale e tecnologie informatiche*, cit., p. 44.

[819] F. BASILE, *Intelligenza artificiale e diritto penale*, cit., pp. 2-3 il quale richiama, a sua volta, le parole di Stephen Hawking durante la Conferenza Zeitgeist a Londra nel maggio 2015: «nell'arco dei prossimi cento anni, l'intelligenza dei computer supererà quella degli esseri umani».

[820] Uno dei settori in cui l'IA ha mostrato quanto possa essere performante è quello dei giochi. «Nel 2011 *IBM Watson* è riuscita a vincere contro i campioni di *Jeopardy*, un complicato quiz televisivo che consiste in una gara di cultura generale tra i concorrenti; nel 2016 e nel 2017 *Google AlphaGo* è riuscita a battere Lee Sedol e Ke Jie, campioni mondiali del complicatissimo gioco Go; nel 2017 *Liberatus*, un sistema sviluppato a Carnegie-Mellon, è riuscito a battere i migliori giocatori di *Texas Hold'em*, una versione molto complessa del gioco del poker» G. ITALIANO, *Intelligenza Artificiale*, cit., pp. 216-217. Per riprendere alcuni concetti analizzati nel Cap. II, Sez. I, Par. 4, potrebbe interessare sapere che, alla base dell'apprendimento di AlphaGo, ci sono i già menzionati sistemi di *reinforced learning* (apprendimento con rinforzo), A. VESPIGNANI, *L'algoritmo e l'oracolo*, cit., pp. 68-69. La versione più evoluta di AlphaGo è AlphaGo Zero, della quale sono state create due versioni: una è stata addestrata anche analizzando partite giocate dall'uomo ed ha appreso più in fretta, l'altra ha imparato da sola, impiegando maggior tempo ma raggiungendo una qualità di gioco migliore. Questo dato costituisce un esempio lampante di come funzionino i *bias* cognitivi umani e di come essi possano essere trasferiti all'IA, R. ROVATTI, *Il processo di apprendimento algoritmico*, cit., p. 34. Il gioco per eccellenza in cui si è registrata la

compiti che richiedono anche uno scarso livello di abilità, come ad es. prendere al volo un oggetto o riconoscere un'immagine[821]. Parlare, a livello generale, di "sostituzione" non dovrebbe sorprendere: ogni rivoluzione tecnologica porta con sé la sostituzione dell'uomo in ambiti che in precedenza erano ad esso esclusivamente riservati[822]. Basti pensare ai robot industriali utilizzati per svolgere mansioni ripetitive e faticose[823]. Anche qui, però, sorge il problema dell'attribuzione della responsabilità nel caso in cui il macchinario cagioni un danno all'uomo che si trova ad interagire con esso[824].

La partita si giocherà, a nostro avviso, sul terreno del rischio consentito (*rectius* sullo spazio che si deciderà di riservargli). Nel momento in cui i sistemi intelligenti raggiungeranno un livello di diffusione tale da essere universalmente considerati forieri di non secondari vantaggi per la comunità, allora sarà più semplice strutturare adeguate norme cautelari e ragionare dell'imputazione della relativa responsabilità per i danni da essa cagionati (sulla falsariga di quanto già avvenuto con la diffusione delle automobili, in quanto prodotti socialmente utili ma pericolosi). «La misura del rischio consentito potrà essere determinata sulla base di regole cautelari predeterminate dal legislatore, che rileveranno sul piano della colpa specifica». Da ciò deriverebbe che, nella misura in cui la società abbia accettato i rischi connessi all'uso dell'IA pur di trarne i relativi benefici, le eventuali «conseguenze negative che dovessero sorgere in conseguenza di (eventuali) errori in fase di programmazione, di sviluppo, di sperimentazione o di produzione degli a.a. [agenti autonomi] ovvero a causa del loro malfunzionamento, laddove siano state rigorosamente osservate tutte le regole cautelari, dovrebbero invece essere accollate alla collettività»[825]. In altri termini, il

---

persistenza dell'IA sono stati gli scacchi. Nel 1997 DeepBlue, un programma scacchistico ideato sempre dalla IBM, battè il campione del mondo di scacchi Garry Kasparov. Dall'esperienza di AlphaGo è nato AlphaZero, il quale ha imparato gli scacchi giocando contro sé stesso con sole 9 ore di addestramento, L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., pp. 5-6.

[821] C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., pp. 119-120; L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., p. 11. Per descrivere tale problema parte della dottrina parla del c.d. paradosso di Moravec, G. ITALIANO, *Intelligenza Artificiale*, cit., p. 213. La dottrina specialistica, riferendosi al fenomeno delle auto a guida autonome, chiarisce che attribuire un significato ai dati raccolti dai sensori è probabilmente la parte più difficile della progettazione di un veicolo completamente autonomo. Proprio nel campo del riconoscimento delle immagini questo potrebbe costituire un rilevante problema, specie in materia di circolazione stradale: basti pensare all'importanza di riconoscere i segnali stradali, i colori dei semafori e la segnaletica orizzontale, H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law,* cit., p. 353.

[822] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 39.

[823] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 2; A. TURANO, *Robotica e roboetica*, cit., p. 130 nota 21.

[824] C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 104. Un esempio di ciò è fornito proprio dallo stesso Hallevy in quale riporta che «nel 1981, un impiegato giapponese di 37 anni di una fabbrica di motociclette è stato ucciso da un robot di intelligenza artificiale che lavorava vicino a lui. Il robot ha erroneamente identificato il dipendente come una minaccia per la sua missione e ha calcolato che il modo più efficiente per eliminare questa minaccia era spingerlo contro una macchina operatrice adiacente. Usando il suo braccio idraulico molto potente, il robot ha spinto il lavoratore sorpreso contro un macchinario operante, uccidendolo all'istante, e poi ha ripreso i suoi compiti senza che nessuno interferisse con la sua missione» G. HALLEVY, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, cit., pp. 171-172.

[825] Per entrambe le citazioni I. SALVADORI, *Agenti artificiali,* cit., p. 117.

costo sociale dei danni provocati dall'IA dovrà essere allocato in capo a colui il quale beneficia dell'uso di un bene intrinsecamente pericoloso.

La via da imboccare sarà quella della minimizzazione del rischio, accollandolo al soggetto che trae un vantaggio dall'uso dell'IA e che, pertanto, è anche la persona maggiormente idonea a gestire il relativo rischio, secondo il principio del *cuius commoda eius et incommoda*[826]. In tal senso si potrebbe pensare di ragionare nel seguente modo: nella misura in cui la "deviazione" dell'agire pre-programmato dell'IA sia prevedibile ed evitabile, allora sarà il produttore a dover essere considerato responsabile, altrimenti spetterà al soggetto danneggiato l'onere di sopportare il "costo" dell'evento lesivo cagionato dall'IA[827], chiaramente nella misura in cui la persona offesa abbia tratto a sua volta un vantaggio, anche solo potenziale, dall'interazione col sistema intelligente. La stella polare delle riflessioni sul tema non potrà che essere quella della delimitazione dell'area del rischio consentito: «l'accettabilità del rischio dipende dalla possibilità – per chi sostiene le perdite – di ricevere anche dei vantaggi. Laddove ciò non accada, il rischio risulterà inaccettabile da chi ne è interessato»[828].

Non è ancora tempo di trarre frettolose conclusioni sulla soggettività artificiale e sulla loro responsabilità. È, piuttosto, il momento di continuare a porsi domande, di formulare ipotesi[829] e di comprendere come il diritto penale possa contribuire allo sviluppo di questi delicati quanto intricati temi.

19. Intelligenza artificiale come "vittima" di reato.

Per completezza vorremo dedicare un breve riferimento anche alla possibile qualificazione dell'IA come "vittima" di reato, non solo in quanto la dottrina si è occupata (anche) di questo tema, ma soprattutto perché ci sembra il naturale completamento della trattazione svolta nelle pagine precedenti. Immediata testimonianza di ciò è determinata dal riemergere della già affrontata questione concernente il riconoscimento di una personalità elettronica in capo all'IA[830].

Ad avviso di parte della dottrina, infatti, rendere l'intelligenza artificiale giuridicamente responsabile servirebbe anche a "garantirla" dal c.d. *ius utendi et abutendi* di colui il quale vanti un diritto su di essa[831]. È infatti ben possibile che l'IA finisca per essere "vittima" di un uso improprio da parte del suo utilizzatore[832]. Il problema del riconoscimento della personalità giuridica in capo ai sistemi intelligenti, infatti, non passa soltanto dalla loro responsabilizzazione e dall'attribuzione in capo a questi ultimi dei relativi diritti, ma anche dalle tutele da riconoscere in capo ad essi. In tal senso è stato osservato che considerare i sistemi intelligenti come meritevoli di tutela passerebbe pur sempre per il tramite della

[826] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 11.
[827] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 359.
[828] J. YATES, *Paura e società del rischio. Un'intervista a Ulrich Beck*, in *Lo Sguardo – Rivista di filosofia*, 2016, p. 213.
[829] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 29.
[830] Cap. II, Sez. III, Par. 17.3.
[831] U. RUFFOLO, Machina delinquere potest?, cit., p. 304.
[832] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 370; prende in considerazione i reati commessi "nei riguardi" dei sistemi intelligenti anche U. PAGALLO, *Saggio sui robot e il diritto penale*, cit., p. 607.

coscienza e della sensibilità sociale[833]. Per vero in dottrina è stato fatto notare che la personalità – o, più in generale, le forme di protezione – sono state riconosciute, piuttosto che sulla base dell'intelligenza, in considerazione del carattere "senziente" dell'entità di riferimento[834]. L'esempio più confacente che potremmo fornire in tal sede è quello degli animali i quali, nonostante il loro carattere senziente, continuano però ad essere considerati come "oggetti" nel nostro ordinamento: meritevoli di protezione ma non per questo da considerare come soggetti di diritto[835].

Su questa scia si potrebbe ritenere che il riconoscimento di una personalità giuridica in capo all'IA non sarebbe necessario in quanto, più semplicemente, i sistemi intelligenti potrebbero essere considerati oggetto di tutela penale in quanto beni patrimoniali. Potrebbero infatti assurgere al rango di beni giuridici meritevoli di tutela penale sulla scorta del fatto che il diritto penale non sarebbe nuovo al riconoscimento di tutele in capo ad entità non-umane, basti pensare a tal proposito, al disposto del Titolo IX *bis* rubricato "Delitti contro gli animali". Tuttavia la questione realmente problematica non concerne tanto la considerazione dell'IA come oggetto di tutela quanto, piuttosto, la sua qualificazione come "vittima di reato"[836].

Far assurgere i sistemi intelligenti ad autonomi centri di tutela comporterebbe, in prospettiva *de iure condendo*, l'incriminazione dell'uomo per i danni cagionati all'IA[837]. Per alcuni tale prospettiva potrebbe non incontrare troppi ostacoli, in considerazione del fatto che «la qualità di vittima del reato non presuppone il possesso degli stessi requisiti psicologici della responsabilità penale»[838]. Ad avviso di altri, invece, l'IA non potrebbe essere equiparata all'uomo quale destinataria degli effetti della normativa penale, non potendo pertanto essere considerata come "soggetto passivo" di un illecito tradizionalmente indirizzato ad una persona: *a contrario*, si tratterebbe di un caso di analogia in *malam partem*[839]. Secondo altri ancora, invece, sarebbe il caso di evitare pregiudizi antropocentrici atti ad ostacolare il riconoscimento della qualità di vittima potenziale del reato in capo ai sistemi intelligenti[840].

In realtà sono stati proposti diversi esempi in dottrina volti a palesare, tanto *de iure condito* quanto *de iure condendo*, come i sistemi intelligenti possano costituire oggetti (o soggetti) di tutela penale. V'è chi si è concentrato sulla portata "oggettiva" del sistema, assimilando i *software* a programmi informatici

---

[833] U. RUFFOLO, *Il problema della "personalità elettronica"*, cit., p. 76.
[834] U. RUFFOLO, *Il problema della "personalità elettronica"*, cit., p. 78; ID., Machina delinquere potest?, cit., p. 298.
[835] F. CAROCCIA, *Soggettività giuridica dei robot?*, cit., p. 236.
[836] S. RIONDATO, *Robotica e diritto penale*, cit., p. 600, l'A. chiarisce che «Tale prospettiva comporta l'attribuzione di diritti e interessi ai robot e rientra nel più vasto problema (…) relativo al fatto che i robot possano essere considerati soggetti del reato»; ID., *Robot: talune implicazioni di diritto penale*, cit., p. 86.
[837] U. PAGALLO, *Saggio sui robot e il diritto penale*, cit., p. 606.
[838] S. RIONDATO, *Robot: talune implicazioni di diritto penale*, cit., p. 91; R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 263.
[839] I. SALVADORI, *Agenti artificiali*, cit., p. 111.
[840] G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 10, il quale richiama a sua volta U. RUFFOLO, *La "personalità elettronica"*, in U. RUFFOLO, *Intelligenza artificiale. Il diritto, i diritti, l'etica*, cit., p. 232 e U. RUFFOLO, *Intelligenza Artificiale, machine learning e responsabilità da algoritmo*, in *Giur. it.*, 2019, p. 1703.

suscettibili di alterazione o danneggiamento ex art. 635 *bis* c.p. o ex art. 635 *ter* c.p., ove si tratti di *software* utilizzati da un ente di pubblica utilità. Del pari un attacco hacker indirizzato verso un sistema intelligente potrebbe integrare il reato di accesso abusivo a sistema informatico ex. art. 615 *ter* c.p. Ancora, i *software agents* potrebbero essere considerati meritevoli di protezione penale in quanto prodotto di una creazione intellettuale, la cui riproduzione sarebbe pertanto punita ai sensi dell'art. 171 *bis* della legge sulla protezione del diritto d'autore[841].

Altri hanno preso in esame una prospettiva più "soggettiva" dell'IA, considerando non tanto il sistema intelligente in sé quanto, piuttosto, la percezione sociale del disvalore dell'azione lesiva di cui l'agente artificiale può essere destinatario. Proviamo a spiegarci meglio. In dottrina sono stati presi in considerazione i programmi di *doll therapy* e *pet therapy* che possono essere utilizzati come strumento terapico per soggetti vulnerabili. Ci si è chiesti in tal senso se una loro distruzione integri il reato di danneggiamento o se, in considerazione del legame affettivo che si può venire ad instaurare con l'IA, si possa arrivare a parlare dei reati di maltrattamento contro familiari e conviventi o contro gli animali[842]. Ci si è poi interrogati sul ruolo dei c.d. *sexbots*, intesi come robot con cui sarebbe possibile instaurare relazioni affettive, chiedendosi come comportarsi nel caso in cui uno di questi *sexbots* avesse le fattezze di un minore. Ad avviso della dottrina la disciplina degli atti sessuali con minori in carne ed ossa (che sarebbero indubbiamente da considerarsi illeciti), non sarebbe così facilmente estendibile al medesimo atto commesso ai danni di un'entità sprovvista di una integrità psico-fisica da tutelare[843]. Tale questione ci conduce direttamente al tema più "avveniristico" che si comincia ad affrontare in dottrina, ossia quello del c.d. "stupro robotico"[844]. La criminalizzazione di tali condotte è stata percepita come una visione moralizzante del diritto penale, volta a perseguire non tanto la tutela di un ben individuato bene giuridico quanto, piuttosto, un rimprovero al "tipo di autore" giudicato in base ad una condotta considerata socialmente riprovevole[845].

Rifuggendo, dunque, da un moralismo giuridico che non trova cittadinanza nel nostro ordinamento, dobbiamo chiederci se incriminazioni di tal fatta mirerebbero a tutelare direttamente l'IA o, più probabilmente, un impianto valoriale pur sempre umano, secondo un interrogativo simile a quello che ha accompagnato l'introduzione dei reati di maltrattamento contro gli animali: ad essere vittima del reato è l'animale (*rectius* il sistema intelligente) in sé o il sentimento dell'uomo dei suoi confronti?

---

[841] Per questi e altri esempi v. I. SALVADORI, *Agenti artificiali*, cit., pp. 111 ss.

[842] F. BASILE, *Intelligenza artificiale e diritto penale*, cit., pp. 32-33.

[843] I. SALVADORI, *Agenti artificiali*, cit., pp. 114, l'A. giunge ad una conclusione opposta (ossia favorevole alla criminalizzazione dell'atto) nel caso in cui il *sexbot* sia utilizzato per produrre immagini di pornografia minorile, applicando il disposto dell'art. 600 *quater*.1 c.p. Sul tema, più in generale, v. anche C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., pp. 117-118.

[844] J. DANAHER, *Robotic Rape and Robotic Child Sexual Abuse: Should They be Criminalized?*, in *Criminal Law and Philosophy*, 2017, pp. 71 ss.; M.H. MARAS, L. R. SHAPIRO, *Child Sex Dolls and Robots: More Than Just an Uncanny Valley*, in *Journal of Internet Law*, 2017, pp. 3 ss.

[845] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 3 nota 2; F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 33.

CAPITOLO III
DALLA TEORIA ALLA PRATICA. ALCUNE CONCRETE IMPLICAZIONI PENALISTICHE DELL'INTELLIGENZA ARTIFICIALE

SOMMARIO: 1. Non solo teoria – 2. Le auto a guida autonoma: un preliminare inquadramento normativo – 2.1 Una questione di riconoscimento sociale – 2.2 Profili penalistici dei veicoli semi-autonomi – 2.3 Le *driverless car* e i limiti del diritto penale – 2.4 L'etica delle auto autonome: il dilemma del carrello ferroviario – 3. L'intelligenza artificiale nel settore medico – 3.1 Il supporto intelligente in fase di diagnosi – 3.2 Chirurgia robotica e gradi di responsabilità – 3.3 Bionica, biorobotica e *human enhancement*: fra limiti e confini mobili – 3.4 La robotica assistenziale – 3.5 Profili penalistici comuni – 3.6 Un approccio proattivo: dal consenso del paziente all'impossibile sostituzione del medico.

1. Non solo teoria.

Vorremmo dedicare la presente parte della nostra analisi ad una breve rassegna dei concreti ambiti applicativi in cui l'intelligenza artificiale (e le sue implicazioni giuridiche) possono trovare fertile terreno di sviluppo. I settori in cui l'IA può trovare applicazione sono i più disparati e le relative conseguenze giuridiche possono rivelarsi quanto più complesse[1]: basti pensare alla sfera economica o assistenziale, all'ambito militare e medico, financo a quello dell'apprendimento[2].

La diffusione su larga scala dei sistemi intelligenti può, però, portare con sé non soltanto conseguenze dannose[3] ma anche usi criminali (come invero già accennato)[4]. L'IA, infatti, può essere utilizzata per la produzione di contenuti illeciti[5], per accelerare e automatizzare le procedure illecite di bagarinaggio online e per realizzare abusive manipolazioni di mercato[6]. Con particolare riferimento a tale ultimo settore è il caso di constatare che, se da un lato gli algoritmi possono costituire strumento utile in materia di finanza, accelerando le operazioni umane e rendendo il mercato più "decifrabile"[7], di contro essi possono eseguire transazioni finanziarie, essendo a loro rimessa non soltanto l'esecuzione bensì la stessa decisione di realizzare suddette transazioni. Ci stiamo, infatti, riferendo al c.d. fenomeno dell'*High Frequency Trading* (HFT), il cui impiego può rendere

[1] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 419.
[2] M.B. MAGRO, *Robot*, cit., pp. 1192 ss.
[3] A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 100.
[4] Cap. II, Sez. III, Par. 16.
[5] I. SALVADORI, *Agenti artificiali*, cit., p. 109.
[6] F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 26.
[7] C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 104. Sull'impatto dei sistemi di *machine learning* nel mondo della finanza G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 5 il quale afferma che «abbiamo già moltissimi strumenti basati interamente su machine learning e che non richiedono alcun intervento umano, come ad esempio i robo-advisor, che sono dei consulenti finanziari digitali, abbiamo piattaforme digitali di valutazione dei rischi (*risk assessment*), abbiamo strumenti per la gestione del portafoglio di investimenti, per stimare l'affidabilità creditizia (*credit score*) delle società e delle persone, per la rilevazione di frodi, abbiamo anche piattaforme di *algorithmic trading* che analizzano velocemente enormi quantità di dati ed effettuano in maniera automatica acquisti e vendite di titoli e azioni».

possibili evolute esecuzioni dei classici modelli di turbativa del mercato, nonché metodologie inedite di manipolazione dei titoli[8].

Tuttavia l'impiego dell'intelligenza artificiale può essere, altresì, foriero di considerevoli benefici. Essa può svolgere un ruolo di "assistenza" nella maggior parte delle nostre attività quotidiane[9]. Pensiamo agli *Ambient Assisted Living* (AAL), ossia sistemi di assistenza domestica, e all'uso che può esserne fatto per prendersi cura delle persone anziane[10]; ai sistemi di apprendimento automatico che possono essere utilizzati per ridurre il consumo energetico[11]; all'uso di algoritmi predittivi, da parte delle aziende, per sapere in anticipo quando i loro prodotti necessiteranno di manutenzione[12]; all'uso dei sistemi intelligenti nell'ambito dell'impresa[13]. Un ambito di particolare interesse è quello dell'uso dell'IA per il monitoraggio delle infrastrutture pubbliche. L'uso di tali sistemi in questo settore consentirebbe di effettuare una manutenzione "predittiva" e non meramente "correttiva", individuando le problematiche di tali infrastrutture (magari non immediatamente riconoscibili dall'uomo), prevenendo possibili eventi lesivi colposi e riducendo i rischi per la pubblica sicurezza[14]. Il tema delle

---

[8] Purtroppo, per ragioni di sintesi, non ci è possibile dilungarci sull'argomento, per un approfondimento del quale si rinvia, per tutti, a F. CONSULICH, *Il nastro di Möbius. Intelligenza artificiale e imputazione penale nelle nuove forme di abuso del mercato*, in *Banca Borsa Titoli di credito*, 2018, pp. 195 ss. e M. PALMISANO, *L'abuso di mercato nell'era delle nuove tecnologie. Trading algoritmico e principio di personalità dell'illecito penale*, in *Dir. pen. cont. - Riv. Trim.* 2/2019, pp. 129 ss., i quali si concentrano sul ruolo dell'*actio libera in causa*, rispettivamente, il primo a pp. 219 ss. e la seconda a pp. 139 ss.; N. BELLOTTO, *High Frequency Trading. Un'indagine ricognitiva sulla rilevanza penale delle condotte manipolative del mercato realizzate dagli algoritmi*, in P. LAMBRINI (a cura di), *Quaderni del dottorato in giurisprudenza dell'Università di Padova*, Milano, 2021, pp. 11 ss.; R. COOPER, M. DAVIS, B.V. VLIET, *The Mysterious Ethics of High-Frequency Trading*, in *Business Ethics Quarterly*, gennaio 2016, pp. 1 ss.; B. BIAIS, T. FOUCAULT, *HFT and Market Quality*, in *Bankers, Markets & Investors*, n. 128, gennaio-febbraio 2014, pp. 5 ss.; V. CAIVANO, S. CICCARELLI, G. DI STEFANO, M. FRATINI, G. GASPARRI, M. GILIBERTI, N. LINCIANO, I. TAROLA, *Il* trading *ad alta frequenza. Caratteristiche, effetti, domande di* policy, Documenti di discussione CONSOB n. 5, in *SSRN*, dicembre 2012; EUROPEAN SECURITIES AND MARKETS AUTHORITY (ESMA), *High-frequency trading activity in EU equity markets*, 2014.

[9] C. CASONATO, *Potenzialità e sfide dell'intelligenza artificiale*, cit., p. 179.

[10] A proposito del robot Giraff S. BECK, *Google Cars, Software Agents, Autonomous Weapons Systems*, cit., pp. 228-229, «Il robot "Giraff" è uno schermo di computer su ruote telecomandato che può muoversi attraverso l'abitazione di una persona o può essere controllato da qualcuno che si trova in quell'abitazione e può consentire conversazioni online con parenti, operatori pastorali o medici e che consente a un medico di decidere se è necessaria un'ambulanza in caso di emergenza».

[11] Sull'esperienza di Google DeepMind del 2016 v. L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., p. 4.

[12] Sull'esperienza dell'uso degli *advanced analytics* da parte dell'azienda *General Electric* v. G. ITALIANO, *Intelligenza Artificiale*, cit., p. 221.

[13] Vorremmo a tal proposito ricordare l'esperienza di Vital, il robot membro del consiglio di amministrazione della società giapponese *Deep Knowledge*, reclutato per la sua capacità prevenire le tendenze di mercato, U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 622. Si pensi anche alla capacità dei sistemi di *machine learning* di svolgere attività tipicamente manageriali: sull'esperienza della *Bridgewater Associates*, M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 29.

[14] Prendendo le mosse dal disastro della funivia del Mottarone del maggio 2021, analizza approfonditamente il tema dell'impiego dei sistemi intelligenti nel settore delle infrastrutture pubbliche M. DI FLORIO, *Il diritto penale che verrà*, cit., pp. 1 ss., per la distinzione tra manutenzione predittiva e correttiva p. 8.

infrastrutture è stato anche affrontato dalla Proposta di Regolamento per la creazione di regole armonizzate sull'intelligenza artificiale[15], la quale considera "ad alto rischio" i sistemi di intelligenza artificiale utilizzati per garantirne la sicurezza[16].

Come è possibile intuire, non solo l'IA e le sue applicazioni sono dotate di carattere settoriale[17], ma esse sono anche in grado di introdurre innovative occasioni di rischio[18]: basti pensare all'uso dei sistemi di pilotaggio remoto[19], degli algoritmi predittivi in sede processuale per valutare la pericolosità sociale del reo[20], ma anche all'uso delle auto a guida autonoma e all'applicazione dell'IA

---

[15] Cap. I, Sez. II, Par. 9.1.

[16] *Proposta di Regolamento,* cit., Punto 34, ove testualmente si legge «Per quanto riguarda la gestione e il funzionamento delle infrastrutture critiche, è opportuno classificare come ad alto rischio i sistemi di IA destinati a essere utilizzati come componenti di sicurezza ai fini della gestione del traffico stradale nonché della fornitura di acqua, gas, riscaldamento ed elettricità, in quanto un loro guasto o malfunzionamento può mettere a rischio la vita e la salute di un grande numero di persone e provocare perturbazioni significative del normale svolgimento delle attività sociali ed economiche».

[17] C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 1.

[18] U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning,* cit., p. 19.

[19] Sul tema C. CUCCO, *La partita del diritto penale nell'epoca dei "*drones-crimes*"*, in *Dir. pen. cont. - Riv. Trim.* 2/2019, pp. 304 ss. Si dedicano all'argomento anche C. SALAZAR, *Umano, troppo umano,* cit., pp. 273 ss.; I. SALVADORI, *Agenti artificiali,* cit., p. 110; M.B. MAGRO, *Robot,* cit., p. 1207; A. TURANO, *Robotica e roboetica*, cit., p. 143; F. BASILE, *Intelligenza artificiale e diritto penale*, cit., pp. 28-29; A. ZORNOZA, M. LAUKYTE, *Robotica e diritto*, cit., p. 817; F. BASILE, *Intelligenza artificiale e diritto penale*, cit., p. 24 nota 91; R. BORSARI, *Intelligenza Artificiale e responsabilità penale*, cit., p. 263. Sull'uso delle armi autonome v. G. TAMBURRINI, *Etica delle macchine. Dilemmi morali per robotica e intelligenza artificiale*, Roma, 2020, pp. 97 ss.; D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., pp. 37 ss; R. CINGOLANI, D. ANDRESCIANI, *Robots,* cit., p. 52; U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., pp. 621-622; U. PAGALLO, *Saggio sui robot e il diritto penale,* cit., p. 598. Più in generale, sul tema dell'imputazione della responsabilità E. GRECO, *Profili di responsabilità penale del controllore del traffico aereo. Gestione del rischio e imputazione dell'evento per colpa nei sistemi a interazione complessa*, Torino, 2021.

[20] In tema del rapporto tra intelligenza artificiale e processo penale si connota per una moltitudine di sfaccettature. Non essendoci possibile approfondire il tema per ragioni di sintesi, si rinvia per un approfondimento generale alla corposa bibliografia sul tema: S. QUATTROCOLO, *Forecasting the future while investigating the past. The use of computational models in pre-trial detention decisions*, in *Revista Brasileira de Direito Processual Penal*, 2021, pp. 1859 ss.; EAD., *Artificial intelligence, computational modelling and criminal proceedings*, Svizzera, 2020; EAD., *Equità del processo penale e* automated evidence *alla luce della Convenzione Europea dei Diritti dell'Uomo*, in *Revista ítalo-española de Derecho Procesal*, 2019, pp. 107 ss.; EAD., *Equo processo penale e sfide della società algoritmica*, in *BioLaw Journal*, 2019, pp. 135 ss.; EAD., *Quesiti nuovi e soluzioni antiche? Consolidati paradigmi normativi vs rischi e paure della giustizia digitale "predittiva"*, in *Cass. Pen.*, 2019, pp. 1748 ss.; EAD., *An introduction to AI and criminal justice in Europe*, in *Revista Brasileira de Direito Processual Penal*, 2019, pp. 1519 ss.; EAD., *Intelligenza artificiale e giustizia,* cit.; A.M. MAUGERI, *L'uso di algoritmi predittivi per accertare la pericolosità sociale: una sfida tra* evidence based practices *e tutela dei diritti fondamentali*, in *Archivio Penale*, 1/2021; G. CANZIO, *Intelligenza artificiale, algoritmi e giustizia penale*, in *Sistema Penale*, 8.1.2021; S. QUATTROCOLO, C. ANGLANO, M. CANONICO, M. GUAZZONE, *Technical Solutions for Legal Challenges: Equality of Arms in Criminal Proceedings*, in *Global Jurist*, 2020; G. TUZET, *L'algoritmo come pastore del giudice? Diritto, tecnologie, prova scientifica*, in *MediaLaws*, 16.3.2020; S. SIGNORATO, *Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo*, in *Rivista di diritto processuale*, 2/2020, pp. 605 ss.; O. DI GIOVINE, *Il* judge-bot *e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione giurisprudenziale)*, in *Cass. Pen.*, 3/2020, pp. 951 ss.; L. D'AGOSTINO, *Gli*

nel settore medico. Pur nella consapevolezza che ciascuno di tali temi meriterebbe autonoma trattazione, e senza alcuna pretesa di esaustività, vorremmo sinteticamente analizzare gli ultimi due fenomeni summenzionati, al fine di proporre ulteriori spunti di riflessione su questa sorta di "parte speciale" della materia. Vorremmo in tal sede, dunque, concentrarci sull'applicazione dell'IA nel settore della circolazione stradale ed in quello dell'attività medica in quanto terreni d'elezione della colpa penale, provando a far applicazione – in questi specifici campi d'analisi – dei più generali approfondimenti svolti nelle pagine precedenti in ordine al possibile uso dei classici meccanismi di imputazione penale.

2. Le auto a guida autonoma: un preliminare inquadramento normativo.

Il settore delle *self-driving cars* è probabilmente quello che desta maggior interesse. Dopo essere passati dai veicoli condotti dall'intelligenza animale ai veicoli governati dall'intelligenza umana ci troviamo, nuovamente, a cedere il controllo ad un'intelligenza altra rispetto alla nostra, ossia quella artificiale[21]. Se, da un lato, ci si dovrà interrogare sulla "capacità di mediazione giuridica" delle norme già vigenti in materia di circolazione stradale rispetto alla materia dei veicoli integrati da sistemi intelligenti[22], di contro ci si dovrà confrontare con i nuovi testi normativi che sono già stati emanati e con quelli che (indubbiamente) vedranno la luce.

Un primo passo verso questa nuova strada è rappresentato dall'emendamento introdotto nel 2016 all'art. 8 della Convenzione internazionale sulla circolazione stradale di Vienna del 1968, il quale ha in parte eliminato il riferimento all'obbligo di controllo del conducente sul veicolo, in modo da consentire

---

algoritmi predittivi per la commisurazione della pena, in *Dir. pen. cont. - Riv. Trim.* 2/2019, pp. 354 ss.; B. OCCHIUZZI, *Algoritmi predittivi: alcune premesse metodologiche*, in *Dir. pen. cont. - Riv. Trim.* 2/2019, pp. 392 ss.; C. PARODI, V. SELLAROLI, *Sistema penale e intelligenza artificiale: molte speranze e qualche equivoco*, in *dirittopenalecontemporaneo.it*, fasc. 6/2019, pp. 47 ss.; M. GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei* risk assessment tools *tra Stati Uniti ed Europa*, in *dirittopenalecontemporaneo.it*, 29.5.2019; G. CONTISSA, G. LASAGNI, G. SARTOR, *Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo*, in *Diritto di Internet*, 4/2019, pp. 619 ss.; G. RICCIO, *Ragionando su intelligenza artificiale e processo penale*, in *Archivio Penale*, 3/2019; G. ZARA, *Tra il probabile e il certo. La valutazione del rischio di violenza e di recidiva criminale*, in *dirittopenalecontemporaneo.it*, 20.5.2016. Nella letteratura straniera sul tema v. V. CHIAO, *Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice*, in *International Journal of Law in Context*, 2019, pp. 126 ss.; D. KEHL, P. GUO, S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*, in *Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School*, 2017.

[21] U. RUFFOLO, *Self-driving car, auto driverless e responsabilità*, in U. RUFFOLO, *Intelligenza artificiale e responsabilità*, cit., p. 37; U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles*, cit., p. 1706. Dello stesso A. v. U. RUFFOLO, *Intelligenza Artificiale ed* automotive*: le responsabilità da veicoli* self-driving *e* driverless, in U. RUFFOLO, *Intelligenza artificiale. Il diritto, i diritti, l'etica*, cit., pp. 153 ss. e U. RUFFOLO, *Le responsabilità da produzione, proprietà e "conduzione" di veicoli autonomi*, in U. RUFFOLO, *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, cit., pp. 163 ss.

[22] U. RUFFOLO, *Self-driving car, auto driverless e responsabilità*, cit., p. 45; U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles*, cit., p. 1704.

l'operare dei sistemi autonomi[23]. In realtà, pur a seguito di tale emendamento, parte della dottrina ritiene che la Convenzione di Vienna continui ad imporre la presenza di un "conducente sentinella" a bordo del veicolo, anche ove dovesse trattarsi di un'auto *driverless*[24]. In realtà però potrebbe non essere un fuor d'opera attendersi che, in un futuro non troppo prossimo connotato da una capillare diffusione delle auto senza conducente, si provveda ad una nuova modifica della Convenzione che non lasci dubbi sul punto. In un simile contesto, ci si chiede se l'utilizzatore del veicolo autonomo «dovrebbe allora vedersi responsabilizzato per gli errori di conduzione automatizzata negli stessi termini in cui è responsabilizzato il "conducente" che adotta una guida umana non assistita»[25].

Anche l'Unione Europea non ha mancato di far sentire la propria voce in materia di circolazione stradale[26], anche nella sua forma automatizzata[27].

---

[23] *Convenzione sulla circolazione stradale*, conclusa a Vienna l'8 novembre 1968. Art. 8 (rubricato "Conducenti") comma 5 *bis* (introdotto dagli emendamenti del 26 marzo 2014, in vigore dal 23 marzo 2016 (RU 2016 1019): «I sistemi di bordo che influiscono sulla guida del veicolo sono considerati conformi al paragrafo 5 del presente articolo e al primo paragrafo dell'articolo 13 se sono conformi alle disposizioni in materia di costruzione, montaggio e utilizzo previste negli strumenti giuridici internazionali riguardanti i veicoli a ruote e gli equipaggiamenti e componenti montati e/o utilizzati sugli stessi. I sistemi di bordo che influiscono sulla guida del veicolo e non conformi alle disposizioni in materia di costruzione, montaggio e utilizzo summenzionate sono considerati conformi al paragrafo 5 del presente articolo e al primo paragrafo dell'articolo 13 se possono essere neutralizzati o disattivati dal conducente». Tale neointrodotto comma costituisce una parziale deroga a quanto stabilito al comma 1 del medesimo articolo, il quale sancisce che «*Ogni veicolo in movimento* o ogni complesso di veicoli in movimento *deve avere un conducente*» e al comma 5, il quale chiarisce che «*Ogni conducente deve avere costantemente il controllo del proprio veicolo* o deve poter guidare i propri animali», corsivi nostri. Sul punto v. G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture*, Modena, 2018, p. 96. Invero, ancor prima della Convenzione di Vienna, altre due convenzioni internazionali erano intervenute sul tema della circolazione stradale al fine di stabilire, negli Stati aderenti, regole uniformi per incrementare la sicurezza su strada, ci stiamo riferendo alla Convenzione di Parigi del 24.4.1926 e alla Convenzione di Ginevra del 19.9.1949.

[24] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles*, cit., p. 1711.

[25] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles*, cit., p. 1708.

[26] Basti riportare a tal proposito Direttiva 2007/46/CE del Parlamento europeo e del Consiglio del 5 settembre 2007 *che istituisce un quadro per l'omologazione dei veicoli a motore e dei loro rimorchi, nonché dei sistemi, componenti ed entità tecniche destinati a tali veicoli*; alla *Relazione della Commissione al Parlamento europeo e al Consiglio, Salvare vite umane: migliorare la sicurezza dei veicoli nell'UE,* Bruxelles, 12.12.2016 COM(2016) 787 final, cui ha fatto seguito la *Risoluzione del Parlamento europeo, Salvare vite umane: migliorare la sicurezza dei veicoli nell'UE*, del 14 novembre 2017 la quale, al Punto 17, afferma che «dovrebbe essere *obbligatoria l'installazione* unicamente *dei sistemi di assistenza alla guida* che contribuiscono in modo significativo e scientificamente provato al miglioramento della sicurezza stradale, presentano un rapporto costi/benefici positivo e sono pronti per la commercializzazione», corsivo nostro; al Regolamento n. 79 della Commissione economica per l'Europa delle Nazioni Unite (UNECE) - *Disposizioni uniformi relative all'omologazione dei veicoli per quanto riguarda lo sterzo*, del 16 ottobre 2018.

[27] Pensiamo alla Direttiva 2010/40/UE del Parlamento europeo e del Consiglio del 7 luglio 2010 *sul quadro generale per la diffusione dei sistemi di trasporto intelligenti nel settore del trasporto stradale e nelle interfacce con altri modi di trasporto*; alla già esaminata *Risoluzione del Parlamento europeo,* cit., la quale, dal Punto 24 in poi, dedica la sua attenzione ai veicoli autonomi; alla Comunicazione della Commissione al Parlamento europeo, al Consiglio europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni, *Verso la mobilità automatizzata: una strategia dell'UE per la mobilità del futuro*, Bruxelles, 17.5.2018 COM(2018) 283 final.

Particolarmente rilevante in tal senso risulta la Risoluzione del Parlamento europeo sulla guida autonoma nei trasporti europei[28], la quale evidenzia come lo sviluppo dell'intelligenza artificiale nel settore dei trasporti sia in grado di migliorarne la sicurezza, riducendo il numero delle vittime della strada. Ciò anche in considerazione del fatto che i sistemi di assistenza alla guida[29] (di cui l'UE auspica un uso obbligatorio) hanno già dimostrato la loro utilità in materia[30]. L'Unione annovera nella nozione di trasporto autonomo «tutte le forme a pilotaggio remoto, automatizzate e autonome di trasporto su strada, ferroviario, aereo, marittimo e per vie navigabili interne», evidenziando come all'evoluzione tecnologica dei veicoli dovrebbe corrispondere uno sviluppo delle infrastrutture secondo canoni comuni a livello internazionale[31]. La presente Risoluzione prende espressamente in esame la questione dell'attribuzione della responsabilità in caso di incidenti causati da veicoli autonomi, invitando la Commissione a ripensare la normativa vigente o, se del caso, ad emanarne una nuova al fine di stabilire in modo chiaro il riparto delle responsabilità[32].

Volgendo lo sguardo alla prospettiva nazionale, invece, non possiamo fare a meno di menzionare il Decreto Smart Road, pubblicato nel 2018 dal Ministero delle Infrastrutture e dei Trasporti, il quale definisce il "veicolo a guida automatica"[33] come «un veicolo dotato di tecnologie capaci di adottare e attuare comportamenti di guida senza l'intervento attivo del guidatore, in determinati ambiti stradali e condizioni esterne»[34]. Il presente decreto individua tra le proprie finalità la valorizzazione delle infrastrutture esistenti, nonché il loro adeguamento tecnologico, prevedendo altresì un articolato meccanismo di autorizzazioni volto a concedere la possibilità di effettuare la sperimentazione di tali veicoli autonomi su strada[35].

---

[28] *Risoluzione del Parlamento europeo del 15 gennaio 2019 sulla guida autonoma nei trasporti europei* (2018/2089(INI)), (2020/C 411/01).

[29] Sul punto la dottrina di settore afferma che «sebbene la tecnologia attualmente consentita sia limitata a funzionalità come il cruise control adattivo, l'assistenza al parcheggio con sterzo automatizzato e l'assistenza al mantenimento della corsia, i veicoli completamente autonomi, che possono guidare in qualsiasi luogo in cui è legale guidare e prendere le proprie decisioni senza l'intervento umano, o quasi completamente autonomi, in grado di controllare il veicolo in tutti gli ambienti tranne alcuni, come il maltempo, e che non richiedono l'attenzione del conducente umano, potrebbero essere in grado di prendere parte al traffico ordinario nei prossimi decenni» H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law*, cit., pp. 341-342.

[30] *Risoluzione del Parlamento europeo sulla guida autonoma nei trasporti europei*, cit., Considerando B, D e E.

[31] *Risoluzione del Parlamento europeo sulla guida autonoma nei trasporti europei*, cit., Considerando O e Punto 14.

[32] *Risoluzione del Parlamento europeo sulla guida autonoma nei trasporti europei*, cit., Punto 21.

[33] Volendo condividere la semantica d'oltreoceano, la National Highway Traffic Safety Administration statunitense (da qui in poi sinteticamente "NHTSA") definisce le auto a guida autonoma come quei veicoli in cui: «alcuni aspetti di una funzione di controllo critica per la sicurezza (ad es. sterzo, acceleratore o frenata) si verificano senza l'input diretto del conducente», NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION, *Preliminary Statement of Policy Concerning Automated Vehicles*, 30.5.2013, p. 3.

[34] Decreto 28 febbraio 2018, *Modalità attuative e strumenti operativi della sperimentazione su strada delle soluzioni di Smart Road e di guida connessa e automatica*, G.U. n. 90 del 18.04.2018, art. 1, lett. f).

[35] Art. 3 e artt. 9 ss. *Decreto Smart Road*. Nel quadro che abbiamo finora delineato di particolare interesse risulta essere l'art. 12, il quale stabilisce che «Ai fini dell'autorizzazione alle prove su

Come anticipato dalla Risoluzione pocanzi sinteticamente esaminata, l'avvento delle auto a guida autonoma richiederà anche la predisposizione di infrastrutture all'avanguardia, con capacità di connettività non indifferenti. Ciò significa che le nostre strade dovranno essere adeguate per consentire alle macchine a guida autonoma di viaggiare in un ambiente ad esse favorevole. In altre parole, si dovrà "*avvolgere*" l'ambiente intorno all'IA per consentirle di adattarsi e di sfruttare le sue capacità[36]. In realtà un recepimento a livello interno di tale assunto è rinvenibile anche all'art. 2 del Decreto Smart Road, ove si chiarisce che per le infrastrutture stradali dovrà essere compiuto un «processo di trasformazione digitale orientato a introdurre piattaforme di osservazione e monitoraggio del traffico, modelli di elaborazione dei dati e delle informazioni, servizi avanzati ai gestori delle infrastrutture, alla pubblica amministrazione e agli utenti della strada, nel quadro della creazione di un ecosistema tecnologico favorevole all'interoperabilità tra infrastrutture e veicoli di nuova generazione»[37].

Come è possibile intuire da queste sintetiche battute, pur mancando ancora un approccio sistemico[38], il tema non è sfuggito all'attenzione delle istituzioni le quali, pur forse non avendo una complessiva e ben delineata percezione del fenomeno, stanno iniziando a tracciare le linee guida dell'argomento, intuendone la portata innovativa e le potenzialità incisive.

2.1. Una questione di riconoscimento sociale.

In realtà, a ben vedere, quello delle auto a guida autonoma non costituisce affatto un fenomeno di nuovo conio. Già nel 2005 un'auto costruita a Stanford si aggiudicò la *Grand Challenge*, una competizione di guida robotica indetta dalla DARPA (*Defence Advanced Research Project Agency*), guidando per 131 miglia nel deserto. Nel 2007 un'auto costruita a Carneige-Mellon vinse la DARPA

---

strade pubbliche, il sistema di guida automatica oggetto di sperimentazione deve: (…) d) essere dotato di *protezioni intrinseche di sicurezza* atte a garantire l'integrità dei dati e la sicurezza delle comunicazioni e tali da scongiurare accessi non autorizzati e, in ogni caso, da vanificarne gli effetti dannosi o pericolosi», così facendo il decreto prende in seria considerazione la necessità di prevenire illeciti attacchi hacker diretti ai software delle auto intelligenti; «e) essere in grado, per tutta la durata delle prove, di *registrare dati dettagliati*», così prevenendo l'esaminato fenomeno del *black box effect* e potendo ricostruire la dinamica di eventuali incidenti. L'Art. 16 sancisce che «2. il titolare dell'autorizzazione alla sperimentazione su strada per tutta la durata dell'autorizzazione, è tenuto a produrre e consegnare al soggetto autorizzante: a) il rapporto puntuale su eventi o *problematiche* di qualsiasi natura che hanno coinvolto il sistema sperimentato e che possono avere risvolti ai fini della *sicurezza* anche solo potenziali» richiamando alla nostra memoria non soltanto il ruolo della classica posizione di garanzia, declinata in termini di posizione di controllo, ma anche il ruolo cardine del principio di precauzione, con riferimento all'inciso relativo alle problematiche in materia di sicurezza che siano "anche solo potenziali". Il principio di precauzione può essere richiamato anche, a nostro avviso, a proposito dell'art. 18 il quale stabilisce che «il soggetto autorizzante può sospendere o revocare l'autorizzazione se ravvisa, anche a seguito di inadempienze del soggetto autorizzato e di segnalazioni relative a quanto emerso in sede di controlli su strada, che il proseguimento delle sperimentazioni *può* causare un *rischio* per la sicurezza della circolazione», corsivi nostri.

[36] Per un approfondimento del concetto di "*enveloping*" cfr. L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., p. 11.

[37] Art. 2 *Decreto Smart Road*.

[38] U. RUFFOLO, *Self-driving car, auto driverless e responsabilità*, cit., pp. 35 ss. Anche G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* cit., p. 125 evidenzia che «molte sono le normative applicabili, poca è la chiarezza».

Urban Challenge percorrendo 55 miglia in una zona urbana nel traffico e con la normale segnaletica[39]. Ad oggi sono moltissimi i progetti di ricerca[40] sviluppati in materia, per quanto inevitabilmente quelli più noti al grande pubblico (nonché protagonisti dei primi incidenti stradali)[41] siano quelli portati avanti dai c.d. "big" del settore: Google Car (ora Waymo), Uber e Tesla[42].

In dottrina è stata tracciata una distinzione di massima tra *selfdriving car* e *driverless car*: la prima richiederebbe ancora la presenza vigile del conducente, pur potendo quest'ultimo rimanere inerte; la seconda, invece, si candida a divenire destinataria esclusiva della circolazione stradale in quanto, essendo sprovvista di

[39] G. ITALIANO, *Intelligenza Artificiale,* cit., p. 216.

[40] Sui progetti incoraggiati dalla Commissione Europea riguardanti la valutazione delle cause dei sinistri stradali, segnatamente TRACE (*Traffic Accident Causation in Europe*) e ASSESS (*Assessment of Integrated Vehicle Safety System*), G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* cit., p. 105.

[41] Ci limitiamo in tal sede a riportare sinteticamente alcuni degli incidenti stradali causati da veicoli a guida autonoma che hanno avuto maggior impatto mediatico.

L'11 dicembre 2021 a Parigi un taxi Tesla ha accelerato improvvisamente causando un incidente dove sono rimasti coinvolti circa 20 feriti ed un deceduto. Il tassista, dal canto proprio, si è difeso riferendo di aver provato a frenare senza esito alcuno. I vertici Tesla, di contro, hanno affermato che, a seguito di una preliminare indagine, era emerso che non fosse occorso alcun malfunzionamento tecnico del veicolo. La notizia è consultabile sul sito https://news.sky.com/story/tesla-accident-leaves-one-dead-and-20-injured-in-paris-prompting-taxi-firm-to-suspend-use-of-model-12496220#.

Il 7 maggio 2016 a Williston in Florida un uomo, a bordo della sua Tesla (Model S) è rimasto vittima di un tragico incidente. L'autopilot del veicolo non è riuscito a visualizzare un camion a cagione della sua eccessiva altezza, andandosi ad incastrare nella parte inferiore del mezzo. I sensori dell'auto, inoltre, hanno confuso il colore bianco del rimorchio con il colore del cielo. Dalle indagini è emersa anche la disattenzione del conducente, il quale non si si sarebbe accorto del camion. Ciò sarebbe stato dimostrato dal fatto che il freno non risultava essere stato neanche toccato. Menziona l'incidente D. TAFANI, *Sulla moralità artificiale,* cit., p. 86. Per un maggiore approfondimento v. https://www.teslaclub.it/la-responsabilita-dell-incidente-mortale-con-il-pilota-automatico-tesla.html.

Il 18 marzo 2018 a Tempe in Arizona una donna è stata travolta da un'auto Uber (Volvo XC90) mentre attraversava la strada fuori dalle strisce pedonali (il sistema, infatti, non era stato programmato per riconoscere i pedoni fuori dalla segnaletica orizzontale). La commissione della NTSB (*National Transportation Safety Board*) ha ripartito la colpa tra l'uomo e la macchina in quanto ha ritenuto che la pilota tester a bordo dell'auto fosse distratta dal telefono ma anche che la procedura di valutazione del rischio per la sicurezza del gruppo *Advanced Technologies* di Uber fosse inadeguata. Tale incidente viene riportato anche da F. BASILE, *Intelligenza artificiale e diritto penale,* cit., p. 24; M.B. MAGRO, *Decisione umana e decisione robotica,* cit., p. 4; A. VENANZONI, *Intersezioni costituzionali - Internet e Intelligenze Artificiali tra ordine spontaneo, natura delle cose digitale e garanzia dei diritti fondamentali*, in *Forum di Quaderni Costituzionali*, 27.4.2018, p. 4. Per un approfondimento v. https://www.repubblica.it/tecnologia/2019/11/21/news/uber_auto_a_guida_autonoma_responsabilita_umana_e_del_software_per_l_incidente_mortale-241593528/.

Il 23 marzo 2018 a Mountain View in California ha trovato la morte il conducente di una Tesla (Model X). L'autopilot del veicolo risultava attivo ma il conducente, essendo intento a giocare al cellulare e, quindi, non avendo le mani sul volante, ha contribuito alla verificazione dell'evento. Accenna all'incidente anche A. CAPPELLINI, *Profili penalistici delle self-driving cars*, cit., p. 331 nota 32. https://www.hdmotori.it/tesla/articoli/n516697/tesla-model-x-problemi-autopilot-incidente-2018/.

Per un elenco degli incidenti registrati che hanno coinvolto le auto a guida autonoma cfr. https://www.ranker.com/list/self-driving-car-accidents/eric-vega.

[42] A. CAPPELLINI, *Profili penalistici delle self-driving cars*, cit., p. 326.

comandi manuali, non richiederebbe più la figura del guidatore[43]. Ad oggi un'indicazione sistemica sul punto ci perviene dagli Stati Uniti d'America, ove la *SAE International* (*Society of Automotive Engineer*) ha redatto gli *standard J3016*, i quali individuano 6 livelli di automazione, dal livello 0 (nessuna automazione di guida) al livello 5 (automazione di guida completa)[44]. Man mano che si avanza dal primo all'ultimo livello di automazione il ruolo del conducente sia va sempre più dissolvendo e, con esso, la possibilità di considerarlo pacificamente come autore materiale di un eventuale incidente causato mediante (o dal?) il veicolo autonomo.

La dottrina ha individuato tre fasi di funzionamento di tali veicoli: la *sensing phase*, rimessa ai sensori dell'auto, durante la quale quest'ultima raccoglie informazioni su cosa la circonda; la *planning phase*, in cui il veicolo, tramite i propri software, elabora le informazioni raccolte e determina un piano d'azione (ad es. in quale direzione è più sicuro muoversi e a quale velocità); infine la *acting phase*, ossia la fase in cui le componenti hardware (es. motore, freno ecc.) eseguono il piano d'azione ideato dal software[45]. Gli operatori del settore hanno rilevato che probabilmente uno dei compiti più difficili da realizzare in fase di progettazione di questi sistemi sia proprio quello di dare un senso ai dati raccolti dai loro sensori[46]. Il problema non è di poco conto. Ci sono diversi oggetti che un veicolo autonomo deve essere in grado di riconoscere: distinguere un segnale stradale da un altro, percepire il rosso del semaforo e non confonderlo dal verde, rilevare la segnaletica orizzontale, differenziare un'auto comune da un'ambulanza

---

[43] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles*, cit., p. 1704.

[44] SAE J3016, *Levels of driving automation*. I livelli 0, 1 e 2 sono caratterizzati dalla presenza di un conducente attivo, chiamato non soltanto ad agire (ad es. sterzando, frenando o accelerando) ma anche a supervisionare le funzioni di supporto al conducente. Al livello 3 le funzioni di guida automatizzata sono attive ma spetterà al conducente (potendolo ancora considerare tale) riprendere i comandi della vettura ove sia essa stessa a richiederlo. Il livello 4 non richiede una condotta di guida attiva (essendo possibile che manchino pedali o sterzo), potendo però circolare solo al ricorrere di determinate condizioni. Il livello 5, invece, possiede tutte le caratteristiche del livello 4 potendo, in aggiunta, circolare senza limiti di condizioni. Questi standard sono stati aggiornati in via di specificazione nel 2021 e sono consultabili al sito https://www.sae.org/standards/content/j3016_202104/. Tale classificazione è stata accolta anche in NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION, *Preliminary Statement of Policy Concerning Automated Vehicles*, cit., pp. 4 ss. Per un approfondimento v. A. CAPPELLINI, *Profili penalistici delle self-driving cars*, cit., pp. 328 ss.; D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto*, cit., p. 42; C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 104 nota 5; D. TAFANI, *Sulla moralità artificiale*, cit., p. 85 nota 12; G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* p. 104; P. SEVERINO, *Intelligenza artificiale*, cit., p. 533 nota 6.

[45] H. SURDEN, M.A. WILLIAMS, *Technological opacity, predictability, and self-driving cars*, in *Cardozo Law Review*, 2016, p. 141.

[46] H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law*, cit., p. 353-354. Ci sono però momenti della circolazione stradale che difficilmente possono essere colti dalle auto a guida autonoma e che, invece, per un uomo sarebbero intuitivi. La dottrina da ultimo citata infatti afferma che «lo scopo di promuovere un traffico sicuro ed efficiente richiede che un AV [*autonomous vehicle*] abbia la capacità di riconoscere e conformarsi ai segnali sociali. Ad esempio, dovrebbe essere in grado di interpretare i gesti o il contatto visivo con i partecipanti umani al traffico. Mentre per gli umani ciò è generalmente semplice, per gli AV anche questo sembra essere un problema complesso. Più in generale, gli AV dovrebbero avere la capacità di anticipare il comportamento dei partecipanti al traffico umano, poiché questa capacità è essenziale per evitare incidenti» p. 354.

alla quale dare la precedenza o distinguere un pedone sul ciglio della strada da un membro delle forze dell'ordine in divisa che intima all'auto di fermarsi[47].

È stato ravvisato come la scelta di immettere sul mercato veicoli con progressivi gradi di automazione sia stata probabilmente volta a facilitare l'approccio del conducente ai sistemi di guida autonoma più sofisticati. Se le tecnologie attualmente circolanti lasciano residuare in capo all'uomo il ruolo di "decisore finale" – restando dunque ferma l'attribuzione della responsabilità in capo a quest'ultimo con riguardo agli eventuali incidenti stradali occorsi – col sopraggiungere dei più sviluppati sistemi autonomi saranno essi stessi a porre coattivamente in essere "misure di *recovery*", volte cioè a rimettere il veicolo in sicurezza, indipendentemente dall'intervento dell'uomo[48]. Ancora, con il progredire di questi sistemi di guida sarà possibile non solo collegare questi ultimi ad internet ma anche connetterli tra di loro: un veicolo autonomo, infatti, potrebbe rispondere in modo più appropriato alle problematiche del traffico se conoscesse in anticipo dagli altri veicoli informazioni utili a migliorare il proprio itinerario o la sua velocità di crociera[49].

Tutti questi fattori sono volti ad ottimizzare la circolazione su strada e a ridurre il margine di errore umano ma, al contempo, non possiamo fare a meno di notare come essi, da un lato, "inducano" il conducente ad essere meno vigile quando si trova alla guida (ammesso che questa espressione sia ancora adoperabile per quanto riguarda le auto *driverless*)[50] e, dall'altro, riducano considerevolmente – fino quasi a sopprimerlo – il potere di controllo dell'uomo sul veicolo e, conseguentemente, la possibilità di evitare un evento lesivo dovuto a un malfunzionamento della macchina. Precipitato logico-giuridico di tale mancanza di controllo sarà la corrispondente difficoltà di imputazione della responsabilità penale in caso di incidente stradale cagionato dal veicolo autonomo: riprendendo il fulcro del ragionamento svolto nel capitolo precedente, chi dovrà essere considerato responsabile per un illecito commesso dall'IA?

Anche qui, a nostro avviso, buona parte della partita si giocherà sul campo del riconoscimento sociale di tali veicoli e della relativa perimetrazione del rischio consentito. Avveduta dottrina ha evidenziato come «la *tollerabilità sociale di un rischio* sovente affondi le sue *radici* più profonde non tanto nella sua dimensione oggettiva, scientifica, quanto piuttosto in *deformanti pregiudizi e paure anche irrazionali*, in larga parte legati fortemente al *momento comunicativo del supposto pericolo*»[51]. Non a caso si registra attualmente una certa diffidenza sul punto, la quale, inevitabilmente, porta con sé un atteggiamento di *precauzione*. Se da un lato tale *modus operandi* si rivela condivisibile, in quanto comunque connesso a una condizione di incertezza scientifica che caratterizza la materia in esame, occorre nondimeno rifuggire dalla china scivolosa delle possibili derive preclusive del principio di precauzione. Un eccessivo ricorso a tale ultimo principio comporterebbe l'astensione dallo sviluppo di un settore che, invece, si candida ad

[47] L. GOMES, *Hidden obstacles for Google's self-driving cars*, in *MIT Technology Review*, 2014. «Pedestrians are detected simply as moving, column-shaped blurs of pixels (…), the car wouldn't be able to spot a police officer at the side of the road frantically waving for traffic to stop».
[48] G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* pp. 101-102.
[49] S. BECK, *Google Cars, Software Agents, Autonomous Weapons Systems*, cit., p. 230.
[50] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles*, cit., p. 1710.
[51] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 332, corsivo nostro.

essere il futuro della mobilità mondiale. Rinunciare a sviluppare le ricerche in materia comporterebbe rinunciare a tutti i potenziali vantaggi sociali che essa porterebbe con sé[52].

Tra i *vantaggi* dell'automazione della circolazione stradale la dottrina ha individuato l'incremento della sicurezza stradale[53] (basti pensare, a titolo esemplificativo, all'eliminazione degli incidenti causati da distrazione, guida in stato di ebbrezza o sotto l'effetto di sostanze psicotrope)[54], l'aumento dell'offerta di mobilità[55], la diminuzione delle emissioni ambientali nocive[56] e l'incremento dell'efficienza della circolazione veicolare[57]. Se, da un lato, l'introduzione di veicoli autonomi e semi autonomi avrà il grande pregio di ridurre i sinistri stradali (*macrofenomeno*), di contro con essi verrà introdotto un minor numero di incidenti che andrà a ledere soggetti che, magari, non sarebbero mai rimasti coinvolti in alcun incidente se tali sofisticati veicoli non fossero stati immessi su strada (*microfenomeno*)[58]. È, infatti, evidente che l'introduzione di questi nuovi

[52] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 340.

[53] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., p. 43; G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* p. 95. Secondo la Commissione Europea circa il 95% degli incidenti stradali è da ricondurre al fattore umano, *Relazione della Commissione al Parlamento Europeo e al Consiglio. Salvare vite umane: migliorare la sicurezza dei veicoli nell'UE,* 12.12.2016, cit., p. 4.

[54] Una totale eliminazione di tale rischio sarebbe riferita esclusivamente alle auto completamente autonome, in quanto per i veicoli semi-autonomi rimarrebbe in capo al conducente un «obbligo di sorveglianza del mezzo durante le fasi di crociera, la perdita di reattività connessa all'uso di alcool o sostanze stupefacenti sarebbe comunque fonte di un rischio non consentito» A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 342. Sulla proposta di collegare un alcol test direttamente al veicolo e di renderne l'effettuazione obbligatoria prima dell'accensione C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?,* cit., p. 1937 nota 90.

[55] Di questo incremento dell'offerta di mobilità potrebbero usufruire alcune categorie di soggetti particolarmente svantaggiati, ad esempio disabili, anziani, giovani senza patente, S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., p. 413.

[56] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 330.

[57] Un interessante episodio dimostra come l'eccessiva sicurezza di un'auto a guida autonoma possa andare ad impattare sull'efficienza della sua circolazione. Un'auto autonoma di Google è stata fermata dalla polizia californiana perché, per motivi di sicurezza, andava troppo lenta. Il veicolo aveva formato dietro di sé una lunga coda di macchine nel traffico in quanto guidava a 24 kmh, H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law*, cit., p. 346. Altra ipotesi di intralcio alla circolazione stradale si verificherebbe ove, ad esempio, un veicolo restasse fermo dietro un altro che sta scaricando della merce senza effettuare una manovra di sorpasso. Certo è che non effettuare tale ultima manovra difficilmente aumenterà i rischi per la circolazione stradale ma, di certo, ne costituirà intralcio, G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* p. 97. Simili condotte costituiscono di certo intralcio alla circolazione stradale il quale, nel nostro ordinamento, è anche passibile di una sanzione amministrativa. Ci stiamo riferendo al disposto dell'art. 141 co. 6 del Codice della Strada il quale sancisce che «Il conducente non deve circolare a velocità talmente ridotta da costituire intralcio o pericolo per il normale flusso della circolazione». Non sfuggirà ad un attento lettore che il riferimento al "conducente" effettuato dal nostro codice potrebbe non adeguarsi totalmente a circostanze in cui il conducente manchi o sia ridotto a mero passeggero. La necessità che le auto a guida autonoma non costituiscano intralcio (oltre che pericolo) per la circolazione stradale è chiarito anche dall'art. 12 lett. a) del *Decreto Smart Road.*

[58] U. RUFFOLO, *Self-driving car, auto driverless e responsabilità,* cit., pp. 39-40. Alle vittime «però, in termini pratici, poco importa che la generale pericolosità da circolazione risulti diminuita grazie all'introduzione di forme di guida automatizzata, se quel tipo di sinistro che le ha colpite non si sarebbe verificato (o avrebbe avuto molto minori possibilità di verificarsi) qualora in circolazione vi fossero stati veicoli più tradizionali, caratterizzati dall'essere, in generale, meno

veicoli porterà con sé anche nuove tipologie di *rischi*, basti pensare ai problemi che queste macchine possono porre con riguardo a specifici errori di malfunzionamento[59] o ad attacchi hacker volti ad intaccarne la *cybersecurity*[60]. A fronte di innovative tipologie di rischi si va ampliando la "divaricazione" tra le prestazioni di cui è capace la vettura e il grado di conoscenza che ne possiede il suo proprietario[61].

In tal senso, però, occorre mettere in evidenza due passaggi fondamentali. *In primis* vorremmo sottolineare che la circolazione stradale delle auto a guida autonoma si caratterizza per una "pericolosità settoriale ed asimmetrica", ossia strettamente connessa alle peculiarità di suddetti veicoli[62]. *In secundis* ci sembra inappuntabile la considerazione di quella parte di dottrina la quale evidenzia che, se è vero che gli operatori del settore delle auto a guida autonoma creano ed introducono rischi per la vita e l'incolumità altrui, non bisogna dimenticare che lo stesso vale ancora per la produzione e la vendita delle auto c.d. tradizionali, riservate alla guida dell'uomo[63].

Nulla di nuovo sotto il sole. Il progredire dell'evoluzione tecnologica porta con sé, endemicamente, rischi più o meno nuovi, che la società ha sempre trovato il modo di affrontare ove si tratti di un campo rischioso ma di indubbia utilità sociale come la circolazione stradale. Probabilmente non si tratterà di capire quale tipologia di guida sia maggiormente foriera di eventi lesivi bensì, più semplicemente, di prendere atto che potremmo limitarci ad assistere ad una «mera sostituzione degli errori umani con gli errori delle macchine»[64].

Vi sarà un lungo periodo di transizione in cui le auto autonome circoleranno insieme alle auto tradizionali. Durante questo lasso di tempo si dovrà ancora fare i conti con l'errore umano, il che indurrà a chiederci come saranno gestite in questo periodo le questioni inerenti l'attribuzione della responsabilità per gli eventuali incidenti stradali verificatisi[65]. «Finché le *autonomous cars* circoleranno a fianco dei veicoli tradizionali, rimarranno sempre dei fattori di rischio connessi a tale commistione»[66], in quanto la crescita dei benefici connessi all'uso di auto a guida autonoma sarà proporzionata alla loro diffusione[67].

---

sicuri ma, in particolare, privi della specifica pericolosità che ha causato quegli incidenti» U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles,* cit., p. 1711.

[59] Pensiamo, ad esempio, ad un errato riconoscimento delle immagini da parte di autoveicoli autonomi che possa produrre risultati errati e, conseguentemente, incidenti anche gravi, C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 120.

[60] I veicoli autonomi possono subire attacchi da parte di hacker, i quali potrebbero non solo impossessarsi di informazioni riservate all'utente del veicolo, ma anche assumere il controllo di quest'ultimo, A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 345. Sulla possibilità di un controllo del veicolo da remoto in grado di cagionare un danno e sull'accesso abusivo alle informazioni registrate dal veicolo S. BECK, *Google Cars, Software Agents, Autonomous Weapons Systems*, cit., p. 231. Sul tema v. anche I. SALVADORI, *Agenti artificiali,* cit., p. 103. Non possiamo in tal sede fare a meno di chiederci se una simile condotta criminosa possa integrare una nuova ipotesi di accesso abusivo ad un sistema informatico o telematico ex art. 615 *ter* c.p.

[61] G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* p. 105.

[62] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles,* cit., p. 1711.

[63] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., p. 430.

[64] D. TAFANI, *Sulla moralità artificiale,* cit., p. 86.

[65] U. RUFFOLO, *Self-driving car, auto driverless e responsabilità,* cit., p. 42.

[66] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 330.

[67] G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* p. 95.

La via da percorrere sarà quella del bilanciamento di interessi, non solo tra i benefici e i pericoli che verranno introdotti da queste nuove tecnologie, ma anche tra le esigenze delle aziende che producono e immettono sul mercato questi veicoli e le istanze di tutela degli utenti della strada[68]. Occorrerà, infatti, «non dilatare, né normativamente né interpretativamente, previsioni di responsabilità scoraggianti o ritardanti l'ingresso delle nuove tecnologie»[69], auspicando che i benefici che tali vetture sono in grado di introdurre nella società siano considerati più "attraenti" dei possibili rischi[70]. Il ragionamento da seguire, infatti, dovrebbe essere del seguente tenore: dovremmo consentire l'ingresso sul mercato delle auto senza conducente in quanto, nonostante il rischio che esse cagionino incidenti stradali sia ineliminabile, esse sono pur sempre in condizione di garantire un minor numero di sinistri rispetto a quanto non faccia la tradizionale guida dell'uomo.

Tale operazione di bilanciamento tra benefici e rischi connessi all'immissione su strada delle auto a guida autonoma consentirà di delimitare (come pocanzi anticipato) l'area del rischio consentito. Tale valutazione dovrà tener conto non soltanto, da un punto di vista scientifico, della componente oggettiva del rischio, ma anche della percezione soggettiva che di esso avrà la società[71]. Tale ultimo aspetto sarà probabilmente l'ostacolo più difficile da superare, in quanto si registra già sul punto una certa resistenza fondata su «pregiudizi tecnologicamente illogici, eppure ingiustificatamente penetrati nel comune sentire (…) eppure, la incidenza statistica di tali sinistri è già oggi comparativamente molto bassa, a parità di chilometri percorsi, rispetto a quella da ordinaria circolazione»[72].

Occorrerà dunque rifuggire dalla "cultura del sospetto"[73] per andare a ricercare una *intermediate solution*, che introduca magari un margine di tolleranza per alcuni errori commessi in fase progettazione e programmazione di tali auto[74] e che faccia confluire la possibile imprevedibilità del comportamento dei veicoli autonomi nell'"orbita del rischio consentito"[75].

2.2. Profili penalistici dei veicoli semi-autonomi.

L'evoluzione di tali mezzi di trasporto è destinata ad incidere tanto sulla materia civile[76] quanto su quella penale. I diversi gradi di automazione cui

---

[68] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., p. 430.
[69] U. RUFFOLO, *Self-driving car, auto driverless e responsabilità,* cit., p. 39.
[70] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 329, il quale afferma che «Oltretutto, è ragionevole supporre, sia pur parlando in termini generalissimi, che il passare del tempo gradualmente addolcirà le stesse ritrosie sociali e i timori legati alla nuova tecnologia. Via via che i consociati inizieranno a familiarizzare con la guida autonoma, l'"ignoto tecnologico" a essa correlato diventerà sempre meno ignoto. Ciò, evidentemente, permetterà prospettive autorizzatorie sempre più ampie».
[71] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 329.
[72] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles,* cit., p. 1705.
[73] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles,* cit., p. 1710.
[74] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., pp. 430-431.
[75] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 341.
[76] «In private law, the possibility of holding entities other than natural persons liable for damages is nothing extraordinary; thus, the system can accommodate robot liability with no great difficulty» S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., pp. 414-415; sul tema, oltre ai già citati U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles,* cit., p.

abbiamo accennato portano con sé sfide diverse, con domande e risposte diverse. Ci si sente in tal sede di condividere l'impostazione metodologica di quella dottrina ad avviso della quale il problema dell'imputazione della responsabilità penale in caso di omicidio o lesioni stradali vada affrontato prendendo le mosse dalla distinzione tra veicoli semi-autonomi e totalmente autonomi[77], secondo «un'ottica gradualistica»[78] del controllo di tali mezzi.

Dal livello di automazione 0 al livello 2 il conducente riveste ancora il suo ruolo tradizionale: egli mantiene il costante controllo della vettura, potendo al più usufruire di alcune funzionalità di supporto alla guida (ADAS: *Advanced Driver Assistance Systems*)[79]. In queste ipotesi sembra pacifico affermare che nulla muti rispetto alla classica osservanza delle regole sulla circolazione stradale tradizionalmente rimessa al conducente, nonché idonea, in caso di una loro violazione, a integrarne la responsabilità colposa. Nessuna forzatura si rileverebbe, infatti, nell'applicare all'occorrenza gli artt. 589 *bis* e 590 *bis* c.p.

Il livello di automazione 3 inizia a palesare una maggiore delegazione di compiti alla vettura. «La principale distinzione tra il livello 2 e il livello 3 è che al livello 3 il veicolo è progettato in modo tale che il conducente non debba monitorare costantemente la carreggiata durante la guida»[80], ma ciò non significa che egli non debba comunque restar vigile e pronto a riprendere il controllo manuale del veicolo ove lo stesso lo richieda. Si usa a tal proposito parlare di manovre di *hand over* e *take over*, ossia di cessione e di ripresa del controllo della vettura a seguito della richiesta di quest'ultima. «Tali manovre creano complicati problemi in tema di responsabilità nel caso siano causative di un danno. Infatti, per assumere la responsabilità del conducente bisogna anche assumere che questo: a. abbia il tempo necessario per riprendere la conduzione del mezzo, aderendo alla richiesta di intervento del sistema; b. sia stato adeguatamente avvisato dai sistemi della vettura della richiesta; c. conosca esattamente i limiti del mezzo». L'uomo dovrà dunque abituarsi a non rivestire più una posizione di comando sul sistema intelligente, dovendo piuttosto dialogare con esso[81].

---

pp. 1704 ss. e U. RUFFOLO, *Self-driving car, auto driverless e responsabilità,* cit., pp. 31 ss., v. anche G. CALABRESI, E. AL MUREDEN, *Driverless cars. Intelligenza artificiale e futuro della mobilità*, Bologna, 2021; A. AMIDEI, *Robotica intelligente e responsabilità,* cit., p. 81; per qualche cenno a proposito del sistema di assicurazione v. M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 370.

[77] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 334.

[78] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1751.

[79] Basti pensare ai sistemi di «*adaptive cruise control*, *brake assistance*, *park assistance*, *lane keeping*», i quali in genere richiedono la supervisione del conducente, G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* pp. 111-112.

[80] NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION, *Preliminary Statement of Policy Concerning Automated Vehicles,* cit., p. 5. Parte della dottrina evidenzia come, nonostante tali sistemi nascano con il fine di ottimizzare la condotta di guida del conducente, il loro operare potrebbe impedire manovre "scorrette ma salvavita": «si pensi al comunissimo ABS, che, modulando le nostre frenate, impedisce, ad esempio, la frenata "panico" su ghiaccio evitando incidenti, ma impedisce altresì al pilota molto esperto di utilizzare scientemente la medesima manovra di frenata al deliberato fine di provocare una sbandata controllata idonea a scongiurare una collisione» U. RUFFOLO, *Self-driving car, auto driverless e responsabilità,* cit., p. 42.

[81] G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* pp. 98 ss., per la citazione p. 98. L'A. parla a tal proposito della c.d. *moral crumple zone* per spiegare come in capo al conducente si accavallino contemporaneamente una serie di problemi da risolvere non sempre del tutto esplicabili, p. 99. L'A. chiarisce più avanti che «il conducente necessita di adeguata preparazione

Con riferimento a tale livello di automazione sembra residuare una considerevole porzione di controllo in capo al conducente che, pertanto, dovrebbe ancora essere considerato tale quando si trova a bordo di un veicolo con automazione 3. Colui il quale si pone alla guida di un veicolo "intelligente" risulterà infatti esposto non soltanto alla responsabilità derivante dalla propria condotta di guida, ma anche a quella connessa alla condotta difettosa (e a tratti imprevedibile) dell'auto dotata di IA. Egli, infatti, potrà essere considerato non solo come conducente ma anche come destinatario di una posizione di controllo nei confronti del veicolo in grado di realizzare una "circolazione intelligente"[82].

Finché l'auto sarà dotata dei classici comandi manuali il responsabile di un eventuale sinistro occorso resterà verosimilmente il conducente (potendo quest'ultimo essere ancora considerato tale). Con particolare riferimento al livello di automazione 3 il conducente assume, infatti, un ruolo "potenziale", essendo costui «ancora investito di una *posizione di garanzia* rispetto agli altri utenti della strada, nonché rispetto ai propri passeggeri. Più in particolare, l'*obbligo di controllo* – di quella fonte di pericolo che è la *semi-autonomous car* stessa – verrebbe certamente rinvenuto nelle clausole generali iperelastiche di cui agli artt. 140 e 141 del Codice della strada (…) difficilmente il legislatore potrebbe costruire un'area di non punibilità per i "conducenti potenziali" di tali mezzi, configurando standard di cautela inferiori»[83].

La questione si fa ancor più problematica quando viene in considerazione il livello di automazione 4, il quale non richiede più che il conducente (ammesso che sia ancora possibile definirlo tale) si metta alla guida. Confrontando gli standard SAE J3016[84] sembra che in questo livello di automazione possano collocarsi anche veicoli senza pedali e, addirittura, senza conducente. Pare opportuno, con riferimento a questa tipologia di veicoli, iniziare a riflettere sulla possibile delimitazione di un parziale spazio di non punibilità per il conducente il quale, in tal sede, riveste un ruolo marginale ed eccezionale[85].

Insomma, la figura del conducente come "*human in command*" inizia a vacillare. Se da un lato vi è chi ritiene che, finché ci si confronterà con veicoli dotati di un livello di automazione pari a 3 o 4, il guidatore avrà sempre un ruolo di comando, essendo a quest'ultimo richiesto di restare vigile e di riprendere, ove necessario, il controllo del mezzo[86], dall'altro lato v'è chi afferma che «se al guidatore non è richiesto di monitorare il traffico sino a una richiesta di riassunzione della funzione di guida, lo stesso non può più essere ritenuto in concreto "*human in command*" né dunque (penalmente) responsabile di eventuali causazioni lesive occorse sino a quel momento, ove appunto il controllo sulla attività rischiosa era delegato alla macchina *AI driven* legalmente autorizzata»[87].

Ad un incremento del livello di automazione di tali veicoli non potrà che conseguire una sovrapposizione dei modelli imputativi commissivi ed omissivi: finché in capo al conducente rimarrà il totale controllo del veicolo, la violazione

---

per controllare l'innaturale ed improvvisa correzione della "sua" manovra: se il sistema interviene, un conducente disattento può commettere errori di guida, frenando improvvisamente» p. 102.

[82] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles,* cit., p. 1707.

[83] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., pp. 334-335.

[84] SAE J3016, *Levels of driving automation*, cit.

[85] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 335.

[86] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles,* cit., p. 1705.

[87] V. MANES, *L'oracolo algoritmico,* cit., p. 550 nota 13.

di una regola cautelare stradale da questi compiuta rientrerà in un contesto commissivo; nel momento in cui, invece, al conducente verrà chiesto di limitarsi a supervisionare il corretto funzionamento della macchina, il mancato adempimento di tale compito integrerà una condotta di tipo omissivo[88]. Ciò invero era già stato accennato in alcune precedenti battute, in quanto nella misura in cui al conducente non venga più richiesto di "condurre" il veicolo bensì, soltanto, di "controllarne" l'agire e di "impedire" i relativi danni a terzi, quest'ultimo verrebbe a rivestire una posizione di garanzia in termini di posizione di controllo su una fonte di pericolo[89].

Chiaro è che, per poter effettivamente parlare di una posizione di garanzia, è necessario che il garante abbia concretamente la possibilità materiale di agire e di compiere la condotta che avrebbe avuto l'obbligo giuridico di realizzare. La dottrina classica esclude che possa essere mosso un rimprovero al garante nell'eventualità in cui quest'ultimo risulti impossibilitato per le sue condizioni psico-fisiche o nel caso manchino le condizioni indispensabili per poter porre in essere l'azione doverosa[90]. Volendo attualizzare tale principio generale con degli esempi inerenti la materia in esame, potremmo affermare che non sarà possibile muovere un rimprovero al conducente di un veicolo dotato di automazione a livello 3 che non riprenda il controllo manuale dell'auto dietro richiesta di quest'ultima in quanto colto da un malore improvviso; del pari non sarà possibile contestare al "conducente" di un veicolo dotato di automazione livello 4 di non aver sterzato per impedire l'impatto col pedone, se l'auto è sprovvista di volante. Ancora, come anticipato pocanzi, abbiamo detto che è necessario che il "conducente" abbia la possibilità di reagire alla richiesta del sistema di riprendere i comandi. È stato stimato che per compiere tale operazione l'uomo necessiti di almeno 6 secondi di tempo che, in un contesto emergenziale su strada, possono essere troppi[91]. In assenza del tempo necessario a riprendere il controllo non sembra, dunque, che al conducente possa essere mosso alcun rimprovero per il sinistro occorso in quanto la condotta alternativa lecita (ossia il recupero del controllo) non sarebbe valsa ad evitare l'evento (in quanto, magari, la richiesta di riprendere i comandi da parte della vettura è sopraggiunta troppo tardi).

Sembra che, in questi termini, i tradizionali strumenti penalistici siano ancora in grado di mostrare la loro efficacia e di adeguarsi all'evoluzione scientifica del

---

[88] A. CAPPELLINI, *Profili penalistici delle self-driving cars*, cit., pp. 335-336. In verità la sussistenza di questa alternativa di imputazione tra condotta commissiva e omissiva non è nuova alla materia della circolazione stradale: «Altri casi di posizione di controllo su fonti di pericolo connesse a cose, trovano esemplificazione nella circolazione stradale. Se nei casi in cui l'uso di un autoveicolo conduce ad un evento lesivo, la responsabilità normalmente si incardina su un'*azione positiva*, tuttavia si danno ipotesi nelle quali il rimprovero penale assume ad oggetto una *condotta omissiva*. Si pensi, ad es., alla posizione di garanzia rivestita dal proprietario, anche nel caso in cui il veicolo sia condotto da altri, rispetto allo stato di manutenzione del mezzo: qui l'obbligo di garanzia a carico del proprietario (peraltro espressamente sancito dall'art. 2054 comma 4 c.c.) – la cui trasgressione può comportare una responsabilità penale per omesso impedimento dell'evento – trova giustificazione proprio nella circostanza che il veicolo, come fonte potenziale di eventi dannosi a carico di terzi, rientra nella sua sfera di appartenenza e, dunque, di controllo» G. FIANDACA, *Il reato commissivo mediante omissione*, cit., p. 192, corsivo nostro.

[89] C. PIERGALLINI, *Intelligenza artificiale*, cit., p. 1753.

[90] G. FIANDACA, E. MUSCO, *Diritto Penale*, cit., p. 629.

[91] S. BECK, *Robotics and Criminal Law. Negligence, Diffusion of Liability and Electronic Personhood*, in E. HILGENDORF, J. FELDLE, *Digitization and the Law*, Baden, 2018, p. 43.

settore in esame. Qualche novità potrebbe invece rinvenirsi in materia colposa. Restano fermi i classici margini di responsabilità colposa del conducente nel caso in cui quest'ultimo, avendo la possibilità e il dovere di intervenire, resti inattivo[92]: «l'evento mortale andrà imputato al conducente se si dimostra che con la dovuta diligenza avrebbe potuto (e dovuto) attivarsi per prendere il controllo dell'auto ed evitare l'investimento»[93].

Elementi di novità potrebbero sorgere con riferimento alla perimetrazione delle regole cautelari. Vigilare sull'operato di una macchina in grado di guidare in quasi totale autonomia costituisce di certo condotta non tipizzata da alcuna regola cautelare ma che è verosimilmente destinata a trovare una sua collocazione in ragione della specificità tecnica di questi veicoli. Ricorrere ai crismi della colpa generica potrebbe anche andar bene in un primo momento, ma non ci sembra possa costituire soluzione di lunga durata. Si renderà probabilmente necessario introdurre nuove regole cautelari: basti pensare all'opportunità di dar corpo, nero su bianco, alla regola di riprendere il controllo del veicolo in caso di richiesta di quest'ultimo o in caso di emergenza, statuendo un espresso dovere di monitoraggio.

Al ricorrere di nuove regole seguiranno verosimilmente nuovi problemi giuridici. Potremmo infatti chiederci come si comporteranno i limiti obiettivi al dovere di diligenza. Ad esempio, in quale momento il conducente sarà tenuto a riprendere il controllo del veicolo? Ancora, quale forma potrebbe assumere il principio di affidamento[94]? Nonostante nella materia della circolazione stradale si sia sempre registrata una certa ritrosia a lasciare margini di operatività al principio di affidamento – stante la tendenza a pretendere dal conducente che quest'ultimo preveda qualsivoglia possibile manovra inconsulta o evento lesivo[95] – adesso sarebbe possibile ripensare questo principio, riferendoci all'affidamento riposto dall'uomo nei confronti del buon funzionamento della macchina e chiedendoci quale incidenza potrebbe avere tale affidamento su un eventuale giudizio di responsabilità penale a carico del conducente.

Non saranno, poi, insolite ipotesi in cui la colpa possa essere equamente ripartita tra l'agire (o l'omettere) umano e il malfunzionamento del veicolo autonomo. Ci troveremo davanti ad ipotesi inedite di *culpa in interagendo* quando si dovesse dimostrare che l'evento lesivo sia stato causato in cooperazione tra uomo e auto autonoma[96]. Ad esempio, potranno esservi ipotesi di condotte "iniziate" dalla macchina e "concluse" dall'uomo. Anche per tale ragione appare di fondamentale importanza chiarire la ripartizione dei compiti tra conducente e sistemi di guida autonomi[97]. Al fine di realizzare un ottimale riparto della responsabilità sarebbe quantomeno opportuno comprendere chiaramente quali

[92] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 8.
[93] I. SALVADORI, *Agenti artificiali*, cit., p. 102.
[94] Si chiedono in che misura il conducente possa fare affidamento sull'auto a guida autonoma e in quali circostanze egli debba riprendere il controllo del veicolo S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., p. 435.
[95] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 336.
[96] V. MANES, *L'oracolo algoritmico,* cit., p. 549. «One should be aware that even when there is a human in the loop one cannot speak of a human decision anymore, but a decision made by human and machine collaboratively» S. BECK, *Robotics and Criminal Law,* p. 43.
[97] G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* pp. 104-105.

momenti della circolazione stradale siano da ricondurre all'uomo e quali al veicolo intelligente.

Stante la difficoltà di operare un simile riparto, si potrebbe ritenere che, «una volta certificata l'affidabilità dell'*autonomous vehicle*, autorizzata l'immissione dello stesso sul mercato e consentita quindi la circolazione, l'eventuale causazione di eventi dannosi o pericolosi sia da ascriversi all'area del "rischio consentito" (*Socially Accepted Risk*) ovvero sia da ascriversi (al produttore, come accennato, o) a chi concretamente si è assunto quel rischio, mettendosi alla guida»[98]. Non crediamo che i tempi siano ancora maturi per far rientrare i sinistri stradali causati dalle auto a guida autonoma nell'alveo del rischio consentito, proprio in quanto i contorni di quest'ultimo non hanno ancora preso forma. Riteniamo piuttosto di condividere l'opinione di quella parte di dottrina ad avviso della quale «appare dunque molto probabile che si opterà per mantenere a ogni costo la (s)confortante figura "parafulmine" del "conducente", capace di attirare su di sé le colpe per qualunque evento occorso, indipendentemente dalla sussistenza di un reale e ragionevolmente esigibile potere di controllo degli eventi»[99], così però rischiando di dar vita (ancora una volta) ad inammissibili forme di responsabilità oggettiva.

2.3. Le driverless car *e i limiti del diritto penale.*

Le auto completamente autonome (per intenderci, quelle rientranti nel livello di automazione 5) spingono il diritto penale tradizionale fino ai propri limiti[100]. Mancando i comandi sul veicolo verrà meno ogni possibile ruolo attivo dell'uomo a bordo della macchina[101]. Con il crescere dell'automazione si registrerà una sempre più massiccia cessione di sovranità dal conducente umano a quello artificiale il quale, come abbiamo visto, non può essere direttamente responsabilizzato. Ciò condurrà, quasi inevitabilmente, ad uno spostamento della responsabilità dall'uomo dietro il volante (non essendo forse più possibile considerarlo un guidatore) al costruttore del veicolo[102]. Anche in tal sede viene chiamata in causa una compagine di soggetti (es. programmatori e costruttori del veicolo) che, tradizionalmente, sarebbero tenuti fuori dall'imputazione penale per un sinistro stradale.

Occorre però chiedersi a quali condizioni tali soggetti possano essere ritenuti penalmente responsabili per il danno causato da un veicolo autonomo: «ad esempio, l'ingegnere che aiuta a sviluppare un'auto a guida autonoma dovrebbe incorrere in una responsabilità penale per omicidio se l'auto "prende una decisione" che causa la morte di un passante?»[103]. In un contesto come quello in

---

[98] V. MANES, *L'oracolo algoritmico,* cit., p. 551.

[99] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 337. Sull'opportunità di estendere la responsabilità penale dell'utente considerandolo responsabile per ogni azione delegata alla macchina (fatti salvi gli errori di costruzione) S. BECK, *Intelligent agents and criminal law. Negligence, diffusion of liability and electronic personhood*, in *Robotics and Autonomous Systems*, 2016, p. 141.

[100] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., p. 435.

[101] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 338. L'A. afferma in altro scritto che «nel caso in cui, invece, l'automazione del mezzo sia completa, al punto da essere questo addirittura privo di comandi, l'impossibilità oggettiva di interventi umani di emergenza impedisce di contestare penalmente una qualche colpa dell'utilizzatore, ridotto qui a mero passeggero» ID., Machina delinquere non potest?, cit., p. 8

[102] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles,* cit., p. 1705.

[103] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., p. 425.

esame, in cui dunque il conducente è ridotto a mero "passeggero", gli eventi lesivi occorsi saranno verosimilmente da ricondurre ad un malfunzionamento del sistema intelligente, chiamando conseguentemente in causa (anche qui) la responsabilità per danno da prodotto[104]. Come invero già accennato, si tratta di contesti produttivi complessi, i quali annoverano tra le proprie fila professionalità differenti: anche ove dovesse essere individuabile con certezza la tipologia di errore di programmazione del veicolo che ha determinato l'incidente, la difficoltà di individuare il diretto responsabile e i tradizionali meccanismi di imputazione della responsabilità penale per danno da prodotto comporterebbero la diretta attribuzione del fatto in capo ai vertici aziendali[105].

Ferme restando le considerazioni svolte nelle pagine precedenti[106], ci sembra di poter aggiungere in tal sede che la responsabilità di cui al d.lgs. 231/2001 mal si attagli alla responsabilità colposa per circolazione stradale: occorrerebbe (anche qui) annoverare tra i reati presupposto della responsabilità delle persone giuridiche i delitti di cui agli artt. 589 *bis* c.p. e 590 *bis* c.p., ferma restando la necessità di ricondurre l'evento lesivo ad una pecca organizzativa dell'ente e alla necessità che il reato sia stato commesso nel suo interesse. In altre parole, a meno di non voler ripensare radicalmente la disciplina in materia, non ci sembra che i tempi siano ancora maturi per un'imputazione di tal fatta.

Occorre inoltre evidenziare che, nel nostro ordinamento, «il rispetto degli *standard* per l'omologazione non basta ad escludere la ipotizzabilità di un difetto del prodotto»[107]. I criteri stabiliti per ottenere l'omologazione di un prodotto, infatti, sono criteri "statici", concernenti la sicurezza *ex ante* del bene in questione. La difettosità (e la relativa pericolosità) di un prodotto possono, infatti, sorgere *ex post* – anche in considerazione del rischio da sviluppo – specie se ci confrontiamo con sistemi di guida autonoma dotati di un'intelligenza artificiale che ingloba algoritmi di *machine learning* che le consentono di imparare dall'esperienza in modo non prevedibile dal suo programmatore e dal suo utilizzatore. In altri termini, «la "sicurezza" garantita dagli *standard* non elide la pericolosità di una particolare attività produttiva e del conseguente prodotto, così come non esclude che un prodotto omologato come "sicuro" si possa poi rivelare difettoso e dannoso. (…) Così, un veicolo potrebbe essere conforme agli *standard* minimi di omologazione e poi rivelarsi suscettibile di hackeraggio»[108]. La questione si fa quanto più complessa considerando che tale innovativa tipologia di produzione è in grado di corresponsabilizzare il produttore di una componente immateriale del veicolo che, purtuttavia, può avere un ruolo determinante nella causazione del sinistro. Ci stiamo riferendo all'autore dell'algoritmo, il quale sarà verosimilmente destinatario di una estensione della disciplina sulla responsabilità per danno da prodotto difettoso[109].

Il vero problema, tuttavia, si pone «nei casi in cui il sinistro *non* si sia verificato a causa di un errore certo e riconoscibile nella programmazione: casi,

---

[104] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1753.
[105] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 339.
[106] Cap. II, Sez. II, Par. 12.
[107] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles,* cit., p. 1710.
[108] U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles,* cit., p. 1709.
[109] U. RUFFOLO, *Self-driving car, auto driverless e responsabilità,* cit., p. 48. Sul concetto di "intelligenza esterna" v. G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture,* p. 125.

dunque, tali da non "reggere" l'addebito dell'evento lesivo, financo in base a quei canoni giurisprudenziali che giudicano sufficienti dei profili di tipicità oggettiva e soggettiva "affievoliti" per imputare il danno da prodotto»[110], acuendo la logica ascrittiva di tale specifica forma di responsabilità penale. La mancata individuazione del fattore scatenante il sinistro, unita agli oscuri meccanismi di *black box* che connotano tali tecnologie, rischiano di rendere sostanzialmente impossibile ricostruire il nesso causale tra condotta ed evento, in assenza del quale non sarà possibile muovere alcun rimprovero penale[111].

Proprio in considerazione dell'impossibilità per l'operatore dietro il veicolo dotato di IA di azzerare i possibili rischi derivanti dalla guida autonoma si riprospetta una duplice alternativa a noi ormai nota (ed invero un po' provocatoria): da un lato si potrebbe esentare il produttore da qualsivoglia responsabilità, specie sulla scorta del fatto che la macchina è in grado di agire in autonomia; di contro si potrebbe affermare il dovere in capo al produttore di prevedere ogni possibile evento lesivo di cui il veicolo intelligente potrebbe rendersi "autore", considerandolo dunque sempre responsabile[112].

Nessuna delle due soluzioni appare astrattamente percorribile. Mandare sempre esente da responsabilità il produttore lascerebbe la vittima del sinistro sprovvista di adeguata tutela penale; per contro, considerare il produttore sempre responsabile porterebbe alla costruzione di una specifica forma di responsabilità per posizione. Occorre però cercare di ragionare a livello pratico[113]. Come anticipato, per delimitare lo standard di diligenza del produttore si potrebbe valutare di delineare un certo margine di tolleranza per alcuni errori commessi dall'operatore in fase progettazione o di programmazione di tali auto[114]. Ad esempio, si potrebbe ridurre il dovere di diligenza all'impiego delle migliori conoscenze e tecnologie disponibili al momento della produzione della macchina, fermo restando il dovere di ritirare tempestivamente il veicolo nel caso in cui quest'ultimo, una volta immesso sul mercato, palesi la sua difettosità[115]. La misura della diligenza richiesta al produttore dipenderà, inoltre, dal livello di autonomia di cui è dotata la macchina. Confrontandoci con le auto completamente autonome ci troveremo, inevitabilmente, davanti a residue ipotesi «*in cui il danno rimane sostanzialmente (o presuntivamente) riconducibile alla macchina stessa*: casi destinati a rimanere penalmente irrilevanti»[116]. Ritorna, a ben vedere, la già affrontata questione del "danno anonimo"[117]. Tale *responsibility gap* finirebbe per

[110] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 339.
[111] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., p. 431. Gli AA. considerano più avanti l'agire del sistema intelligente come fattore interruttivo del nesso causale, p. 342.
[112] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., p. 427.
[113] Sul permanere della "responsabilità da cosa" in capo al detentore, al possessore, al proprietario o al custode di un'auto totalmente autonoma, U. RUFFOLO, *Self-driving car, auto driverless e responsabilità,* cit., p. 50.
[114] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., p. 431.
[115] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., p. 434. Gli AA. si concentrano anche sul mancato ritiro dal mercato di veicoli intelligenti difettosi che abbiano già manifestato segni di difettosità e sulla relativa imputazione in capo al produttore degli eventi lesivi derivati dal mancato ritiro del prodotto.
[116] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 340.
[117] Cap. II, Sez. II, Par. 14.

gravare esclusivamente sulla vittima la quale, patendo un danno ingiusto, potrebbe al più aspirare ad un risarcimento in sede civile[118].

Occorrerebbe ripensare la materia della *product liability* tenendo conto non soltanto della necessità di non disincentivare (con forme di responsabilità oggettiva dei produttori) la realizzazione dei veicoli autonomi, ma anche dell'importanza di immettere sul mercato prodotti sicuri, tutelando le potenziali vittime della strada rispetto ad un rischio ancora non socialmente accettato. Il nodo davvero problematico di queste vicende risiede nella difficoltà di ricondurre il sinistro ad un agente umano: da un lato non potrà essere considerato responsabile il conducente, in quanto divenuto mero passeggero; dall'altro non potrà essere sempre considerato responsabile il produttore, specie nel caso in cui la macchina abbia agito in modo non prevedibile.

Come più volte anticipato, la materia in esame porta con sé problematiche politiche, prima ancora che giuridiche. A nostro avviso, solo una volta risolto a monte il problema di politica legislativa afferente all'accettazione dei veicoli autonomi da parte della società e alla relativa delimitazione dell'area del rischio consentito, sarà possibile iniziare a ragionare su un'ottimale allocazione del rischio[119]. La questione verterà dunque sul livello di "tollerabilità sociale"[120] che i consociati saranno disposti a sopportare, fino forse al raggiungimento del momento in cui le auto a guida autonoma dovessero essere talmente diffuse da considerare un incidente da esse causato come un raro malfunzionamento considerato (anche dalla vittima) come socialmente accettato, nonché compensato dai relativi vantaggi di tale tipologia di veicoli[121]. Fino ad allora, occorrerà rifuggire da un lato dalla creazione di zone franche di responsabilità per i produttori e, dall'altro, da pericolosi vuoti di tutela per le vittime[122], andando a ricercare un'equa via mediana.

---

[118] È stato proposto, anche qui, di gestire il problema del risarcimento del danno derivato dalla circolazione di veicoli autonomi tramite coperture assicurative obbligatorie o mediante la creazione di un apposito fondo patrimoniale per ristorare le vittime di tale tipologia di sinistri, U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles*, cit., p. 1706.

[119] V. MANES, *L'oracolo algoritmico,* cit., p. 551.

[120] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1750.

[121] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., pp. 434-436.

[122] Ad avviso di parte della dottrina esonerare "l'uomo dietro la macchina" da qualsivoglia responsabilità non sarebbe soluzione ottimale. Intanto perché «il danno causato può infatti essere il risultato di una programmazione negligente e non di qualche imprevedibile stranezza dell'auto. In secondo luogo, fintanto che il robot stesso non può essere ritenuto penalmente responsabile, la vittima (e la società) possono trovarsi di fronte a un gap di responsabilità; sarebbe impossibile non ritenere penalmente responsabili né la macchina né le persone dietro di essa, anche per gravi danni causati dal robot. Un tale vuoto di responsabilità potrebbe causare un precipitoso calo del sostegno alle invenzioni robotiche. Queste considerazioni sconsigliano in generale l'assoluzione degli operatori del robot da responsabilità per danni causati dal robot stesso» S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?,* cit., p. 432. Gli AA. in altra parte del lavoro chiariscono inoltre che il fondamento di questa «responsabilità penale da prodotto non è la creazione illecita di un rischio ma il mero fatto che il produttore, nel perseguimento di interessi economici, crea legalmente un rischio per il grande pubblico rilasciando un Agente Intelligente le cui reazioni non possono essere previste con sicurezza e controllate. L'unicità di questo concetto di responsabilità penale è il fatto che un atto perfettamente legale - la commercializzazione di un'auto a guida autonoma conforme allo stato attuale delle conoscenze e delle tecnologie - può innescare la responsabilità penale per omissione. Potrebbe essere difficile per l'industria automobilistica accettare questo ampio ambito di responsabilità. Ma le vittime di incidenti causati

2.4. L'etica delle auto autonome: il dilemma del carrello ferroviario.

Le questioni etico-sociali afferenti alla materia delle auto a guida autonoma sono, invero, quelle che hanno maggiormente destato l'attenzione degli studiosi[123]. Torna in tal sede il *machine ethics*, ossia quella disciplina che studia come far sì che i sistemi autonomi intelligenti si comportino in modo eticamente responsabile[124]. Ci si dovrebbe aspettare, in altri termini, che il conducente "artificiale" prenda le stesse decisioni che prenderebbe il conducente umano[125]. I più evoluti sistemi intelligenti (ci riferiamo in particolare alle auto a guida completamente autonoma) possono assumere "decisioni" valutando diverse opzioni, ma occorrerà probabilmente insegnare loro come reagire, ed eventualmente risolvere, un dilemma morale, scegliendo magari di desistere dal proprio obiettivo ove esso comporti un prezzo troppo alto: ad esempio, sebbene un'auto a guida autonoma sia programmata per raggiungere la sua destinazione velocemente e senza deviazioni, può (e deve) essere programmata in modo tale che non passi sopra i pedoni che bloccano la strada[126].

Il problema non è solo teorico. Il Massachusetts Institute of Technology (M.I.T.) di Boston ha dato vita al progetto "*Moral Machine*"[127], «destinato a orientare i programmi di intelligenza artificiale nella percezione e soluzione di dilemmi etici. (…) La *Moral machine*, consultabile interattivamente, pone numerosi tipi di problemi etici e si propone di raccogliere, a livello planetario, l'opinione di milioni di persone in merito alle scelte moralmente preferibili, così da poter istruire di conseguenza i programmi di intelligenza artificiale»[128]. In altre parole, stante l'assenza di principi etici universalmente condivisi[129] si cerca, mediante tale esperimento, di creare un «modello computazionale della scelta

---

da auto a guida autonoma malfunzionanti troverebbero altrettanto difficile accettare una situazione in cui, in assenza di conducente, nessuno sia ritenuto responsabile dei danni causati» p. 428.

[123] U. RUFFOLO, *Self-driving car, auto driverless e responsabilità,* cit., p. 36.

[124] H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law,* cit., p. 343.

[125] U. RUFFOLO, *Self-driving car, auto driverless e responsabilità,* cit., p. 41.

[126] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., p. 422.

[127] La Moral Machine è consultabile al sito https://www.moralmachine.net. Sul tema v. E. AWAD, S. DSOUZA, R. KIM, J. SCHULZ, J. HENRICH, A. SHARIFF, J.F. BONNEFON, I. RAHWAN, *The Moral Machine experiment,* in *Nature,* 2018, vol. 563, pp. 59 ss.; J.D. GREENE, *Our driverless dilemma. When should your car be willing to kill you?,* in *Science,* 24.6.2016, vol. 352, pp. 1514 ss.; J.F. BONNEFON, A. SHARIFF, I. RAHWAN, *The social dilemma of autonomous vehicles,* in *Science,* vol. 352, 24.6.2016, pp. 1573 ss.

[128] M. PAPA, *Future crimes,* cit., pp. 11-12.

[129] Ciò è stato confermato da uno studio condotto sui dati registrati dal Moral Machine relativamente alle risposte fornite dagli utenti di 130 Paesi differenti e dal quale è emersa una sorta di divisione in tre distinti "gruppi morali": un gruppo Occidentale, un gruppo Orientale e un gruppo Meridionale. «Ad esempio, la preferenza nel risparmiare le persone più giovani piuttosto che quelle più anziane è molto meno pronunciata nei Paesi del cluster Orientale e molto più alta per i Paesi del cluster Meridionale. Lo stesso vale per la preferenza nel salvare persone dotate di uno status superiore. Allo stesso modo, i Paesi del cluster Meridionale mostrano una preferenza molto più debole nel risparmiare gli esseri umani rispetto agli animali domestici, a differenza degli altri due cluster. Solo la (debole) preferenza nel risparmiare i pedoni rispetto ai passeggeri e la (moderata) preferenza nel risparmiare chi rispetta la legge a fronte di chi non lo fa sembrano essere condivise nella stessa misura in tutti i cluster. Infine, osserviamo alcune particolarità sorprendenti, come la forte preferenza nel risparmiare le donne e la forte preferenza nel risparmiare le persone in forma nel cluster meridionale» E. AWAD, S. DSOUZA, R. KIM, J. SCHULZ, J. HENRICH, A. SHARIFF, J.F. BONNEFON, I. RAHWAN, *The Moral Machine experiment,* cit., p. 61.

sociale»[130]. L'analisi dei risultati raccolti dal *Moral Machine* ha mostrato l'esistenza di alcuni principi etici largamente condivisi e che potrebbero fungere da elementi costitutivi per fondare una discussione universale sull'"etica delle macchine", ossia la preferenza nel risparmiare le vite umane (rispetto agli animali), la preferenza nel risparmiare più vite (a fronte di singoli sacrifici) e la preferenza nel risparmiare le giovani vite (rispetto agli anziani)[131].

L'alternativa della collisione inevitabile[132] di fronte alla quale viene posto il veicolo autonomo è nota a molti come il "dilemma del carrello"[133]: trattasi di un'*impasse* morale in cui ogni scelta porta a un sacrificio, insomma, una "*lose-lose situation*"[134]. Ove gli interessi in contrasto dovessero essere gerarchicamente disomogenei, *nulla quaestio*, verrebbe certamente sacrificato il bene di minor valore (ammesso che possa esserci consenso universale sul c.d. "male minore")[135]. Il vero problema sorgerebbe nella misura in cui si trovassero a collidere interessi gerarchicamente disomogenei. La dottrina che si è occupata dal tema ha ritenuto non percorribili né la via del criterio quantitativo – che predilige, nell'ottica della minimizzazione del danno[136], la vita dei molti a danno dei pochi, stante il valore unico della vita di ognuno – né la via *randomica*, che cioè rimetta la scelta della macchina alla pura casualità[137]. Il problema è già stato affrontato nell'ordinamento tedesco il quale, con l'emanazione del Rapporto della Commissione etica su "*Automated and connected driving*" ha vietato qualsiasi distinzione basata sulle caratteristiche personali (ad es. età, sesso, costituzione fisica o psichica), giustificando per contro una programmazione generale volta a ridurre il numero di lesioni personali[138].

Agli occhi del penalista tali riflessioni possono evocare una rinnovata lettura dello stato di necessità: ove non sia possibile evitare un danno a una persona, il veicolo sacrificherà il passante o il passeggero? La scelta egoistica del conducente di salvare sé stesso, a ben vedere, parrebbe lecita proprio alla luce dello stato di

---

[130] D. TAFANI, *Sulla moralità artificiale,* cit., p. 87.

[131] E. AWAD, S. DSOUZA, R. KIM, J. SCHULZ, J. HENRICH, A. SHARIFF, J.F. BONNEFON, I. RAHWAN, *The Moral Machine experiment,* cit., p. 63. Le preferenze rivelate dalla Moral Machine risultano strettamente correlate con le caratteristiche economiche e culturali di questi Paesi e restituiscono una spiegazione delle risposte fornite dagli utenti. Ad esempio, i membri di culture tipicamente individualistiche preferiscono risparmiare un maggior numero di persone a fronte del sacrificio del singolo. Inoltre i membri delle culture collettivistiche, particolarmente sensibili al rispetto dovuto ai membri più anziani della comunità, mostrano una preferenza meno accentuata per salvare le vite dei giovani. Da qui, non a caso, la preferenza accordata dai paesi del cluster Orientale nei confronti della scelta di risparmiare gli anziani a danni dei più giovani. Ancora, i cittadini dei paesi più poveri, con istituzioni più deboli, sono più tolleranti nei confronti dei pedoni che attraversano illegalmente, rispetto ai cittadini di Paesi con istituzioni più solide, p. 62.

[132] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., p. 45.

[133] D. EDMONDS, *Uccideresti l'uomo grasso? Il dilemma etico del male minore*, Milano, 2014.

[134] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 12.

[135] G. FORNASARI, *Dilemma etico del male minore e* ticking bomb *scenario riflessioni penalistiche (e non) sulle strategie di legittimazione della tortura,* Napoli, 2020, pp. 17 ss.

[136] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., p. 45.

[137] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 14.

[138] FEDERAL MINISTRY OF TRANSPORT AND DIGITAL INFRASTRUCTURE, *Ethics Commission. Automated and connected driving*, 2017, Punto 9.

necessità cogente in cui si verrebbe a trovare[139]. Già da queste sintetiche battute pare intuitivo come queste rivisitazioni del *trolley problem* non siano mere congetture, assumendo piuttosto un ruolo determinante in ordine alla percezione della pericolosità delle auto autonome: «la prospettiva di una standardizzazione legale di questi "algoritmi di necessità" – magari ancorandola, com'è stato proposto, al criterio etico consequenzialista del minimo danno – inevitabilmente esporrebbe gli utenti al tanto paventato rischio di essere, in situazioni eccezionali, "sacrificati" dalla vettura»[140]. Tale opzione comporterebbe l'accoglimento di un approccio *paternalistico*, ossia volto a risolvere a monte i dilemmi morali che possono venire in gioco, con una soluzione "imposta dall'alto"[141].

Da esso va distinto un approccio *liberale* (o a "etica differenziata")[142] in cui sia consentito ad ogni utente di programmare il proprio veicolo affinché, davanti ad un dilemma morale, decida secondo il volere del suo proprietario[143], come una sorta di "manopola etica"[144]. Parte della dottrina ha sostenuto che, ove un simile scenario fosse possibile, da un punto di vista strettamente giuridico, la condotta del veicolo sarebbe determinata da una scelta consapevole del suo proprietario, il quale dovrà essere considerato come il solo responsabile dei danni cagionati dalla macchina, ferma restando la possibilità per il conducente di invocare a sua difesa lo stato di necessità ove siano stati cagionati danni a terzi per salvare sé stesso[145].

Il dilemma del carrello ferroviario[146] è dunque destinato a trovare nuova linfa in quella che potrebbe essere definita come l'«etica degli incidenti stradali»[147] commessi dalle *self driving cars*. Il terreno si fa ancor più accidentato e, perciò, occorre restare coi piedi per terra: «un veicolo a guida autonoma non è un agente morale in senso proprio e dunque, nei casi in cui si trovi di fronte alla prospettiva di un incidente inevitabile, non è (…) un soggetto che si trovi di fronte a un

---

[139] D. TAFANI, *Sulla moralità artificiale,* cit., p. 92. Sulla questione della situazione di necessità v. anche A. CAPPELLINI, Machina delinquere non potest?, cit., p. 6 nota 11.

[140] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 333.

[141] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., p. 46.

[142] A. CAPPELLINI, *Profili penalistici delle self-driving cars,* cit., p. 333.

[143] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., p. 46.

[144] «The knob gives the passenger the option to select one of three settings: 1. Altruistic Mode: preference for third parties; 2. Impartial Mode: equal importance given to passenger(s) and third parties; 3. Egoistic Mode: preference for passenger(s)» G. CONTISSA, F. LAGIOIA, G. SARTOR, *The Ethical Knob: Ethically-Customisable Automated Vehicles and the Law*, in *Artificial Intelligence and Law*, 2017, p. 369.

[145] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., p. 47. «Bastano le norme ed i principi relativi allo "stato di necessità" (art. 54 c.p.) per mandare indenne da responsabilità il conducente che travolga pedoni disattenti (bambini o meno) improvvisamente sbucati sulla strada, se posto di fronte alla alternativa di precipitare in un burrone» U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles*, cit., p. 1708.

[146] U. RUFFOLO, *Self-driving car, auto driverless e responsabilità,* cit., p. 40; U. RUFFOLO, E. AL MUREDEN, *Autonomous vehicles*, cit., p. 1705; U. PAGALLO, *Intelligenza Artificiale e diritto,* cit., p. 618, «Si pensi a un autoveicolo autonomo lungo un percorso in cui si trovano quattro persone legate e incapaci di muoversi: tra queste ultime e la macchina c'è però un'altra strada in cui giace, a sua volta, un altro individuo nelle medesime condizioni. Cosa dovrebbe fare il nostro autoveicolo (che, poi, significa: come dovremmo programmarlo)? Dovrebbe deviare dal suo percorso, in chiave utilitaristica, al fine di uccidere una sola persona, al posto di quattro? E che dire, nel caso in cui le scelte da soppesare mettano o meno a loro volta a rischio la vita del passeggero?».

[147] D. TAFANI, *Sulla moralità artificiale,* cit., p. 86.

dilemma morale. È una macchina, che reagirà sulla base di istruzioni programmate e, in quanto dotata di intelligenza artificiale, sulla base degli esiti, solo in parte prevedibili, del suo apprendimento automatico»[148].

Nonostante si registri attualmente una certa attenzione per i cosiddetti "algoritmi morali", la reale priorità dovrebbe essere lo studio di "algoritmi legali", volti cioè a rendere il comportamento del veicolo autonomo conforme al codice della strada. Con particolare riferimento alla circolazione stradale, ad avviso della dottrina, non rischieremmo neanche di trovarci davanti ad una difformità tra norme giuridiche ed etiche in quanto, per i veicoli autonomi, i requisiti legali ed etici finirebbero spesso per sovrapporsi, fino a coincidere, poiché l'"etica del traffico" deriva dagli stessi valori promossi dal codice della strada, ossia una circolazione sicura ed efficiente[149].

Inoltre, anche ove si dovesse riuscire ad "insegnare" al veicolo autonomo come rispettare le regole del codice della strada, resterebbe da considerare che una sicura circolazione stradale si fonda anche su principi non scritti[150]: ad esempio, se l'auto autonoma è stata addestrata a riconoscere solo determinati segnali, come potrà riconoscere la condotta di chi suoni il clacson sventolando un fazzoletto bianco dal finestrino per palesare agli altri conducenti che ha urgenza di passare per recarsi in ospedale?

I nodi problematici da sciogliere sono ancora molti ma su tale ultimo punto pensiamo possa esserci largo consenso: più che parlare di dilemmi morali sarebbe il caso si pensare a "standard di sicurezza funzionali", che istruiscano il veicolo a gestire i rischi ed, eventualmente, ad arrestarsi in caso di impossibilità di risoluzione degli stessi. Non sarebbe forse il caso di pensare ad un "algoritmo sacrificale" che stabilisca una gerarchia di valori delle vite umane che (fortunatamente) non esiste. L'assenza di una scala di valori che consenta di scegliere chi sacrificare non costituisce una lacuna dell'ordinamento, quanto piuttosto una mancata risposta per una domanda che non dovrebbe essere posta[151].

3. L'intelligenza artificiale nel settore medico.

L'avvento dell'intelligenza artificiale nel settore medico può portare con sé una moltitudine di benefici[152]. Ciò è già stato avvertito, invero, dall'ormai nota

---

[148] D. TAFANI, *Sulla moralità artificiale,* cit., p. 84.

[149] H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law,* cit., pp. 359-360. Proprio al fine di "disegnare" un algoritmo in grado di rendere il comportamento del veicolo autonomo conforme al codice della strada l'A. propone il modello del *regimentation approach*, del *reasoning approach* e del *training approach*, cui si rimanda per un approfondimento, pp. 351-352.

[150] H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law,* cit., p. 360.

[151] D. TAFANI, *Sulla moralità artificiale,* cit., p. 93 e p. 89.

[152] A. SPINA, *La medicina degli algoritmi: Intelligenza Artificiale, medicina digitale e regolazione dei dati personali*, in F. PIZZETTI, *Intelligenza artificiale*, cit., p. 320. Pensiamo, a titolo meramente esemplificativo, ai robot chirurghi che eseguono i comandi del chirurgo, ai robot fisioterapici usati per la riabilitazione dei pazienti, ai nano robot usati in fase di diagnosi, alle protesi per potenziare il corpo umano, V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina*, in G. ALPA, *Diritto e intelligenza artificiale*, cit., p. 489. Oppure, ancora, pensiamo alle c.d. macchine "intelligenti", come i sintetizzatori vocali o le *Brain Computer Interfaces*, che consentono a malati coscienti e sprovvisti dell'uso della parola di poter comunicare con terzi, C. SALAZAR, *Umano, troppo umano,* cit., pp. 270-271 a proposito dei casi Nuvoli e Welby.

Risoluzione del Parlamento europeo recante norme di diritto civile sulla robotica, la quale esamina le potenzialità dei robot impiegati nel settore dell'assistenza, della chirurgia, della diagnostica e degli interventi migliorativi del corpo umano[153].

Le problematiche connesse all'ingresso dell'IA in medicina sono invero state affrontate anche a livello interno, basti pensare al parere del Comitato nazionale per la Bioetica (CNB) e del Comitato nazionale per la Biosicurezza, le Biotecnologie e le Scienze della vita (CNBBSV) intitolato "Intelligenza Artificiale e Medicina: aspetti etici". Quest'ultimo, pur salutando con favore i progressi e le opportunità dell'IA nel campo della medicina, si prefigge l'obiettivo di «identificare le condizioni etiche per uno sviluppo della IA che non rinunci ad alcuni aspetti della nostra umanità, in un nuovo "umanesimo digitale", *per una medicina "con" le macchine e non "delle" macchine*»[154]. Il parere affronta espressamente anche la questione dell'attribuzione della responsabilità per un eventuale danno al paziente, evidenziando l'importanza di un lavoro interdisciplinare tra scienze giuridiche e scienze mediche «che veda le due competenze "parlarsi fra loro", anche ai fini di delineare il futuro assetto delle possibili molteplici responsabilità mediche connesse con la IA»[155].

Lo sforzo unificatore del diritto in questo specifico ambito dipenderà altresì da un ulteriore fattore: mentre la scienza medica si caratterizza già di per sé come un settore "unico", a qualsiasi latitudine, il diritto invece si connota per un approccio differenziato a seconda dell'ordinamento che viene preso in considerazione. L'unitarietà del fenomeno richiede una soluzione giuridica che sia, a sua volta, quanto più unitaria possibile: per questo motivo il soggetto probabilmente più adeguato a svolgere tale compito di unificazione sarebbe il legislatore sovranazionale, sul quale ricadrebbe l'arduo compito di trovare una convergenza di vedute tra le logiche e le norme dei diversi sistemi giuridici[156].

Chiaro è che raggiugere tale punto d'incontro non sarà compito semplice, specie in considerazione del fatto che le diverse tipologie dei sistemi intelligenti impiegati in ambito medico sono dotate delle proprie peculiarità e delle rispettive ricadute dal punto di vista giuridico. Ad esempio, una preliminare distinzione problematica potrebbe essere quella che intercorre tra robot che agiscono in totale autonomia e robot che agiscono al fianco dell'uomo[157]: torna anche qui il *leitmotiv*

[153] *Risoluzione del Parlamento europeo,* cit., Punti 31 ss.

[154] COMITATO NAZIONALE PER LA BIOETICA (CNB), COMITATO NAZIONALE PER LA BIOSICUREZZA, LE BIOTECNOLOGIE E LE SCIENZE DELLA VITA (CNBBSV), *Intelligenza Artificiale e Medicina: aspetti etici*, 29.5.2020, p. 9, corsivo nostro. Il parere si concentra, in particolar modo, sul ruolo dell'IA nella relazione medico-paziente, sull'affidabilità e sull'opacità degli algoritmi intelligenti, sul rapporto tra IA e trattamento dei dati personali, sul rispetto del consenso informato e del principio di autodeterminazione, sul problema dell'attribuzione della responsabilità e sulla formazione del personale medico.

[155] COMITATO NAZIONALE PER LA BIOETICA (CNB), COMITATO NAZIONALE PER LA BIOSICUREZZA, LE BIOTECNOLOGIE E LE SCIENZE DELLA VITA (CNBBSV), *Intelligenza Artificiale e Medicina: aspetti etici*, cit., p. 15, corsivo nostro.

[156] U. RUFFOLO, *Le responsabilità da intelligenza artificiale nel settore medico e farmaceutico*, in U. RUFFOLO, *Intelligenza artificiale e responsabilità*, cit., p. 55.

[157] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 490.

della nostra indagine, ossia l'imputazione della responsabilità penale (stavolta in ambito medico)[158] in caso di danno cagionato da o mediante l'IA.

3.1. Il supporto intelligente in fase di diagnosi.

Uno dei settori in cui vengono maggiormente utilizzati i sistemi intelligenti è quello del supporto in fase di diagnosi e di strutturazione dei piani terapeutici[159]. Molti di tali sofisticati *medical devices* si servono degli ormai noti algoritmi di *machine learning*, i quali «lavorano su grandissime quantità di dati provenienti da pazienti, costruiscono i loro modelli, e dai primi esperimenti riescono ad avere accuratezze confrontabili e a volte superiori a quelle dei medici»[160].

Gli algoritmi di *machine learning* e di *deep learning* sono molto sfruttati in fase di diagnostica, in quanto il loro *modus operandi* è quello che maggiormente si avvicina alle capacità percettive e deduttive del nostro sistema nervoso[161]. Tale

---

[158] Posta l'ampiezza della letteratura sul tema della responsabilità medica, senza alcuna pretesa di esaustività, si rinvia a D. MICHELETTI, *La normatività della colpa medica nella giurisprudenza della Cassazione*, in S. CANESTRARI, F. GIUNTA, R. GUERRINI, T. PADOVANI (a cura di), *Medicina e diritto penale*, Pisa, 2009, pp. 247 ss.; M. BILANCETTI, *La responsabilità civile e penale del medico*, Padova, 2010; R. BLAIOTTA, *La responsabilità medica. Nuove prospettive per la colpa*, in *penalecontemporaneo.it*, 23.3.2012; G. DE FRANCESCO, *L'imputazione della responsabilità penale in campo medico-chirurgico: un breve sguardo d'insieme*, in *Riv. it. med. leg.*, 2012, pp. 969 ss.; O. DI GIOVINE, *La responsabilità penale del medico: dalle regole ai casi*, in *Riv. it. med. leg.*, 2013, pp. 78 ss.; L. RISICATO, *L'attività medica di équipe tra affidamento ed obblighi di controllo reciproco. L'obbligo di vigilare come regola cautelare*, Torino, 2013; A. MANNA, *Medicina difensiva e diritto penale*, Pisa 2014; L. RISICATO, *Il nuovo statuto penale della colpa medica: un discutibile progresso nella valutazione della responsabilità del personale sanitario*, in *La Legislazione Penale*, 5.6.2017.

[159] «L'intelligenza artificiale si sta notevolmente affermando, peraltro, anche con riguardo alla fase diagnostica, nel cui ambito s'è sperimentato con successo l'uso di micro-sistemi robotizzati. Si pensi, per esempio, all'utilizzo delle cosiddette *smart medical capsules*: ovvero, microcapsule endoscopiche, dotate di telecamere, utilizzate, in particolare, per la diagnosi delle patologie dell'apparato intestinale. Ma non basta. Sempre più incoraggianti risultano, altresì, i risultati ottenuti, sempre in ambito diagnostico, dalla *nano*-robotica: tanto che, in futuro, vi saranno dispositivi diagnostici in miniatura, con dimensioni inferiori a 100 nm, in grado di essere iniettati direttamente nel flusso sanguigno o di essere ingeriti, senza alcuna difficoltà, dal paziente» C. IAGNEMMA, *I 'robot medici': profili problematici in tema di alleanza terapeutica e di responsabilità penale*, in *Corti supreme e salute*, 2/2020, p. 5; «in una simulazione i cui risultati sono stati pubblicati nell'agosto del 2017, il sistema Watson dell'IBM ha diagnosticato e proposto una terapia per un malato di tumore cerebrale in dieci minuti, mentre un team interdisciplinare di umani ha impiegato centosessanta ore per svolgere lo stesso compito. I medici, però, sono giunti ad una combinazione di terapie più efficace rispetto a quella proposta da Watson; risultato che ha suggerito una collaborazione macchina-uomo, più che una sostituzione dell'elemento umano» C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 120; C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1918; G. MOBILIO, *L'intelligenza artificiale*, cit., p. 402; A. VESPIGNANI, *L'algoritmo e l'oracolo,* cit., p. 72. Sul tema V. HARISH, F. MORGADO, A. STERN, S. DAS, *Artificial Intelligence and Clinical Decision Making: The New Nature of Medical Uncertainty*, in *Academic Medicine*, 2020, pp. 31 ss.; N. SIMAAN, M.R. YASIN, L. WANG, *Medical Technologies and Challenges of Robot assisted Minimally Invasive Intervention and Diagnostics*, in *Ann. Rev. Control, Robotics, and Autonomous Systems*, 2018, fasc. 1, pp. 465 ss.; H. LIANG, B. TSUI, H. NI, et al., *Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence*, in *Nature Medicine*, 25, 11.2.2019, pp. 433 ss.

[160] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 4.

[161] A. PERIN, *Standardizzazione, automazione e responsabilità medica. Dalle recenti riforme alla definizione di un modello d'imputazione solidaristico e liberale*, in *BioLaw Journal*, 1/2019, p. 229.

tipologia di sistemi intelligenti è dotata di capacità di autoapprendimento e di autocorrezione che gli consentono di ridurre l'errore umano in fase diagnostico/terapeutica, di migliorarne l'accuratezza e di riuscire a produrre diagnosi precoci, così da evitare il declino di certe patologie. È possibile distinguere due tipi di tali *devices*: da un lato ci sono i sistemi che si servono del *machine learning* per analizzare i dati sanitari, raggruppare le caratteristiche di vari pazienti e dedurre i probabili esiti di determinate malattie; dall'altro abbiamo i sistemi di elaborazione del linguaggio naturale, i quali «elaborano dati non strutturati (ad es. note cliniche e riviste mediche) al fine di sviluppare dati medici strutturati»[162].

Come ormai abbiamo imparato a conoscere, il problema principale di questi strumenti intelligenti risiede nella mancata esplicabilità del processo logico seguito dal sistema per raggiungere un dato output, non consentendo così di effettuare una valutazione clinica sulla bontà del risultato[163]. Si tratta di una declinazione settoriale dell'ormai noto fenomeno del *black box,* inteso in questo contesto come *black box medicine*, il quale «mette in discussione gli elementi tradizionali del metodo scientifico e della medicina quali la trasparenza, la dimostrabilità, la generalizzabilità dei risultati»[164].

I classici problemi di trasparenza che connotano tali tecnologie generano "dubbi di credibilità" nei confronti dei sistemi di supporto alle decisioni cliniche, per questo motivo gli studiosi hanno individuato le caratteristiche che devono essere possedute da questi sistemi affinché il loro output possa essere accettato: «le scatole nere sono inaccettabili: un CDSS [*clinical decision support systems*] richiede trasparenza in modo che gli utenti possano comprendere la base di qualsiasi consiglio o raccomandazione che viene offerto; il tempo è una risorsa scarsa: un CDSS dovrebbe essere efficiente in termini di tempo richiesto e deve integrarsi nel flusso di lavoro del frenetico ambiente clinico; complessità e mancanza di utilizzabilità ne ostacolano l'uso: un CDSS dovrebbe essere intuitivo e semplice da apprendere e utilizzare in modo che non sia richiesta una formazione importante e risulti facile ottenere consigli o risultati analitici; rilevanza e intuizione sono essenziali: un CDSS dovrebbe riflettere una comprensione del dominio pertinente e del tipo di domande per le quali è probabile che i medici richiedano assistenza; la trasmissione di conoscenze e informazioni deve essere rispettosa: un CDSS dovrebbe offrire consigli in un modo che riconosca l'esperienza dell'utente, chiarendo che è progettato per informare e assistere ma non per sostituire un medico; le basi scientifiche devono essere solide: un CDSS dovrebbe avere prove scientifiche rigorose e sottoposte a revisione paritaria che ne stabiliscano la sicurezza, la validità, la riproducibilità, l'utilizzabilità e l'affidabilità»[165].

Oltre ai classici problemi di trasparenza (che connotano, invero, ogni sistema intelligente), nel settore in esame si pongono ulteriori problematiche. Preliminarmente è stato correttamente fatto osservare in dottrina che, se un

[162] F.C. LA VATTIATA, *Artificial Intelligence in Healthcare: Risk Assessment and Criminal Law*, in *Diritto Penale e Uomo*, 2.12.2020, pp. 6-7.
[163] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 494.
[164] A. SPINA, *La medicina degli algoritmi*, cit., p. 320.
[165] E.H. SHORTLIFFE, M.J. SEPÚLVEDA, *Clinical Decision Support in the Era of Artificial Intelligence*, in *JAMA*, 5.11.2018, p. 2199.

modello diagnostico dimostra la sua efficacia con riferimento ai dati di una certa popolazione, non è detto che fornisca eguali prestazioni con riferimento alla popolazione di un'area geografica differente[166].

Un secondo problema attiene alla necessità di una formazione "tecnica" da parte dei medici. L'interpretazione dei criteri su cui si basano gli output dei sistemi di *deep learning* costituisce attività complessa perfino per coloro che hanno dato vita all'algoritmo[167]: per tale ragione emerge ancor più chiara in questo frangente l'importanza di un lavoro sinergico tra operatori sanitari e informatici, al fine di raggiungere un proficuo intersecarsi di conoscenze.

Altro aspetto che meriterebbe di essere adeguatamente valorizzato concerne il ruolo dei dati[168]. Tali sistemi intelligenti dovrebbero essere sottoposti ad un addestramento costante sulla base di dati clinici quanto più possibile aggiornati[169], chiaramente forniti dal sanitario in carne ed ossa. Per quanto in realtà sia stato fatto oculatamente osservare in dottrina che, proprio in ragione del fatto che le informazioni somministrate al sistema intelligente provengono da una catalogazione umana, ciò impedirebbe di escludere un residuo margine di errore[170].

Tali considerazioni costituiscono probabilmente il sintomo dell'impossibilità di sostituire completamente il medico umano con quello "artificiale", non solo sulla scorta di un'inopportunità rilevata a livello istituzionale[171], ma anche in base a valutazioni tanto di tipo etico quanto giuridico. Ove la figura del medico dovesse essere "annullata", riducendo quest'ultimo a mero esecutore materiale della "decisione" assunta dalla macchina, verrebbe meno uno dei poli della relazione medico-paziente. Accogliere una simile prospettiva significherebbe rinunciare al momento della comunicazione tra medico e paziente, il quale è stato normativamente riconosciuto come "tempo di cura"[172]. A sostegno dell'impossibilità di una completa sostituzione del sanitario in carne ed ossa milita un altro, non secondario, fattore: «è stato osservato come molte attività della medicina diagnostica – quale, tipicamente, la lettura di una radiografia – non si realizza sulla base di ragionamenti necessariamente lineari, ma anche di carattere *ipotetico-intuitivo*»[173], realizzando un'attività "creativa" di cui l'IA non sarebbe

[166] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 5.
[167] A. PERIN, *Standardizzazione, automazione e responsabilità medica,* cit., p. 230.
[168] Sulla nuova fisionomia della materia della protezione dei dati personali nello sviluppo dei sistemi di IA, A. SPINA, *La medicina degli algoritmi,* cit., pp. 322 ss.
[169] F.C. LA VATTIATA, *Artificial Intelligence in Healthcare,* cit., p. 7.
[170] A. PERIN, *Standardizzazione, automazione e responsabilità medica,* cit., p. 230.
[171] Ciò è chiarito invero anche dalla *Risoluzione del Parlamento europeo,* cit., la quale, al Punto 33, afferma che «l'utilizzo delle tecnologie in questione *non* debba *sminuire o ledere il rapporto medico-paziente*, bensì fornire al medico un'*assistenza nella diagnosi* e/o nella cura del paziente allo scopo di ridurre il rischio di errore umano e di aumentare la qualità della vita e la speranza di vita», corsivo nostro. In tal senso si esprime COMITATO NAZIONALE PER LA BIOETICA (CNB), COMITATO NAZIONALE PER LA BIOSICUREZZA, LE BIOTECNOLOGIE E LE SCIENZE DELLA VITA (CNBBSV), *Intelligenza Artificiale e Medicina: aspetti etici,* cit., p. 10 il quale afferma che «L'automazione nell'acquisizione e interpretazione dei dati, nell'elaborazione delle diagnosi e nell'individuazione delle terapie o nella stessa effettuazione degli interventi non può essere completamente indipendente dall'uomo, ma richiede una costante verifica, pertanto non esclude la specificità della relazione tra medico e paziente».
[172] Art. 1, comma 8, della legge 22 dicembre 2017, n. 219, *Norme in materia di consenso informato e di disposizioni anticipate di trattamento*, G.U. n. 12 del 16.1.2018.
[173] A. PERIN, *Standardizzazione, automazione e responsabilità medica,* cit., p. 230, corsivo nostro.

capace. Ciò risulta tanto più vero quando ci si riferisce a patologie con riferimento alle quali non sussistono ancora evidenze scientifiche certe e, rispetto alle quali, la macchina non sarebbe in grado di pensare "fuori dagli schemi"[174].

Dunque, per quanto il medico possa nutrire una seria aspettativa nei confronti della bontà e della correttezza della soluzione proposta dall'IA, ci si chiede fino a che punto «sia possibile sostenere la riconoscibilità della necessità di disattenderne l'indicazione», rispondendo magari per colpa grave solo in caso di inequivoca riconoscibilità dell'errore o dell'inadeguatezza dell'indicazione. Al fine di cercare di individuare i limiti al suddetto principio di affidamento, in dottrina è stato ipotizzato di prendere in considerazione, ad esempio, l'uso del *device* intelligente al di fuori del proprio autorizzato ambito di utilizzo, nonché l'affidamento riposto in un output nonostante sussistessero dubbi sulla validità scientifica dei dati utilizzati per il *training* del sistema[175].

In altri termini, non sembra concretamente sostenibile nel settore *de quo* una sostituzione dell'operatore umano a tutto vantaggio di quello artificiale, dovendo quest'ultimo restare mero supporto nei confronti del medico in carne ed ossa, il quale continua a rimanere il principale responsabile per gli eventuali danni riportati dai pazienti. Proseguendo nello studio della materia in esame avremo modo di vedere come, in realtà, tale schema sembri ripetersi.

3.2. Chirurgia robotica e gradi di responsabilità.

L'attività chirurgica si è col tempo andata evolvendo, passando dal tradizionale coinvolgimento esclusivo della manualità umana, alla chirurgia mininvasiva a quella laparoscopica. Tutte queste ipotesi restavano accomunate dall'indubbia riconducibilità dell'eventuale evento lesivo al medico in carne ed ossa. L'avvento dei robot chirurghi[176] e la comprensione della loro modalità di

---

[174] C. IAGNEMMA, *I "robot medici"*, cit., pp. 11-12, la quale riporta proprio «quanto accaduto nella fase iniziale dell'epidemia da Sars-Cov2. Il primo caso di contagio in Italia è stato diagnosticato grazie alla perspicacia di un giovane medico, la dottoressa Malara, che ha deciso di 'forzare' il protocollo previsto in materia: senza che vi fosse alcun segnale evidente che suggerisse di discostarsi dalla prassi. Qualsiasi intelligenza artificiale non sarebbe riuscita a conseguire tale risultato, essendo incapace di "pensare – come, più volte, ha dichiarato la stessa dottoressa Malara – *all'impossibile*"». Sull'assenza di linee guida idonee a fronteggiare l'emergenza da coronavirus L. RISICATO, *La metamorfosi della colpa medica nell'era della pandemia*, in *disCrimen*, 25.5.2020, pp. 3 ss.

[175] A. PERIN, *Standardizzazione, automazione e responsabilità medica*, cit., p. 232.

[176] «Il primo utilizzo di una macchina robotica in chirurgia si è avuto all'inizio degli anni Ottanta, quando negli Stati Uniti è stato introdotto il sistema PUMA 560 (*Programmable Universal Manipulation Arm*), al fine di realizzare biopsie celebrarli di alta precisione. Da allora la tecnologia ha sperimentato, in materia, una rapida evoluzione, con l'introduzione di sistemi assai più sofisticati: si pensi, per esempio, al robot *ACRobot*, adoperato nella microchirurgia ortopedica, al robot *NeuroArm*, che agisce per limitare i tremori degli operatori negli interventi cerebrali, oppure alla macchina *ROBODOC*, utile nelle operazioni del femore e del ginocchio» C. IAGNEMMA, *I "robot medici"*, cit., p. 4; nella letteratura specialistica v. M. GARBEY, B. LEE BASS, S. BERCELI, C. COLLET, P. CERVERI, *Computational Surgery and Dual Training: Computing, Robotics and Imaging*, New York, 2013; D.A. HASHIMOTO et al., *Artificial Intelligence in Surgery: Promises and Perils*, in *Annals of Surgery*, 2018, pp. 70 ss.; sul tema v. anche E. MACRÌ, A. FURLANETTO, *I robot tra mito e realtà nell'interazione con le persone, negli ambienti sociali e negli ospedali. Un approccio tra* risk management *e diritto*, in *Riv. it. med. leg.*, 2017, fasc. 3, pp. 1045 ss.

funzionamento è, invece, in grado di incidere sui meccanismi di imputazione della responsabilità[177].

Sono stati individuati in dottrina 6 livelli di automazione della robotica chirurgica: i sistemi al livello 0 si limitano a rispondere al comando impartito dal medico; i robot chirurghi al livello 1 forniscono un'assistenza meccanica al chirurgo il quale mantiene il controllo del sistema; i sistemi al livello 2 eseguono in autonomia i compiti affidati dall'uomo, non essendo richiesto a quest'ultimo di esercitare un controllo continuo; i robot al livello 3 sono dotati di "autonomia condizionale", potendo formulare diverse strategie d'azione, pur essendo rimessa all'uomo la scelta finale sulla strategia da mettere in pratica; al livello 4 il sistema può assumere le sue decisioni, sempre sotto la supervisione del medico; il livello 5 si caratterizza per un'autonomia completa in grado di svolgere un intervento chirurgico integralmente da solo[178]. Lo sviluppo degli ultimi due livelli, in particolare, si colloca ancora in un futuro non troppo prossimo.

Tra i sistemi al livello 0 spicca il robot chirurgo Da Vinci, ossia «una macchina che garantisce al chirurgo una visione tridimensionale ad alta definizione unita a una strumentazione flessibile, avente sette gradi di movimento, che riproduce i movimenti di una mano umana, consentendo così la realizzazione di interventi complessi con estrema precisione»[179]. Il chirurgo umano mantiene sul suo "collega artificiale" un controllo significativo, non sollevando dunque particolari problematiche in punto di attribuzione della responsabilità in caso di eventi lesivi[180]. Rileva dunque in tal senso l'ordinaria diligenza che deve essere tenuta dal medico-chirurgo al fine di non incorrere in responsabilità connesse al malfunzionamento del sistema o al suo scorretto utilizzo: ad esempio la preliminare effettuazione di un *check* di sicurezza prima di ogni intervento, il costante monitoraggio della macchina e lo svolgimento di un periodo di *training* chirurgico[181].

Dal livello di automazione 1 in poi non si tratta più dei c.d. "dispositivi *slave*", in quanto essi sono in grado di "correggere" il chirurgo il quale, al fine di esercitare un potere di comando sul congegno, dovrà essere in condizione di "superare" il sistema riprendendo il controllo. Al livello 2 al medico non viene più richiesto di supervisionare costantemente il dispositivo intelligente, infatti in dottrina è stato a tal uopo proposto di «modulare opportunamente la vigilanza percettiva e cognitiva richiesta all'operatore affinché sia garantito un controllo umano significativo dei robot chirurgici che si collocano nell'ampia classe dei sistemi ad autonomia L2». Oltre ai problemi connessi al controllo esercitabile su questi sistemi, il livello 3 si caratterizza per un'ulteriore difficoltà, ossia

---

[177] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 491, per quanto l'A., in realtà, si riferisca alla responsabilità civile.

[178] G.Z. YANG et al., *Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy*, in *Science Robotics*, 2017, pp. 1-2. Nella dottrina giuridica v. anche D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., p. 47; C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1757.

[179] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 491. Sul tema v. anche C. IAGNEMMA, *I "robot medici"*, cit., p. 4, cui si rinvia per la letteratura specialistica.

[180] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., p. 48.

[181] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 493.

interpretare le decisioni proposte dall'IA, stante l'ormai nota opacità dei sistemi di *machine learning* e *deep learning* [182].

Gli studiosi hanno individuato una (ulteriore) nuova frontiera del settore nella "telechirurgia", ossia la possibilità per il medico di eseguire interventi chirurgici a distanza, evitando al paziente di spostarsi dal luogo di degenza[183] e, magari, anche di mettere in collegamento più medici consentendogli di collaborare all'operazione, ciascuno dal proprio terminale[184].

L'uso della robotica nel settore della chirurgia può certamente essere considerato foriero di benefici, basti pensare all'incremento della precisione delle incisioni determinato dall'eliminazione del tremore della mano[185] e alla possibilità per il medico di maneggiare una strumentazione più pratica e leggera che lo affatichi meno[186]. Occorre però evidenziare che spesso i costi per l'acquisto e la manutenzione di tali sofisticati robot risulta particolarmente proibitivo, senza tener conto delle spese per la formazione del personale sanitario deputato ad utilizzarli[187].

Tuttavia, anche in tale branca della medicina, un dato ci sembra allo stato incontrovertibile: il robot-chirurgo deve restare un ausilio, non potendo sostituire il medico umano[188]. Quest'ultimo dovrà continuare a rivestire una posizione di controllo significativa da cui ben potrebbero scaturire standard di diligenza *ad hoc* ed il cui mancato rispetto sarebbe idoneo a fondare la responsabilità penale del medico dietro il sistema intelligente. In dottrina è stato evidenziato come, in prospettiva *de iure condendo*, tali standard potrebbero costituire oggetto di specifiche linee guida[189], «il cui rispetto è suscettibile di escludere la punibilità del medico ai sensi dell'art. 590-sexies del codice penale e può venire in rilievo come elemento a favore del medico nei giudizi di responsabilità civile»[190].

Non solo, la materia delle linee guida viene qui in rilievo sotto un altro punto di vista. Se prima ci siamo confrontati con la possibilità per i sistemi intelligenti di "imparare" le regole del codice della strada, in tal sede si prospetta il problema di come far sì che essi imparino il disposto di linee guida che, invero, non costituiscono valore assoluto[191]. Le linee guida sono recentemente state definite dalla giurisprudenza come «raccomandazioni di ordine generale, che contengono

---

[182] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., pp. 48-49, per la citazione p. 49.
[183] C. IAGNEMMA, *I "robot medici"*, cit., p. 3.
[184] A. TURANO, *Robotica e roboetica*, cit., p. 141.
[185] C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 116.
[186] C. IAGNEMMA, *I "robot medici"*, cit., p. 3.
[187] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 492; C. IAGNEMMA, *I "robot medici"*, cit., p. 5.
[188] A. TURANO, *Robotica e roboetica*, cit., p. 141.
[189] Sul tema L.A. TERRIZZI, *Le linee guida in funzione espansiva del diritto penale: quando l'*Unglück *si trasforma in* Unrecht, in *Dir. pen. cont. - Riv. Trim.*, fasc. 7-8/2019, pp. 107 ss.; A.R. DI LANDRO, *Dalle linee guida e dai protocolli all'individualizzazione della colpa penale nel settore sanitario. Misura oggettiva e soggettiva della malpractice*, Torino, 2012; M. CAPUTO, *"Filo d'Arianna" o "Flauto magico"? Linee guida e checklist nel sistema della responsabilità per colpa medica,* in *dirittopenalecontemporaneo.it*, 16.7.2012; P. PIRAS, A. CARBONI, *Linee guida e colpa specifica del medico*, in AA.VV., *Medicina e diritto penale*, cit., pp. 285 ss.
[190] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto,* cit., p. 50.
[191] C. IAGNEMMA, *I "robot medici"*, cit., p. 11.

"regole" cautelari di massima, flessibili e adattabili, prive di carattere precettivo, rispetto alle quali è fatta salva la libertà di scelta professionale del sanitario nel rapportarsi alla specificità del caso concreto, nelle sue molteplici varianti e peculiarità e nel rispetto della "relazione terapeutica" con il paziente»[192]. Tale definizione, probabilmente, mal si concilia con l'esigenza di "codificare" una regola al fine di renderla comprensibile e riproducibile dall'intelligenza artificiale.

Anche in tale frangente emerge chiara l'impossibilità di una completa sostituzione del medico umano, specie se poniamo mente al fatto che, in fin dei conti, l'apprendimento delle strategie di operazione da parte del robot-chirurgo è rimessa alle istruzioni impartite dall'uomo[193]. Sembra dunque che, rispetto all'ambito medico, il ragionamento di fondo non sia troppo distante da quello svolto a proposito delle auto a guida autonoma. Finché l'uomo manterrà una posizione di controllo del sistema intelligente (in tal sede, il robot-chirurgo) la responsabilità per eventuali danni occorsi al paziente sarà da ricondurre a costui (ferma restando l'esclusione della punibilità in caso di osservanza delle linee guida e delle buone pratiche clinico-assistenziali). La responsabilità per danno da prodotto potrebbe invece, verosimilmente, trovare maggiore spazio operativo nel momento in cui dovessero far ingresso nelle sale operatorie robot dotati di livello di automazione 4 o 5, ove il ruolo del medico umano verrebbe sensibilmente ridotto[194].

3.3. Bionica, biorobotica e human enhancement*: fra limiti e confini mobili.*

Uno dei temi che ha destato l'attenzione della letteratura del settore è stato indubbiamente quello della bionica, della biorobotica e delle tecniche di *human enhancement*. Vorremmo tentare di proporre – senza alcuna pretesa di esaustività, stante l'alterità della questione alle nostre competenze – un sommario inquadramento della materia, per poi riflettere sulle problematiche giuridiche ad essa connesse.

La *bionica* è stata definita come «la scienza che studia le funzioni sensorie e motorie degli organismi viventi, al fine di individuare dispositivi tecnologici idonei a riprodurle o potenziarle». La *biorobotica*, invece, è stata intesa come «l'insieme delle teorie e delle applicazioni scientifiche dei sistemi di intelligenza artificiale realizzati nel campo della bionica»[195]. Altra parte della dottrina, invece, non differenzia le due discipline, riferendo in egual modo alla bionica e alla biorobotica la pratica medica con cui una componente meccanica viene "inglobata" all'interno dell'uomo[196]. In ogni caso sembra di poter affermare che la biorobotica si caratterizzi per l'innesto (anche spesso invasivo) di componenti

[192] Cassazione penale, Sez. IV, 03.02.2022, n. 7849, in *DeJure*.
[193] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto*, cit., p. 49.
[194] In realtà la materia della responsabilità per danno da prodotto nel settore medico ha già trovato applicazione nell'ordinamento statunitense, ove si richiede al soggetto danneggiato di dimostrare la riconducibilità dei danni al robot medico in esame. Per una ricostruzione della giurisprudenza americana nel settore dei robot chirurgici v. M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 349 nota 40.
[195] P. MORO, *Biorobotica e diritti fondamentali*, cit., pp. 533-534. Sul tema v. anche A. MONTANARI, *Questioni di tecnoetica in intelligenza artificiale, robotica e bionica*, in P. MORO (a cura di), *Etica, informatica, diritto*, Milano, 2008, pp. 33 ss.
[196] M.B. MAGRO, *Biorobotica*, cit., p. 501.

meccaniche nel corpo umano, il cui controllo sfugge al loro programmatore e che sono volte al recupero o al potenziamento di alcune funzionalità[197]. La biorobotica, inoltre, usa servirsi di interfacce cervello-macchina[198], le quali «consentono di leggere e utilizzare i segnali neurali associati all'attività cognitiva per controllare un arto artificiale o la traiettoria di una piattaforma robotica mobile (ad esempio, l'arto bionico è in grado di riconoscere la volontà del soggetto ed eseguire gli ordini motori del cervello in tempo reale). Esistono anche interfacce cervello-macchina che, convogliando segnali verso il sistema nervoso centrale o periferico di un essere umano, ne modificano significativamente l'attività (come avviene nel caso delle interfacce usate per il controllo del tremore in soggetti affetti dal morbo di Parkinson). Queste ricerche bioniche si propongono soprattutto di *ripristinare* o di *vicariare* funzioni senso-motorie perdute ed aprono la strada al *potenziamento* di apparati senso-motori e cognitivi che funzionano regolarmente»[199].

Tali tecnologie mostrano tutto il loro potenziale nel settore della "robotica riabilitativa e protesica", consentendo al paziente di recuperare funzionalità motorie perdute (magari a seguito di patologie neurologiche) grazie a supporti tecnologici in grado di ripristinare le aree della corteccia cerebrale deputate al controllo del movimento. Si tratta a tutti gli effetti di "protesi intelligenti", in grado cioè di riconoscere il movimento che il paziente intende porre in essere[200], nonché "bioniche", ossia capaci di diventare parte integrante del corpo umano[201].

Il fenomeno dello *human enhancement* è stato definito dal Comitato Nazionale per la Bioetica come l'«uso intenzionale delle conoscenze e tecnologie biomediche per interventi sul corpo umano al fine di modificarne, in senso

---

[197] P. MORO, *Libertà del robot?*, cit., pp. 525-526.

[198] Sul punto v. anche U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale*, human enhancement *e diritti della persona*, in U. RUFFOLO, *Intelligenza artificiale. Il diritto, i diritti, l'etica*, cit., p. 181.

[199] M.B. MAGRO, *Robot*, cit., pp. 1188-1189, corsivo nostro. Ancora, C. SALAZAR, *Umano, troppo umano*, cit., p. 265 riporta il seguente episodio «a un paziente del Policlinico Gemelli di Roma che aveva subito l'amputazione di una mano ne è stata innestata una bionica in grado di sentire gli stimoli tattili attraverso elettrodi intraneurali impiantati nel braccio che assicurano la connessione diretta tra le dita e il cervello, mentre al momento sono in fase di avanzata sperimentazione protesi neuromuscolari e esoscheletri robotici sensibili agli impulsi cerebrali, che riaccendono la speranza di tornare a camminare per le persone costrette su una sedia a rotelle. Risultati di questo genere, come anche quelli raggiunti dagli impianti cocleari e da quelli retinici, spingono a riconsiderare il concetto stesso di *disabilità*, così come l'immissione di "nanomotori" nelle cellule al fine di veicolare in modo non invasivo farmaci nei tessuti potrà, in un futuro prossimo, rivoluzionare le modalità di somministrazione delle cure e, soprattutto, ampliare le possibilità di guarigione per alcune malattie oggi ritenute incurabili». U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona: le frontiere del "transumanesimo"*, in *Giur. it.*, luglio 2019, p. 1659, riportano l'esperimento condotto presso la John Hopkins University che ha portato alla creazione di «un braccio protesico controllabile "con il pensiero", la cui operatività si basa su un *re-mapping* dei nervi umani rimasti integri e sulla loro "fusione" ai "nervi artificiali" del *device* protesico; e ci si prefigura altresì che, una volta raggiunta una più completa integrazione tra componenti neurali artificiali e fisiche, la medesima tecnologia consentirà addirittura ai pazienti di provare sensazioni nell'arto protesico».

[200] «È quanto avviene nelle ipotesi in cui, al fine di recuperare la mobilità degli arti inferiori si applica al paziente l'esoscheletro robotizzato *Lokomat*, che rappresenta un modello assai avanzato di intelligenza artificiale riabilitativa» C. IAGNEMMA, *I "robot medici"*, cit., pp. 5-6.

[201] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale*, human enhancement *e diritti della persona*, cit., p. 181.

migliorativo e/o potenziante, il normale funzionamento»[202]. In dottrina è stato fatto notare come tali tematiche portino con sé una duplice incertezza: da un punto di vista scientifico non sembrano sussistere ancora dati certi in ordine ai possibili effetti benefici o dannosi di tali pratiche; da un punto di vista etico ci si chiede se tale tipologia di interventi particolarmente invasivi sia in grado di modificare la stessa identità personale del soggetto che vi si sottopone[203].

Il principale problema che connota la materia *de qua* consiste nella difficoltà di tracciare una netta linea di demarcazione, universalmente condivisa, tra ciò che rientra nella terapia medica e ciò che invece afferisce al fenomeno dello *human enhancement* in senso stretto[204]. Distinguere ciò che costituisce trattamento terapico (volto a recuperare funzionalità perdute) da un potenziamento *tout court* (volto, invece, ad implementare capacità già esistenti) non costituisce di certo compito semplice[205]: potremmo infatti arrivare a prendere atto dell'inesistenza di una tale linea di confine[206]. Per quanto in realtà è forse il caso di evidenziare che in dottrina sono già state sollevate non poche perplessità in ordine al fondamento giustificativo dei trattamenti potenziativi, a differenza di quelli terapeutici i quali troverebbero copertura costituzionale nel disposto dell'art. 32 Cost.[207]. È stato da altri fatto notare come l'intervento bionico sprovvisto di fini terapeutici potrebbe mettere in pericolo il diritto alla *privacy* e alla libertà personale, ponendo altresì un problema connesso alla natura e ai limiti del consenso prestato ad un intervento bionico[208].

Vediamo, dunque, come la tecnologia abbia già ampiamente dimostrato la sua capacità di disvelare nuovi orizzonti nel campo della medicina. V'è chi, studiando l'aspetto giuridico di tali impianti bionici, non ha esitato ad utilizzare la parola

---

[202] COMITATO NAZIONALE PER LA BIOETICA (CNB), *Neuroscienze e potenziamento cognitivo farmacologico: profili bioetici*, 22.2.2013, p. 5. Più approfonditamente sul tema L. D'AVACK, *Per un uso umano dell'*enhancement, in U. RUFFOLO, *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, cit., pp. 79 ss.; A. D'ALOIA, *I diritti della persona alla prova dello* human enhancement, in U. RUFFOLO, *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, cit., pp. 85 ss.; U. RUFFOLO, A. AMIDEI, *Intelligenza artificiale, biotecnologie e potenziamento: verso nuovi diritti della persona?*, in U. RUFFOLO, *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, cit., pp. 101 ss. Menzionano il fenomeno anche M.B. MAGRO, *Biorobotica*, cit., p. 502; EAD., *Robot*, cit., p. 1189; C. SALAZAR, *Umano, troppo umano*, cit., p. 265, la quale parla dell'*empowerment*; A. AMIDEI, *Robotica intelligente e responsabilità*, cit., p. 79; A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 8; U. RUFFOLO, *Le responsabilità da intelligenza artificiale nel settore medico e farmaceutico*, cit., p. 54; U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona*, cit., p. 1658 parlano di "*tecnologizzazione* dell'uomo", corsivo nostro; G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 3; C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 117. Il Comitato nazionale per la Bioetica si è anche occupato, più nello specifico, del tema dell'*enhancement* in ambito militare, cerando di individuare un sostrato minimo ed essenziale di margini etici al fine di regolamentare la materia *de qua*, v. COMITATO NAZIONALE PER LA BIOETICA (CNB), *Diritti umani, etica medica e tecnologie di potenziamento (enhancement) in ambito militare*, 22.2.2013. Sul tema v. anche C. SALAZAR, *Umano, troppo umano*, cit., p. 269; qualche cenno è rinvenibile anche in U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale*, human enhancement *e diritti della persona*, cit., p. 181.

[203] M.B. MAGRO, *Biorobotica*, cit., pp. 505-506.

[204] A. D'ALOIA, *Oltre la malattia: metamorfosi del diritto alla salute*, in *BioLaw Journal*, 2014, p. 94.

[205] M.B. MAGRO, *Biorobotica*, cit., p. 504.

[206] P. MORO, *Biorobotica e diritti fondamentali*, cit., p. 534.

[207] M.B. MAGRO, *Biorobotica*, cit., pp. 503-504.

[208] P. MORO, *Biorobotica e diritti fondamentali*, cit., p. 535.

cyborg, riferendosi all'ipotesi in cui l'innesto entri in connessione con l'intero organismo umano determinandone un potenziamento delle funzionalità[209]. Sembra che si stia realizzando una sorta di processo osmotico in grado di attenuare le differenze tra le forme di intelligenza naturale e quelle artificiali, "meccanicizzando" l'uomo[210] con estensioni sempre meno "meccaniche" e sempre più "biologiche"[211]. Sorge a questo punto spontaneo chiedersi «fino a che punto potrà continuare a considerarsi "umana" una entità corporea "snaturata", o aumentata, o mantenuta operativa solo grazie ad addizioni o estensioni artefatte?»[212]. La questione attiene, invero, non soltanto alla distinzione tra ciò che debba essere considerato "umano" e ciò che, invece, possa rientrare nel concetto di "disumano", ma anche ad un'esigenza di tutela che vada oltre la "persona", afferendo piuttosto al più lato concetto di "condizione umana"[213].

Gli studiosi che si sono approcciati alla materia hanno rilevato come essa sia in grado di porsi in contrasto con il principio di eguaglianza e con il relativo diritto di accesso a tali sofisticate tecnologie. I costi di queste ultime saranno verosimilmente proibitivi, restando a disposizione di quella limitata parte di cittadinanza in grado di potersi permettere tale tipo di spesa e andando così a creare una sorta di *enhancement divide* tra soggetti "potenziati" e "non potenziati"[214]. Invero tale tipologia di interventi dovrebbe avere uno scopo primario di tipo terapeutico, volto cioè a curare disabilità e patologie al fine di appianare le disuguaglianze tra consociati e non acuirle[215], creando una sorta di "società a due velocità"[216].

Tali riflessioni conducono alla formulazione di una nuova concezione dell'inviolabilità del corpo umano, specie in considerazione del sempre maggiore

---

[209] M.B. MAGRO, *Robot*, cit., p. 1188, l'A. prosegue affermando che «Il dato sorprendente - che distingue l'ibridazione uomo-macchina dalle altre protesi non cibernetiche - risulta proprio nell'interazione tra sistema nervoso, impulsi cerebrali e animazione della protesi, con possibilità di innescare flussi non solo in uscita (cervello-terminazioni nervose-chip-braccio robotico), al fine di comandare il movimento della protesi attraverso gli impulsi cerebrali, ma anche in entrata (braccio robotico, chip, terminazioni nervose, cervello), al fine di restituire al soggetto la percezione del movimento».

[210] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona*, cit., p. 1666.

[211] U. RUFFOLO, *Le responsabilità da intelligenza artificiale nel settore medico e farmaceutico*, cit., p. 55. In dottrina è stato fatto notare come tali invasivi innesti, in grado di divenire parte integrante del corpo umano, pongano problemi giuridici non secondari. Uno fra tutti è quello attinente alla qualificazione giuridica di tali componenti: «questi nuovi arresti della tecnica sono delle cose, che formano oggetto di un *diritto di proprietà* della persona oppure, data la strettissima connessione che in alcuni casi possono avere con le sue *funzioni vitali*, si integrano con essa, diventando parte del corpo? (…) in merito alla natura del danno, l'approccio dovrebbe essere pragmatico, nel senso che quando lo strumento di intelligenza artificiale rappresenti l'unico mezzo che ha il soggetto per svolgere alcune funzioni vitali, allora il danno potrà logicamente essere considerato alla *salute*. Nell'ipotesi invece in cui l'utilizzo dell'intelligenza artificiale persegua solo una funzione *potenziativà* potrebbe aver più senso fare ricorso alla categoria del danno alla *proprietà*» V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 498, corsivo nostro.

[212] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona*, cit., p. 1661.

[213] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona*, cit., p. 1666.

[214] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona*, cit., p. 1668; C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 117.

[215] C. SALAZAR, *Umano, troppo umano*, cit., p. 268.

[216] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale*, human enhancement *e diritti della persona*, cit., p. 204.

margine di operatività che viene lasciato al diritto alla libera autodeterminazione, anche alla luce dell'avvento delle nuove tecnologie[217]. Assistiamo a un progressivo superamento della "visione paternalistica" della tutela del bene salute[218], per accogliere una prospettiva dell'"inviolabilità del corpo" non più caratterizzata dalla tradizionale "intangibilità conservativa" dello stesso[219].

Per quanto, in un primo momento, potremmo essere indotti a individuare i limiti agli atti di disposizione del proprio corpo nel disposto dell'art. 5 c.c., in dottrina è stato fatto notare come, in realtà, tale norma si riferisca agli atti realizzati dall'interessato nei confronti di un altro soggetto e non ai c.d. atti "autodispositivi"[220]. Inoltre i limiti posti dalla succitata norma (ossia la contrarietà alla legge, all'ordine pubblico o al buon costume) sembrano scomparire all'ombra di «una visione del corpo umano come oggetto di un *diritto assoluto incomprimibile* aperto alle scelte di autodeterminazione del singolo»[221].

In tale ottica potrebbe assumere un nuovo significato anche il concetto stesso di "salute" in quanto diritto costituzionalmente tutelato (ex art. 32 Cost.)[222]. L'Organizzazione Mondiale della Sanità ha tradizionalmente definito la salute come «uno stato di completo benessere fisico, psichico e sociale e non semplicemente assenza di malattia o infermità»[223]. Nella misura in cui, dunque, tale benessere fisico e psichico[224] dovesse passare per l'opportunità di un potenziamento umano (determinato dall'esercizio della libertà del singolo alla piena realizzazione di sé)[225] ci si chiede se, spingendoci troppo oltre, non si rischi

---

[217] U. RUFFOLO, *Le responsabilità da intelligenza artificiale nel settore medico e farmaceutico*, cit., p. 58; U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona,* cit., p. 1660.

[218] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona,* cit., p. 1664.

[219] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale,* human enhancement *e diritti della persona*, cit., p. 183.

[220] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale,* human enhancement *e diritti della persona*, cit., p. 187.

[221] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona,* cit., p. 1662, corsivo nostro.

[222] Ciò, invero, è stato fatto notare anche dalla *Risoluzione del Parlamento europeo,* cit., la quale, al Punto 36, ammette che il settore degli sviluppi riparativi e migliorativi del corpo umano sia in grado di «modificare il nostro concetto di corpo umano in salute».

[223] International Health Conference, *Constitution of the world health organization*, New York, 1946, p. 1.

[224] Alla luce di tale ampio concetto di salute si complica ancor di più la distinzione tra un trattamento meramente migliorativo e un trattamento effettivamente curativo. In dottrina è stato proposto come esempio il riconoscimento della legittimità di interventi di rettificazione di sesso i quali, pur essendo estremamente invasivi e pur potendo a prima vista apparire come interventi estetici si palesano, ad un osservatore più attento, come interventi curativi: «per una persona prigioniera di un corpo che le attribuirebbe un'identità sessuale non sentita come propria, risulterebbero connotati da una forte finalità "curativa", in quanto volti al perseguimento del benessere, esteso alla salute psicofisica della persona nella sua interezza» U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale,* human enhancement *e diritti della persona*, cit., p. 198. Gli AA. tornano sul tema in altro scritto ove riflettono sul fatto che la sensazione di costrizione patita da un uomo che si sente prigioniero di un corpo che non sente appartenergli potrebbe rientrare nel nuovo concetto di "malattie dell'animo" U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona,* cit., p. 1665.

[225] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale,* human enhancement *e diritti della persona*, cit., p. 189.

di dar vita a un fenomeno che si ponga in senso contrario al rispetto della dignità umana[226].

Assistiamo ad una sempre crescente "soggettivizzazione" della dignità, strettamente connessa all'idea che di essa ha ciascun consociato[227]. In dottrina si è parlato a tal proposito del c.d. "paradosso della dignità": da un lato troviamo il diritto alla libera autodeterminazione dell'individuo a realizzare la propria identità (secondo una visione soggettivistica), dall'altro la libera ed autonoma realizzazione di sé potrebbe contrastare proprio col concetto stesso di dignità (in ottica oggettivistica). «Nel primo caso l'accento viene posto sugli impulsi e sui bisogni dell'individuo di soddisfare il suo benessere psicofisico, nel secondo su un paradigma di società imposto dall'autorità, in quest'ultimo senso la dignità andrebbe declinata nel dovere di rispettare le diversità i difetti e le debolezze del singolo essere umano»[228].

La realtà è che «il diritto alla piena realizzazione di sé stessi e della propria personalità» costituisce, oggi, una «declinazione dello stesso diritto (non alla mera integrità fisica, ma) alla salute (art. 32 Cost., declinato alla luce del diritto all'autodeterminazione)»[229]. Un'ampia tutela in tal senso è garantita anche dall'art. 8 CEDU il quale, nel sancire il diritto al rispetto della vita privata e familiare, vieta altresì qualsivoglia tipo di ingerenza da parte dell'autorità pubblica nel libero esercizio di tale diritto. Tuttavia, ad uno sguardo più approfondito, sembra che nessuna delle due norme garantisca tutela incondizionata al diritto alla libera autodeterminazione latamente inteso: l'art. 32 Cost., infatti, subordina il diritto alla salute alla tutela dell'interesse della collettività alla salute medesima; l'art. 8 CEDU, invece, ammette l'ingerenza dell'autorità pubblica nella vita privata del cittadino ove essa sia «prevista dalla legge e costituisca una misura che, in una società democratica, è necessaria per la sicurezza nazionale, per la sicurezza pubblica, per il benessere economico del paese, per la difesa dell'ordine e per la prevenzione dei reati, per la protezione della salute o della morale, o per la protezione dei diritti e delle libertà altrui»[230].

In realtà, nonostante l'esistenza di tali argini dal punto di vista normativo, nella giurisprudenza – tanto costituzionale quanto europea – si registra un sempre maggiore riconoscimento al diritto alla libera autodeterminazione ed alla libertà di realizzazione della propria identità personale, sempre però contemperando tali interessi con altri ritenuti egualmente meritevoli di tutela da parte di ciascun ordinamento e senza escludere «l'ammissibilità di ragionevoli limiti al generale diritto all'autodeterminazione»[231]. Sembra dunque che, per quanto ampio possa

---

[226] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 500.

[227] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona,* cit., p. 1664, gli AA. evidenziano, per contro, che «è anche vero, tuttavia, che il medesimo parametro della dignità umana è sovente impiegato, in senso contrario, per individuare il limite tra "lecito" ed "illecito" e talora anche per "proteggere il singolo da sé stesso": sono, così, vietati trattamenti ed interventi che sviliscano la dignità umana».

[228] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., pp. 499-500.

[229] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale,* human enhancement *e diritti della persona*, cit., p. 183.

[230] Sul punto v. anche U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale,* human enhancement *e diritti della persona*, cit., p. 201.

[231] Per una ricostruzione della giurisprudenza della Consulta e della Corte europea dei diritti dell'uomo sull'ampio riconoscimento al diritto di ciascuno all'autodeterminazione v. U. RUFFOLO,

essere l'ambito di operatività di tale diritto fondamentale, esso continui a non caratterizzarsi come diritto assoluto, essendo sempre possibile per l'ordinamento porvi degli argini, purché giustificati e ragionevoli.

Ed è proprio in questa "terra di mezzo" tra consentito e non che entra in gioco il diritto, al quale spetterà l'arduo compito di porre in essere una regolamentazione giuridica proattiva di tali nuovi fenomeni di *enhancement*, tracciando una linea di confine tra potenziamenti leciti e illeciti[232]. A ben pensarci la società di diritto in cui viviamo non è nuova a queste dinamiche, all'alternarsi cioè di fenomeni avversati in un primo momento e, successivamente, regolamentati. Si usa a tal proposito distinguere tra "coscienza reale", la quale annovera nella propria definizione tutto ciò che, in un dato momento storico, trova pacifico riconoscimento sociale, e "coscienza possibile", la quale ricomprende tutti quei fenomeni non ancora socialmente accettati ma che, in prospettiva futuribile, sono forieri di una possibile regolamentazione giuridica. «Così, *la "coscienza possibile" diviene*, con il tempo e con l'evolversi della cultura, *"coscienza reale"*, ed anche la sensibilità giuridica di quella cultura parallelamente si evolve. La contrarietà ai principi di ordine pubblico subisce, così, *frontiere parallelamente mobili*, rendendo accettabile e lecito oggi quanto pareva contrario ad ogni principio ieri; e dunque accettabile domani qualcosa che oggi rifiutiamo (…) Dall'*ostracismo sociale*, e dalla *parallela illiceità*, alla *accettazione sociale* ed al conseguente suo ingresso nel novero delle libertà individuali lecitamente esercitabili»[233]. Anche qui, come per le auto a guida autonoma, la partita si giocherà in prima battuta sul riconoscimento sociale che verrà riservato alle pratiche di *human enhancement*, per poi formulare una corrispondente regolamentazione giuridica.

La questione probabilmente più problematica della materia in esame atterrà alla distinzione tra pratiche consentite e non consentite e all'individuazione dei criteri che verranno adoperati per tracciare tale linea di demarcazione tra ciò che sarà considerato lecito e ciò che, invece, sarà considerato illecito. In dottrina è stato proposto di guardare all'incisività degli effetti di tali pratiche di *enhancement*, vietando interventi particolarmente aggressivi[234], o ancora di impedire pratiche in grado di snaturare la stessa "condizione umana"[235]. Tra i possibili limiti alla libertà di autodeterminazione è stato individuato anche il c.d. potenziamento "ultrattivo", in grado cioè di incidere sul patrimonio genetico di un individuo e, dunque, sui caratteri trasmissibili in via ereditaria. Precludere interventi dalla portata, per così dire, intergenerazionale avrebbe altresì lo scopo

---

A. AMIDEI, *Intelligenza Artificiale*, human enhancement *e diritti della persona*, cit., pp. 189 ss., per la citazione p. 189.

[232] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale*, human enhancement *e diritti della persona*, cit., p. 182.

[233] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona*, cit., p. 1664, corsivo nostro.

[234] Anche se in realtà, a ben vedere, occorrerà fare chiarezza anche su questo punto. In dottrina è stato fatto notare come siano attualmente consentiti interventi di chirurgia plastica estremi, volti a modificare le fattezze di un uomo al fine di farlo assomigliare al proprio idolo o ad un animale. In tale prospettiva, sarebbe contraddittorio ammettere tale tipologia di interventi e non consentire pratiche di potenziamento non curative ma, magari, supportate dall'esercizio del libero diritto all'autodeterminazione. Sul punto U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale*, human enhancement *e diritti della persona*, cit., pp. 198-199.

[235] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona*, cit., p. 1665.

di tutelare le generazioni future, considerate in questo contesto come soggetti endemicamente vulnerabili[236].

Il rispetto per l'"intergenerazionalità" appare invero sancito anche dal Preambolo della Carta di Nizza, ove si evidenzia che il godimento dei diritti sanciti nella Carta medesima sia in grado di far «sorgere responsabilità e doveri nei confronti degli altri come pure della comunità umana e delle generazioni future». È proprio per queste ragioni che in dottrina è stata sostenuta l'opportunità di negare il diritto a sottoporre il proprio corpo ad alterazioni modificative del proprio patrimonio genetico che diventerebbero, pertanto, trasmissibili alle future generazioni. Tale divieto godrebbe della protezione costituzionale dell'art. 32 Cost. nella parte in cui quest'ultimo tutela il diritto alla salute anche alla luce dell'interesse collettivo alla salute stessa[237].

Il su esaminato tema apre la porta a quello che da più parti in dottrina è stato definito come il fenomeno del "transumanesimo" o "post-umanesimo", ossia «una rivoluzione antropologica capace di travalicare il confine tra naturale ed artificiale»[238]. Tale corrente di pensiero, la quale porta con sé implicazioni tanto bioetiche quanto giuridiche, si fonda sull'esaltazione delle potenzialità delle nuove tecnologie per implementare l'evoluzione umana[239]. Per quanto sia stato sostenuto in dottrina che il fenomeno del post-umanesimo sarebbe, in realtà, fondato su una visione riduttiva della personalità e della soggettività umana[240], le tematiche appena affrontate ci insegnano a rivisitare i principi sulla libera autodeterminazione scevri da pregiudizi, tenendo ben presente che «ciò che oggi ci ripugna come pratica "disumana" potrebbe domani divenire socialmente accettato, e dunque considerato ammissibile sul piano del diritto»[241].

3.4. La robotica assistenziale.

Altro terreno fertile di applicazioni e implicazioni dell'intelligenza artificiale in medicina è quello assistenziale. Esistono già diverse tipologie di robot in grado di prestare assistenza ad anziani o disabili, svolgendo classiche attività infermieristiche[242] o aiutando i beneficiari dell'assistenza robotica a svolgere le proprie attività quotidiane in autonomia[243]. Si usa parlare, a tal proposito, dei c.d. *personal care robots* (PCR): si tratta di sistemi robotici in grado di muoversi nello spazio circostante in autonomia svolgendo compiti di vario genere[244].

---

[236] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale,* human enhancement *e diritti della persona*, cit., p. 200. Sul tema v. anche M.B. MAGRO, *Biorobotica*, cit., p. 504.
[237] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona,* cit., p. 1665.
[238] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale,* human enhancement *e diritti della persona*, cit., p. 179. Citazione riportata anche da G. UBERTIS, *Intelligenza artificiale, giustizia penale*, cit., p. 4.
[239] M.B. MAGRO, *Robot,* cit., p. 1189. Sul tema v. anche L. MEZZETTI, *Introduzione*, cit., p. 17.
[240] P. MORO, *Macchine come noi*, cit., p. 47.
[241] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona,* cit., p. 1664.
[242] C. IAGNEMMA, *I "robot medici"*, cit., p. 7, l'A. ricorda che «questa tipologia di intelligenza artificiale è stata particolarmente utile durante l'emergenza sanitaria dovuta alla diffusione dell'infezione SARS-CoV-2: i robot Ivo, Sanbot Elf e Lhf-Connect sono stati impiegati in taluni nosocomi italiani, infatti, per contenere il carico di lavoro del personale infermieristico e ridurre i contatti con i malati contagiosi»; menziona l'impiego dei robot per l'assistenza ai malati anche H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law*, cit., p. 342.
[243] C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 104.
[244] A. TURANO, *Robotica e roboetica*, cit., p. 142.

Tali robot possono essere programmati per realizzare una pluralità di condotte assistenziali, «in una sorta di *reciproco adattamento* tra robot e paziente, anche eventualmente opponendo resistenza rispetto ai movimenti di quest'ultimo o non prestando talora alcuna forma di assistenza onde stimolare il paziente ad effettuare movimenti interamente *motu proprio*; il tutto, modulando l'azione anche sulla base della reattività del paziente alla terapia ed ai suoi progressi e miglioramenti, onde evitare di proseguire modalità terapeutiche non efficaci»[245].

La diffusione di questi sistemi potrebbe diventare sempre più capillare, in particolar modo nel settore riabilitativo[246]. Tali robot mostrano la loro utilità anche in casi di riabilitazione cognitiva, ad esempio nei riguardi di anziani affetti da Alzheimer o di bambini con sindrome autistica[247]. L'interazione uomo-macchina, specie ove siano coinvolti soggetti vulnerabili[248], porta con sé problematiche tanto etico-morali quanto giuridiche.

Preliminarmente è stato fatto notare come tali robot si confrontino con la più intima sfera umana, rendendo pertanto di fondamentale importanza evitare la "disumanizzazione della cura" mediante la sostituzione dell'interazione umana con quella robotica. Si raccomanda pertanto che i robot assistenziali vengano utilizzati comunque come supporto di un'irrinunciabile assistenza umana, la quale si connota altresì per la sua componente morale ed interpersonale[249]. Il tema dei dilemmi morali torna anche in tal sede, potendo la macchina trovarsi a doverne risolvere alcuni: ad esempio, il robot-assistente potrebbe costringere l'anziano allettato che opponga resistenza ad assumere un farmaco[250]?

È proprio in questo contesto che subentrano le problematiche giuridiche. Infatti, nell'esempio pocanzi citato, chi risponderebbe dell'eventuale danno causato dal robot al paziente nel tentativo di costringerlo ad assumere un farmaco o, per contro, nell'eventualità che l'IA non somministri all'anziano i farmaci di cui ha bisogno, così mettendone in pericolo la salute? Torna in tal sede la problematica principale che ha caratterizzato la nostra intera analisi, ossia l'allocazione della responsabilità in caso di verificazione di un evento avverso.

3.5. Profili penalistici comuni.

Vorremmo in tal sede far convogliare alcune considerazioni in punto di imputazione penale di possibili eventi lesivi cagionati "da" o "mediante" i sistemi intelligenti applicati in ambito medico e che potrebbero costituire terreno di riflessione comune per tutti i settori su esaminati.

Anche in tal sede non possiamo fare a meno di chiamare in causa il principio di precauzione in quanto, anche qui, ci muoviamo in un contesto di indubbia incertezza scientifica. Il generale divieto di *neminem laedere* potrebbe giustificare un intervento penale che, però, andrebbe contemperato col rischio di dar vita a divieti "ultraprudenziali" che rischiano di ostacolare le ricerca nel settore dell'intelligenza artificiale applicata all'ambito medico[251].

---

[245] U. RUFFOLO, *Le responsabilità da intelligenza artificiale nel settore medico e farmaceutico*, cit., p. 53.
[246] C. TREVISI, *La regolamentazione in materia di intelligenza artificiale*, cit., p. 8.
[247] C. IAGNEMMA, *I "robot medici"*, cit., p. 6.
[248] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 7.
[249] A. TURANO, *Robotica e roboetica*, cit., p. 143.
[250] P. MORO, *Macchine come noi*, cit., p. 53.
[251] M.B. MAGRO, *Biorobotica*, cit., p. 507.

Per quanto concerne l'imputazione della responsabilità per i danni cagionati dai sistemi di IA, sembra possibile affermare che, finché essi costituiscano meri strumenti nelle mani dell'uomo, sarà quest'ultimo a rispondere degli eventuali danni occorsi, secondo i classici canoni dell'imputazione colposa[252]. Il controllo umano, specie nel contesto della chirurgia robotica, funge da "catalizzatore della responsabilità", anche per quanto concerne malfunzionamenti del sistema riconducibili ad un "malgoverno manutentivo"[253].

La stessa Risoluzione del Parlamento europeo concernente norme di diritto civile sulla robotica predilige uno *human in command approach*, chiarendo che il principio ispiratore del settore dovrebbe essere quello dell'"autonomia supervisionata" dei robot, rimettendo dunque la fase di programmazione iniziale, nonché le scelte finali del trattamento, sempre al sanitario in carne ed ossa[254]. In tale contesto il robot intelligente costituisce *longa manus* del medico[255].

In dottrina è stato a tal proposito rilevato che in realtà, allo stato attuale, la "robotica medicale" non sarebbe ancora capace di agire in totale autonomia, necessitando sempre dell'intervento umano: «di conseguenza, è l'operatore che risponde, quando si verifica un evento avverso, dell'inadeguato utilizzo dei sistemi robotici da cui quest'ultimo sia stato prodotto: e ciò secondo i criteri della responsabilità per colpa»[256].

Tuttavia (come ormai abbiamo imparato) i dispositivi che si servono di algoritmi di *machine learning* e *deep learning* sono in grado di sfuggire al controllo umano, potendo agire in modo autonomo e imprevedibile[257]. La dottrina ha a tal proposito ritenuto che sarebbe ingiusto andare alla ricerca di una forma di colpevolezza in capo al medico nelle ipotesi in cui quest'ultimo si limitasse ad "affidarsi" all'output prodotto dall'IA. Gli oscuri meccanismi di funzionamento di tali tecnologie, uniti alla pluralità di soggetti normalmente coinvolti nella loro creazione, renderanno estremamente complesso individuare il comportamento cui attribuire incidenza causale determinante rispetto alla verificazione dell'evento[258]. La capacità di apprendimento automatico di questi sistemi e la loro imprevedibilità pongono seri problemi di imputazione della responsabilità in capo al programmatore (*rectius*, creatore dell'algoritmo)[259], a meno di non voler intendere la prevedibilità del possibile evento lesivo in termini talmente generici

[252] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 493.
[253] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1757.
[254] *Risoluzione del Parlamento europeo,* cit., Punto 33.
[255] U. RUFFOLO, *Le responsabilità da intelligenza artificiale nel settore medico e farmaceutico*, cit., p. 56.
[256] C. IAGNEMMA, *I "robot medici"*, cit., pp. 7-8.
[257] Ad esempio, circa il rapporto tra intervento bionico e apprendimento automatico, in dottrina è stata posta la domanda, invero un po' provocatoria «l'"*azione*" compiuta con un arto artificiale può essere considerata "*cosciente e volontaria*"?» M.B. MAGRO, *Biorobotica*, cit., p. 503.
[258] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., pp. 494-495. L'A. si concentra più avanti sulla problematica ricostruzione del nesso causale facendo leva sull'imperscrutabilità dell'algoritmo da parte del medico e sulla difficoltà, per il soggetto danneggiato, di individuare il difetto del software, p. 496.
[259] U. RUFFOLO, *Le responsabilità da intelligenza artificiale nel settore medico e farmaceutico*, cit., p. 54.

da includere ogni possibile danno futuro, anche non identificabile *ex ante*, ponendosi però in contrasto con il principio di colpevolezza[260].

La questione è destinata a complicarsi ulteriormente nel momento in cui la lesione o la morte di un paziente fosse da ricondurre a un difetto dell'IA. Torna in campo la materia della responsabilità per danno da prodotto, stavolta con specifico riferimento ai dispositivi medici: «stando alla normativa *de qua*, è compito, anzitutto, del produttore del sistema automatizzato assicurare che questo venga realizzato secondo tutti gli *standard* di sicurezza richiesti dalla legge, seppure non si possa quasi mai garantire l'assoluta assenza di rischi connessi all'impiego del prodotto»[261]. Nel contesto in esame muta anche il concetto stesso di "prodotto" cui siamo tradizionalmente abituati, assumendo quest'ultimo una nozione "ampia". In dottrina è stato fatto notare come in ambito medico si usi ricomprendere nella nozione di "prodotto", ad esempio, l'organo espiantato pronto per il trapianto, il sangue utilizzato per le trasfusioni o, inevitabilmente, le protesi da impiantare, anche quelle intelligenti[262].

Ferma restando la definizione di "prodotto difettoso" fornita dall'art. 117 del codice del consumo, secondo cui deve intendersi come difettoso quel prodotto che non offre la sicurezza che ci si può legittimamente attendere tenuto conto di tutte le circostanze, con riferimento ai prodotti medici la questione sembra un po' complicarsi. «La giurisprudenza, in caso di danni prodotti da dispositivi medici, sembra essersi mossa in questo senso, richiedendo la prova di un prodotto "pericoloso" più che "difettoso". Pericolo che potrebbe annidarsi nell'assenza di linee guida sull'utilizzo di robot che si avvalgono, in autonomia, di algoritmi altamente sofisticati, la cui modalità di funzionamento è spesso sconosciuta al medico»[263].

In tali ipotesi sarà certamente arduo ipotizzare la responsabilità penale del medico che si sia limitato a servirsi di un robot apparentemente privo di difetti nonché dotato della relativa certificazione di conformità. È proprio in simili contesti che viene nuovamente chiamata in causa la responsabilità del produttore, del programmatore e, financo, dell'ente certificatore, per aver immesso sul mercato *devices* difettosi o per non aver monitorato il buon funzionamento dei prodotti mercato, eventualmente procedendo al ritiro degli stessi[264].

In dottrina è stato proposto di individuare il criterio di imputazione nel c.d. "rischio creato", sulla base del quale il danno andrebbe imputato al soggetto che sia «meglio in grado di assicurarsi avverso il rischio incolpevole» nonché di «internalizzare i relativi costi (si tratta in sostanza del principio *cuius commoda eius et incommoda*)»[265]. Indipendentemente dalle scelte di politica criminale per le quali si deciderà di optare nell'andare a individuare il responsabile di un evento

---

[260] C. IAGNEMMA, *I "robot medici"*, cit., p. 9. In tal sede l'A. si concentra diffusamente sulla possibilità di considerare l'IA come diretto autore del reato attribuendo soggettività giuridica ai *devices* medicali. Per una trattazione del tema si rinvia al Cap. II, Sez. III, Par. 17.3.

[261] C. IAGNEMMA, *I "robot medici"*, cit., p. 8, la quale richiama i decreti legislativi 14 dicembre 1992, n. 507 e 25 gennaio 2010, n. 37.

[262] U. RUFFOLO, *Le responsabilità da intelligenza artificiale nel settore medico e farmaceutico*, cit., p. 58.

[263] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 497.

[264] C. IAGNEMMA, *I "robot medici"*, cit., p. 9.

[265] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 496. Per un cenno al tema si rinvia al Cap. II, Sez. III, Par. 18.

lesivo causato dall'IA in ambito medico, ci sembra anche qui che l'avvertimento resti il medesimo: evitare, da un lato, di creare forme di responsabilità oggettiva dei produttori dei *devices* medicali e dei medici che se ne servono, e garantire, dall'altro, un adeguato livello di tutela nei confronti dei pazienti nella loro qualità di potenziali vittime[266].

3.6. Un approccio proattivo: dal consenso del paziente all'impossibile sostituzione del medico.

Il progredire di tali nuove tecnologie nel settore medico renderà necessaria una nuova visione della fondamentale figura del "consenso". È infatti imprescindibile che i medici informino adeguatamente i pazienti in ordine ai possibili benefici e ai relativi pericoli che potrebbero derivare dall'impiego delle nuove tecnologie, nei confronti delle quali potrebbe registrarsi una certa resistenza[267].

Il consenso assumerà un ruolo dirimente anche per l'uso dei sistemi di IA in fase diagnostica. La medicina di precisione sfrutta un enorme quantitativo di informazioni cliniche appartenenti ad altrettanti pazienti i quali, chiaramente, dovranno prestare il proprio consenso all'utilizzo dei loro dati per l'alimentazione delle "biobanche" utilizzate per la ricerca scientifica. In dottrina è stato fatto notare come lo strumento del "consenso specifico" potrebbe costituire un ostacolo rispetto ai possibili benefici che l'impiego dell'IA in fase diagnostica potrebbe portare con sé. Pertanto è stato proposto di sostituire il consenso informato «con uno strumento più flessibile (…) che permetta quella condivisione di dati che potrebbe, proprio grazie all'impiego dell'AI, condurre alla concretizzazione di benefici terapeutici per la nostra o per le successive generazioni»[268].

Il consenso torna alla nostra attenzione anche per quanto concerne le pratiche di *enhancement*. Ad esempio, il consenso dell'avente diritto ex art. 50 c.p., il quale sancisce il principio per il quale *volenti non fit iniuria*, può trovare applicazione solo ove verta su diritti disponibili. Le considerazioni pocanzi svolte su una rinnovata visione del concetto di salute e dell'indisponibilità del proprio corpo, alla luce del principio di autodeterminazione e dei limiti in esso insiti, inducono a chiederci «se sia nella disponibilità dell'interessato cambiare o progettare la propria identità facendo ricorso non soltanto al recupero di funzionalità biologiche ridotte ma al potenziamento di tali funzionalità con l'impianto di un dispositivo bionico»[269].

---

[266] Sull'esigenza di giustizia per i "pazienti-vittime" F.C. LA VATTIATA, *Artificial Intelligence in Healthcare,* cit., p. 15. Sul contemperamento tra ricerca scientifica e progresso tecnologico da un lato, e tutela dei pazienti potenzialmente danneggiati dall'altro, V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 497.

[267] C. IAGNEMMA, *I "robot medici"*, cit., p. 3, la quale richiama a sua volta COMITATO NAZIONALE PER LA BIOETICA (CNB), COMITATO NAZIONALE PER LA BIOSICUREZZA, LE BIOTECNOLOGIE E LE SCIENZE DELLA VITA (CNBBSV), *Intelligenza Artificiale e Medicina: aspetti etici,* cit., p. 17 ove si afferma che «nell'ambito della relazione medico-paziente, informare, soprattutto in questo periodo di transizione, in modo corretto i malati dei rischi e benefici dell'uso della IA con riferimento alle specifiche applicazioni (e anche dei limiti di spiegabilità delle tecnologie "opache"), al fine di garantire la piena consapevolezza delle scelte e assicurando anche percorsi alternativi nella misura in cui emergesse una resistenza all'accettazione delle nuove tecnologie».

[268] C. CASONATO, *Intelligenza artificiale e diritto costituzionale,* cit., p. 108.

[269] P. MORO, *Biorobotica e diritti fondamentali*, cit., p. 535.

Insomma, probabilmente anche in questo ambito le domande aperte sono più delle possibili risposte. Addirittura v'è chi ritiene che il potenziale apporto dell'intelligenza artificiale nel settore medico non andrebbe sopravvalutato, in quanto non è ancora concretamente ipotizzabile una completa sostituzione del medico umano a vantaggio di quello artificiale. Il ruolo del medico in carne ed ossa è destinato a rimanere centrale[270], dovendo sempre residuare in capo a quest'ultimo il c.d. "privilegio di *override*", che consenta cioè all'uomo di superare le scelte proposte dalla macchina e di mantenere "l'ultima parola" su decisioni che possono riguardare la vita o la salute di un paziente[271].

Il settore in esame, a nostro avviso, non è (e non sarà mai) suscettibile di una completa automazione. Intanto perché, come invero già accennato a proposito dell'uso dell'IA in fase di diagnosi, essa non disporrebbe del naturale "sesto senso" dell'uomo. Per quanto un sistema intelligente possa essere addestrato alla comprensione di un cospicuo numero di linee guida, difficilmente esso potrà essere programmato per declinare in ogni modo possibile l'applicazione delle linee guida al caso concreto, non disponendo delle capacità intuitive che spesso consentono di raggiungere una corretta diagnosi oltrepassando i modelli medici prefissati[272].

Inoltre l'uso di tali nuove tecnologie pone un problema di "fiducia", posto che non sarà certo semplice chiedere ad un paziente di fidarsi della terapia proposta da un sistema non del tutto interpretabile dall'uomo[273]. Un eccessivo ricorso ai sistemi di IA rischierebbe di snaturare il rapporto tra medico e paziente, in quanto quest'ultimo è connotato da un qualcosa che non è riproducibile dalla macchina, ossia l'instaurazione di una relazione di fiducia[274].

Dunque, la vera componente umana insostituibile ad opera della macchina è proprio quella dialogico-relazionale, la quale costituisce aspetto essenziale del rapporto medico-paziente. Il dialogo con quest'ultimo non ha il solo fine di scongiurare un'automatica applicazione delle linee guida che, come abbiamo avuto modo di anticipare, non sono volte all'individuazione della miglior strategia terapeutica quanto piuttosto all'ottimizzazione economica delle risorse medicali[275], ma v'è di più. L'attitudine al dialogo è altresì volta al perseguimento della c.d. medicina narrativa, ossia una «modalità di affrontare la malattia tesa a comprenderne il significato in un quadro complessivo, sistemico, più ampio e rispettoso della persona assistita»[276]. Il perseguimento di un approccio medico che

---

[270] U. RUFFOLO, *Le responsabilità da intelligenza artificiale nel settore medico e farmaceutico*, cit., pp. 55-56.

[271] D. AMOROSO, G. TAMBURRINI, *I sistemi robotici ad autonomia crescente tra etica e diritto*, cit., p. 51.

[272] C. IAGNEMMA, *I "robot medici"*, cit., pp. 10-11. L'A. afferma che «Uno stesso processo patologico può manifestarsi secondo un'imprevedibile varietà di modi, non è possibile applicare *in termini meccanicistici* la regola comunemente valida per quella stessa tipologia di casi, ma occorre *modellare* sul paziente *hic et nunc* quella generale indicazione, fino, ove necessario, a disattenderla».

[273] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 5.

[274] C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 117.

[275] L. EUSEBI, *Appunti per una pianificazione terapeutica condivisibile*, in *Riv. it. med. leg*., 2016, fasc. 3, p. 1160.

[276] A. VIRZÌ et al., *Medicina narrativa: cos'è?*, in *Medicina Narrativa*, 2011, fasc. 1, p. 10. Citazione riportata anche da C. IAGNEMMA, *I "robot medici"*, cit., p. 13. Sul tema v. anche C. MAZZUCATO, A. VISCONTI, *Dalla medicina narrativa alla giustizia riparativa in ambito sanitario:*

sia centrato sul paziente[277] necessita di una componente umana che non può essere soddisfatta neanche dalla macchina più sofisticata, non essendo quest'ultima in grado di integrare la dimensione comunicativa dell'alleanza terapeutica, indebolendo così tanto il diritto alla salute quanto quello all'autodeterminazione[278].

Insomma, le possibili applicazioni dell'IA in ambito medico sono in grado di raggiungere obiettivi un tempo molto lontani, tuttavia occorre prendere atto che, ad oggi, nulla di tutto ciò sarebbe possibile senza la gestione di tali sistemi da parte del personale sanitario in carne e ossa[279]. Per quanto il progresso scientifico dei sistemi intelligenti si sia dimostrato in grado di riprodurre alcune componenti tipicamente umane, esso non è ancora in grado di replicare le capacità relazioni o intuitive dell'uomo[280].

Occorre, in ultima analisi, avvicinarsi allo studio della materia *de qua* in modo reattivo e proattivo, per iniziare ad affrontare problematiche che, oggi, appaiono ancora lontane nel tempo, ma che si candidano un domani a impegnare i giuristi[281]. Torna chiara l'esigenza di una stretta cooperazione tra medici, giuristi e informatici per affrontare tali complesse tematiche[282], senza mai perdere di vista il filo conduttore proposto dal Comitato nazionale per la Bioetica e dal Comitato nazionale per la Biosicurezza, le Biotecnologie e le Scienze della vita: «l'esclusione dell'artificiale toglie molte opportunità all'uomo; l'esclusione dell'umano solleva molte criticità dati i limiti dell'artificiale. Bisogna evitare eccessive speranze, ma anche eccessivi timori, con un atteggiamento di fiducia e di cautela»[283].

---

un progetto "integrato" di prevenzione delle pratiche difensive e di risposta alla colpa medica, in *Riv. it. med. leg.*, 2014, fasc. 3, pp. 847 ss.; G. CANZIO, *Medicina e narrativa*, in *Riv. it. med. leg.*, 2014, fasc. 3, pp. 869 ss.; G. ROTOLO, *Profili di responsabilità medica alla "luce" della medicina narrativa*, in *Riv. it. med. leg*., 2014, fasc. 3, pp. 873 ss.

[277] L. BORGHI, E. MOJA, *Medicina centrata sul paziente: uno strumento per ridurre la frequenza delle cause di negligenza?*, in *Riv. it. med. leg*., 2014, fasc. 3, pp. 887 ss.

[278] C. IAGNEMMA, *I "robot medici"*, cit., p. 15.

[279] C. IAGNEMMA, *I "robot medici"*, cit., p. 16.

[280] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 501.

[281] U. RUFFOLO, A. AMIDEI, *Intelligenza Artificiale e diritti della persona,* cit., p. 1660.

[282] U. RUFFOLO, *Le responsabilità da intelligenza artificiale nel settore medico e farmaceutico*, cit., p. 61.

[283] COMITATO NAZIONALE PER LA BIOETICA (CNB), COMITATO NAZIONALE PER LA BIOSICUREZZA, LE BIOTECNOLOGIE E LE SCIENZE DELLA VITA (CNBBSV), *Intelligenza Artificiale e Medicina: aspetti etici*, cit., p. 17.

CONCLUSIONI

Ci troviamo di fronte ad una rivoluzione sociale, prima ancora che normativa. Lo sviluppo della tecnica è destinato ad ampliare le problematiche che il giurista si troverà a dover affrontare[1] e a produrre mutamenti sociali che imporranno, verosimilmente, un adeguamento normativo[2].

Il tema dell'intelligenza artificiale si caratterizza per la sua «portata "intertemporale", sia nel senso che l'AI non sarà una fase passeggera nell'evoluzione tecnologica, ma un modo irreversibile di ridefinire le forme della nostra esistenza; sia nel senso che quello che decidiamo o consentiamo oggi avrà un impatto anche sulle generazioni che verranno dopo di noi, *ora* e *domani*»[3]. È per questo che in dottrina è stato proposto di adottare il c.d. principio di precauzione costituzionale, il quale raccomanda una regolamentazione solida ed effettiva delle nuove tecnologie al fine di evitare che esse, una volta diffuse, possano violare le nostre libertà fondamentali. Tale principio, in altri termini, ruota intorno all'interiorizzazione dei nostri valori da parte dello sviluppo tecnologico[4]. In prospettiva futuribile, occorrerà dunque studiare come fare in modo che i più sofisticati e autonomi sistemi intelligenti si conformino ai valori, alle leggi e all'etica dell'ordinamento in cui sono collocati[5].

Si potrebbe, a tal proposito, meditare di intervenire a monte, delineando una regolamentazione dei sistemi intelligenti che recepisca, da un lato, i principi etici considerati fondamentali, senza limitare, dall'altro, la libertà di ricerca tecnologica. «La ricerca scientifica sull'IA dovrebbe essere contenuta da una regolamentazione che prevenga questi rischi futuri per i diritti e le libertà fondamentali dell'uomo, che limiti la tecnologia di autoapprendimento e che ponga la ricerca robotica a servizio dell'umanità»[6]. Un primo passo in questa direzione è già stato fatto, ad esempio, con la pubblicazione degli "Orientamenti etici per un'IA affidabile", non a caso fatti oggetto di trattazione nell'incipit della nostra indagine[7].

Viviamo in una fase di equilibrio provvisorio[8] che, a nostro avviso, alla luce della continua evoluzione della materia, è destinata a restare tale. Sarà di fondamentale importanza fornire un quadro normativo elastico, che non ostacoli

---

[1] A. CELOTTO, *I robot possono avere diritti?*, cit., p. 98.
[2] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 490.
[3] A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 30.
[4] A. SIMONCINI, *L'algoritmo incostituzionale*, cit., p. 87.
[5] Estremamente chiare sul punto le parole di H. PRAKKEN, *On the problem of making autonomous vehicles conform to traffic law*, cit., pp. 342-343 il quale afferma che «quando vengono utilizzati tali sistemi autonomi, le norme giuridiche non possono più essere considerate come regole del comportamento umano, poiché non sono gli esseri umani ma le macchine che agiscono. Ciò pone il problema di come i sistemi autonomi possano essere progettati in modo tale che il loro comportamento sia conforme alla legge. Si noti che questa domanda deve essere posta indipendentemente dalla questione legale se le macchine possono essere assegnatarie di responsabilità in senso legale. Anche se un essere umano rimane legalmente responsabile o responsabile delle azioni della macchina, l'essere umano deve affrontare il problema di garantire che la macchina si comporti in modo tale che l'uomo responsabile rispetti la legge». Sul punto v. anche R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 52.
[6] M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 21.
[7] Cap. I., Sez. I, Par. 7.
[8] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 371.

l'evoluzione scientifica, nonché sufficientemente generico da potersi adattare alle novità - ad oggi anche imprevedibili - della tecnologia. Il rischio insito nell'altra faccia della medaglia, però, è quello di peccare di eccessiva genericità, rendendo la normativa di settore concretamente inidonea a regolamentare il fenomeno. Per fronteggiare tale esigenza di bilanciamento si potrebbe meditare di sfruttare un meccanismo simile a quello delle norme penali in bianco, ossia redigere una norma generale e astratta, che individui gli elementi essenziali della fattispecie e che rinvii per specificazione ad altro strumento normativo più facilmente aggiornabile, pur con tutti i rischi che tale tecnica legislativa può comportare sul piano del rispetto del principio di sufficiente determinatezza.

Il diritto, finora, si è sempre trovato un passo indietro rispetto all'incedere della tecnologia, costretto a "seguirla" con passo lento e ad intervenire *ex post* a seguito dell'emersione delle problematiche e delle lacune insite in ogni nuovo fenomeno sociale[9]. In tal modo lo Stato ha assunto un ruolo "post-regolatorio"[10] in quanto, non avendo «la forza di indirizzare *ex ante* il progresso tecnologico» si ritrova costretto a dover «disciplinare le conseguenze da esso prodotte»[11]. In dottrina è stato osservato come il diritto non dovrebbe "inseguire" l'evoluzione dell'intelligenza artificiale, dovendo piuttosto dirigerne lo sviluppo in modo proattivo[12], passando «dalla caccia alla guida»[13]. Nella consapevolezza che stare un passo avanti non sarebbe compito semplice per il diritto che, in quanto scienza sociale, è più propenso ad intervenire in una fase susseguente, sarebbe quantomeno auspicabile che esso non perda di vista il fenomeno, rimanendo vigile al punto da non restare più un passo indietro all'evoluzione tecnologica ma da camminare al suo fianco.

Abbiamo imparato a comprendere la "non autosufficienza" del diritto, stante l'indubbia interdisciplinarietà della materia, la quale richiede l'intervento di molti settori diversi da quello giuridico[14]. Si parla già di un "diritto dell'intelligenza artificiale" che, alla luce dell'inestricabile connessione con l'uomo, sarà anche "diritto dell'essere umano"[15].

In questo contesto non resta che chiederci quale ruolo sarà riservato al diritto penale. Secondo alcuni tale ultima branca del diritto potrebbe non essere la via da imboccare: pur essendo possibile considerarlo una "scorciatoia accattivante" «quando si è preteso di combattere "grandi pericoli" con lo strumento penale e, quindi, di risolvere con esso problemi epocali, si è sempre registrato l'insuccesso e – parallelamente – si sono provocate lesioni ai diritti di libertà e arrecati dolori ai singoli»[16]. Inoltre nella materia in esame il diritto penale raggiunge i propri limiti, non essendo possibile fondare la responsabilità penale su presupposti che

[9] G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., pp. 103-104.
[10] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 423.
[11] V. DE BERARDINIS, *L'impiego delle nuove tecnologie in medicina,* cit., p. 489.
[12] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 405.
[13] L. FLORIDI, *Soft Ethics and the Governance of the Digital*, in *Philos. Technol*., 2018, p. 2.
[14] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., pp. 370-371.
[15] C. CASONATO, *Potenzialità e sfide dell'intelligenza artificiale*, cit., p. 182.
[16] L. STORTONI, *Angoscia tecnologica*, cit., p. 89.

prescindano dal rispetto dell'elemento della colpevolezza[17] e del principio di personalità della responsabilità penale[18].

Per quanto sia certamente vero che il diritto penale mantiene il suo ruolo di *extrema ratio* e che esso debba essere maneggiato con estrema cura, ci sembra per contro vero anche il risvolto della medaglia, ossia che nonostante tutte le perplessità che possano sorgere in materia, «rinunciare ad esplorare i nessi tra nuove tecnologie e rinnovamento della fattispecie incriminatrice sarebbe un grave errore»[19]. Il penalista dovrà guardare all'intelligenza artificiale con mente aperta e scevra da pregiudizi, per comprendere quale contributo possa dare il diritto penale "classico" al "*diritto penale che verrà*"[20], magari nella futuribile forma di un "diritto penale robotico"[21].

Ciò che il diritto dovrà certamente aver cura di fare sarà non cedere completamente il controllo a tali sistemi in ambiti particolarmente sensibili, al fine di impedire che l'intelligenza artificiale possa sostituire in toto l'uomo. Basti pensare all'assunzione di decisioni che siano in grado di incidere sulla vita di un soggetto o che siano connotate da una certa discrezionalità. Per quanto, ad avviso di alcuni, la pretesa oggettività dell'IA eliminerebbe il "rumore di fondo" che può condizionare la valutazione umana, riteniamo comunque opportuno il permanere costante della figura dell'uomo, essendo quest'ultimo dotato di connotati non riproducibili dalla macchina (basti pensare all'intuito, alla creatività o all'empatia)[22]. In altre parole, l'IA dovrebbe supportare la decisione umana e non proporsi di sostituirla[23].

Per quanto concerne, invece, le attività quotidiane, esistono già settori in cui l'IA potrebbe "sostituire" l'uomo, come nel caso della guida autonoma[24] o dello svolgimento da parte delle macchine di compiti pericolosi o usuranti[25]. Per certi versi, in realtà, i sistemi intelligenti sono financo in grado di "superare" le capacità umane, basti pensare alla superiore capacità logico computazionale della macchina rispetto a quella umana[26]. Le strabilianti capacità delle macchine sono in grado di ingenerare nell'uomo la c.d. *vergogna prometeica*: quest'ultimo percepisce «la propria subalternità, in quanto novello Prometeo, al mondo delle macchine da lui stesso create, avverte un senso di "dislivello", di non sincronicità,

[17] S. BECK, *Google Cars, Software Agents, Autonomous Weapons Systems*, cit., p. 245.
[18] Sulla problematica declinazione dell'elemento soggettivo in caso di reati commessi "da" o "mediante" l'IA si rinvia al Cap. II, Sez. III, Parr. 17.6-17.6.1.
[19] M. PAPA, *Future crimes*, cit., p. 13.
[20] M. DI FLORIO, *Il diritto penale che verrà*, cit., p. 14.
[21] A. CAPPELLINI, Machina delinquere non potest?, cit., p. 23; I. SALVADORI, *Agenti artificiali*, cit., p. 115. A proposito della fondazione del c.d. "diritto della robotica", U. RUFFOLO, *Per i fondamenti di un diritto della robotica self-learning*, cit., p. 1.
[22] C. CASONATO, *Potenzialità e sfide dell'intelligenza artificiale*, cit., p. 179.
[23] M.B. MAGRO, *Robot*, cit., p. 1187; M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 22.
[24] G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 106.
[25] R. CINGOLANI, D. ANDRESCIANI, *Robots*, cit., p. 39.
[26] M.B. MAGRO, *Robot*, cit., p. 1182; I. SALVADORI, *Agenti artificiali*, cit., p. 115. Parlano, a tal proposito, di "singolarità tecnologica" M. GABBRIELLI, *Dalla logica al deep learning*, cit., p. 22; L. MEZZETTI, *Introduzione*, cit., p. 2; G. MOBILIO, *L'intelligenza artificiale*, cit., p. 406.

tra sé e i propri prodotti meccanici che, sempre più nuovi ed efficienti, lo oltrepassano, facendo sì che si senta "antiquato"»[27].

Per quanto l'intelligenza artificiale sia in grado di elaborare dati e informazioni in modo più efficace di quanto non farebbe l'intelligenza umana, essa difetta della flessibilità del pensiero dell'uomo, della capacità di "riorganizzarsi", potendosi muovere soltanto entro i predeterminati limiti stabiliti dall'algoritmo[28]. Esistono peculiarità umane che restano, a tutt'oggi, appannaggio esclusivo dell'uomo quali l'autocontrollo[29] o l'interazione nelle relazioni sociali[30]. L'IA difetterebbe, inoltre, del c.d. "senso comune", ossia ciò 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»[31].

Dunque, in quest'ottica, non solo l'intelligenza artificiale non può sostituire l'uomo[32] (almeno non ancora)[33], ma non può neanche essere ad esso paragonata, a meno di non voler ridurre drasticamente la poliedricità umana[34]. In altri termini, avvicinarci al tema con apertura mentale «ci impone di evitare ogni fallace pregiudizio antropomorfico ma non ci impedisce di restare "sanamente" antropocentrici»[35], ricordando che siamo noi a servirci di queste nuove tecnologie, non il contrario[36].

La necessità di regolamentare i nuovi rapporti tra esseri umani e macchine non può essere sottovalutata[37]. Nello svolgere tale lavoro normativo il legislatore dovrà altresì tener conto della «percezione collettiva del "rischio artificiale"»[38]. Magari, un domani, la presenza dei sistemi intelligenti nella nostra società verrà considerata come un "normale" rischio nella vita quotidiana. Al momento, però, non sembra che i tempi siano ancora maturi per considerare un sistema di IA potenzialmente pericoloso come un qualcosa che la collettività è tenuta a tollerare, rientrando piuttosto nell'ambito di un rischio "eccezionale"[39]. L'IA porta con sé "dilemmi di civiltà"[40] che andranno affrontati evitando eccessivi catastrofismi da un lato, ma anche atteggiamenti troppo rassicuranti dall'altro[41].

Chiudersi all'utilizzo dell'IA, anche in materia penale, sarebbe "antistorico"[42], «un anacronistico arroccamento su posizioni di retroguardia che

---

[27] L. MEZZETTI, *Introduzione*, cit., p. 13. S. RIONDATO, *Robot: talune implicazioni di diritto penale*, cit., p. 94 parla del timore della possibile perdita di controllo dell'uomo sulle proprie creazioni tecnologiche le quali, un giorno, potrebbero rivoltarsi contro di lui.
[28] C. SALAZAR, *Umano, troppo umano*, cit., p. 260.
[29] P. MORO, *Biorobotica e diritti fondamentali*, cit., p. 542.
[30] M.B. MAGRO, *Robot*, cit., p. 1181.
[31] M.B. MAGRO, *Biorobotica*, cit., p. 512.
[32] V. MANES, *L'oracolo algoritmico*, cit., p. 567; C. CASONATO, *Intelligenza artificiale e diritto costituzionale*, cit., p. 124; C. CAVACEPPI, *L'Intelligenza artificiale applicata al diritto penale*, cit., p. 99.
[33] P. MORO, *Macchine come noi*, cit., p. 49.
[34] M.B. MAGRO, *Decisione umana e decisione robotica*, cit., p. 16. Si rinvia sul punto alle considerazioni svolte al Cap. II, Sez. I, Par. 3.
[35] U. RUFFOLO, Machina delinquere potest?, cit., p. 304.
[36] G. ITALIANO, *Intelligenza Artificiale*, cit., p. 224.
[37] U. PAGALLO, *Intelligenza Artificiale e diritto*, cit., p. 634.
[38] P. SEVERINO, *Intelligenza artificiale*, cit., p. 536.
[39] S. GLESS, E. SILVERMAN, T. WEIGEND, *If robots cause harm, who is to blame?*, cit., p. 433.
[40] G. CORASANITI, *Intelligenza artificiale e diritto*, cit., p. 402.
[41] L. STORTONI, *Angoscia tecnologica*, cit., p. 71.
[42] V. MANES, *L'oracolo algoritmico*, cit., p. 564.

rischierebbero, ben presto, di essere superate dall'impetuoso incedere del processo di trasformazione digitale della nostra società»[43]. Il progresso arriva comunque[44] e non sarebbe saggio farsi trovare impreparati.

L'avvento dell'intelligenza artificiale, con tutte le sue problematiche sociali e giuridiche, non costituisce né la fine della legge[45] né la fine del diritto penale, rappresentando piuttosto la pietra angolare di un diritto penale ragionevole, rivolto all'uomo ma anche aperto alle nuove tecnologie[46]. Il diritto penale andrà inteso in chiave costruttiva, potendolo considerare «utile ed efficace solo ove questo abbracci il conflitto in modo appropriato, attribuisca la responsabilità in modo ragionevole e corrisponda per quanto possibile ai (presunti) valori sociali»[47].

Nessuna delle problematiche analizzate nel corpo della nostra indagine potrà essere risolta soltanto dai giuristi o soltanto dagli informatici, avvertendosi l'urgenza di un lavoro sinergico che consenta agli uni di imparare dagli altri[48]. La necessità di una «mediazione giuridica nella crescita esponenziale del fenomeno A.I. è invocata anche dagli esponenti delle scienze dure che lo studiano o lo generano, e che postulano, in materia, una *legal machinery* non semplicemente reattiva ma fattivamente pro-attiva, capace di precorrere le evoluzioni da disciplinare piuttosto che limitarsi a rincorrerle»[49].

Non esiste confine alle immaginabili applicazioni dell'IA[50] il che, se vogliamo, è naturale frutto del vivere nella c.d. "società delle possibilità"[51] (o forse sarebbe meglio parlare di "società algoritmica"?)[52]. Il futuro, nel bene e nel male, è frutto del passato[53], ed è per questa ragione che le riflessioni che verranno sviluppate e le decisioni che verranno prese oggi avranno un effetto determinante sul domani. Domande tanto ampie non possono pretendere soluzioni definitive, quantomeno non ancora[54], ma ciò non significa che non ci sia già da lavorare. Nello scenario che ci si apre innanzi, risuonano più che mai attuali le parole di Alan Turing: «possiamo vedere nel futuro solo per un piccolo tratto, ma possiamo pure vedere che in questo piccolo tratto c'è molto da fare»[55].

---

[43] P. SEVERINO, *Intelligenza artificiale*, cit., p. 545.
[44] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici,* cit. p. 7.
[45] G. MOBILIO, *L'intelligenza artificiale*, cit., p. 418.
[46] C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1941.
[47] S. BECK, *Google Cars, Software Agents, Autonomous Weapons Systems*, cit., p. 235.
[48] G. ITALIANO, *Intelligenza artificiale, che errore lasciarla agli informatici*, cit., p. 6.
[49] U. RUFFOLO, Machina delinquere potest?, cit., p. 303.
[50] A. VESPIGNANI, *L'algoritmo e l'oracolo,* cit., p. 63.
[51] C. PIERGALLINI, *Intelligenza artificiale,* cit., p. 1771.
[52] M. BASSINI, L. LIGUORI, O. POLLICINO, *Sistemi di Intelligenza Artificiale*, cit., p. 334; G. MOBILIO, *L'intelligenza artificiale*, cit., p. 402; G. ROMANO, *Diritto, robotica e teoria dei giochi*, cit., p. 109; A. D'ALOIA, *Il diritto verso "il mondo nuovo"*, cit., p. 20.
[53] L. FLORIDI, *What the Near Future of Artificial Intelligence*, cit., p. 13; C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1939.
[54] C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale?*, cit., p. 1934.
[55] A.M. TURING, *Computing machinery*, cit., p. 460.

BIBLIOGRAFIA

- ABBOTT R., SARCH A., *Punishing Artificial Intelligence: Legal Fiction or Science Fiction*, in *UC Davis L. Rev.*, 53, 2019, pp. 323 ss.
- AINIS M., *Il regno dell'Uroboro: benvenuti nell'epoca della solitudine di massa*, Milano, 2018.
- AL MUREDEN E., *La sicurezza dei prodotti e la responsabilità del produttore: casi e materiali*, Torino, 2017.
- ALESSANDRI A., *Il nuovo diritto penale delle società*, Milano, 2002.
- ALPA G., *Prefazione*, in RUFFOLO U., *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020, pp. XVII-XVIII.
- ALTAVILLA E., *La colpa*, Torino, 1957.
- AMARA G., *Fra condotta attiva e condotta omissiva: nuovi criteri distintivi e reali conseguenze sul piano dell'imputazione dell'evento*, in *Cass. Pen.*, 2007, pp. 2795 ss.
- AMARELLI G., *Profili pratici della questione sulla natura giuridica della responsabilità degli enti*, in *Riv. it. dir. proc. pen.*, 2006, pp. 151 ss.
- AMIDEI A., *Robotica intelligente e responsabilità: profili e prospettive evolutive del quadro normativo europeo*, in RUFFOLO U., *Intelligenza artificiale e responsabilità*, Milano, 2017, pp. 63 ss.
- AMODIO E., *Rischio penale di impresa e responsabilità degli enti nei gruppi multinazionali*, in *Riv. it. dir. proc. pen.*, 2007, pp. 1287 ss.
- AMOROSO D., TAMBURRINI G., *I sistemi robotici ad autonomia crescente tra etica e diritto: quale ruolo per il controllo umano?*, in *BioLaw Journal*, 2019, pp. 33 ss.
- ANDERSON S.L., *Asimov's "Three Laws of Robotics" and Machine Metaethics*, in *AI & Society*, 22/2008, 10.3.2007, pp. 477 ss.
- ANDERSON S.L., *The Unacceptability of Asimov's Three Laws of Robotics as a Basis for Machine Ethics*, in *Machine Ethics*, a cura di ANDERSON M. e ANDERSON S.L., Cambridge, 2011, pp. 285 ss.
- ANTOLISEI F., *Il rapporto di causalità nel diritto penale*, Padova, 1934.
- ASARO P., *A Body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics*, in LIN P., ABNEY K., BEKEY G.A., *Robot Ethics: The Ethical and Social Implications of Robotics*, Cambridge: MIT Press, 2011, pp. 169 ss.
- AWAD E., DSOUZA S., KIM R., SCHULZ J., HENRICH J., SHARIFF A., BONNEFON J.F., RAHWAN I., *The Moral Machine experiment*, in *Nature*, 2018, vol. 563, pp. 59 ss.
- BALKIN J.B., *The Path of Robotics Law*, in *California Law Review Circuit*, 2015, pp. 45 ss.
- BARBARESCHI S., *Rivoluzione digitale e diritti dei disabili: la tecnologia come fattore inclusivo e la tutela dell'*habeas mentem, in *Gruppo di Pisa, Quaderno monografico*, n. 3, 2021, pp. 229 ss.
- BARBIERI G., *Reato colposo: confini sostanziali tra azione ed omissione e obbligazione giuridica di prevenire l'evento*, in *Cass. pen.*, 2010, pp. 4329 ss.
- BARTOLI R., *Colpevolezza: tra personalismo e prevenzione*, Torino, 2005.
- BARTOLUCCI M.A., *L'art. 8 d.lgs. 231/2001 nel triangolo di Penrose. Tra minimizzazione del rischio-reato d'impresa e "nuove forme" di colpevolezza*, in *dirittopenalecontemporaneo.it*, 9.1.2017.

- BASILE E., *Condotta atipica e imputazione plurisoggettiva: alla ricerca del coefficiente di colpevolezza del concorrente "anomalo"*, in *Riv. it. dir. proc. pen.*, 2005, pp. 1336 ss.
- BASILE F., *Commento all'art. 116 – Reato diverso da quello voluto da taluno dei concorrenti*, in DOLCINI E., GATTA G.L. (diretto da), *Codice penale commentato*, vol. I, IV ed., Milano, 2015, pp. 1852 ss.
- BASILE F., *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Diritto Penale e Uomo*, 29.9.2019.
- BASILE F., *L'alternativa tra responsabilità oggettiva e colpa in attività illecita per l'imputazione della conseguenza ulteriore non voluta, alla luce della sentenza Ronci delle Sezioni Unite sull'art. 586 c.p.*, in *Riv. it. dir. proc. pen.*, 2011, pp. 911 ss.
- BASILE F., *La colpa in attività illecita: un'indagine di diritto comparato sul superamento della responsabilità oggettiva*, Milano, 2005.
- BASILE F., *La responsabilità oggettiva nella più recente giurisprudenza della cassazione relativa agli artt. 116, 584 e 586 c.p.*, in *dirittopenalecontemporaneo.it*, 22.11.2012.
- BASSI A., EPIDENDIO T., *Enti e responsabilità da reato. Accertamento, sanzioni e misure cautelari*, Milano, 2006.
- BASSINI M., LIGUORI L., POLLICINO O., *Sistemi di Intelligenza Artificiale, responsabilità e accountability. Verso nuovi paradigmi?*, PIZZETTI F., *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino, 2018, pp. 333 ss.
- BECK S., *Google Cars, Software Agents, Autonomous Weapons Systems – New Challenges for Criminal Law?*, in HILGENDORF E., SEIDEL U., *Robotics, Autonomics and the Law*, Baden, 2017, pp. 227 ss.
- BECK S., *Intelligent agents and criminal law. Negligence, diffusion of liability and electronic personhood*, in *Robotics and Autonomous Systems*, 2016, pp. 138 ss.
- BECK S., *Robotics and Criminal Law. Negligence, Diffusion of Liability and Electronic Personhood*, in HILGENDORF E., FELDLE J., *Digitization and the Law*, Baden, 2018, pp. 41 ss.
- BELLOTTO N., *High Frequency Trading. Un'indagine ricognitiva sulla rilevanza penale delle condotte manipolative del mercato realizzate dagli algoritmi*, in LAMBRINI P. (a cura di), *Quaderni del dottorato in giurisprudenza dell'Università di Padova*, Milano, 2021, pp. 11 ss.
- BERNARDI A., *La responsabilità da prodotto nel sistema italiano: profili sanzionatori*, in *Riv. trim. dir. pen. econ.*, 2003, pp. 1 ss.
- BERTARINI B., *Tutela della salute, principio di precauzione e mercato del medicinale. Profili di regolazione giuridica europea e nazionale*, Torino, 2016.
- BIAIS B., FOUCAULT T., *HFT and Market Quality*, in *Bankers, Markets & Investors*, n. 128, gennaio-febbraio 2014, pp. 5 ss.
- BILANCETTI M., *La responsabilità civile e penale del medico*, Padova, 2010.
- BISORI L., *L'omesso impedimento del reato altrui nella dottrina e giurisprudenza italiane*, in *Riv. it. dir. proc. pen.*, 1997, pp. 1339 ss.
- BLAIOTTA R., *Causalità giuridica*, Torino, 2010.
- BLAIOTTA R., *La responsabilità medica. Nuove prospettive per la colpa*, in *penalecontemporaneo.it*, 23.3.2012.

- BONNEFON J.F., SHARIFF A., RAHWAN I., *The social dilemma of autonomous vehicles*, in *Science*, vol. 352, 24.6.2016, pp. 1573 ss.
- BORGHI L., MOJA E., *Medicina centrata sul paziente: uno strumento per ridurre la frequenza delle cause di negligenza?*, in *Riv. it. med. leg.*, fasc. 3, 2014, pp. 887 ss.
- BORSARI R., *Intelligenza Artificiale e responsabilità penale: prime considerazioni*, in *MediaLaws*, 20.11.2019, pp. 262 ss.
- BRICOLA F., *Aspetti problematici del cd. rischio consentito nei reati colposi*, in CANESTRARI S., MELCHIONDA A. (a cura di), *Scritti di diritto penale*, Milano, 1997, vol. I, tomo I, pp. 67 ss.
- BRICOLA F., *Responsabilità penale per il tipo e per il modo di produzione*, in AA.VV., *La responsabilità dell'impresa per i danni all'ambiente e ai consumatori*, Milano, 1978, pp. 75 ss.
- BRUSCO C., *Rischio e pericolo, rischio consentito e principio di precauzione. La c.d. "flessibilizzazione delle categorie del reato"*, in *Criminalia*, 2012, pp. 383 ss.
- BURCHARD C., *L'intelligenza artificiale come fine del diritto penale? Sulla trasformazione algoritmica della società*, in *Riv. it. dir. proc. pen.*, 2020, pp. 1909 ss.
- BUTTERFIELD A., NGONDI G.E., *A Dictionary of Computer Science*, Oxford, 2016.
- CADOPPI A., CANESTRARI S., MANNA A., PAPA M., *Diritto Penale*, I, Milano, 2022, pp. 654 ss.
- CAIVANO V., S. CICCARELLI S., DI STEFANO G., FRATINI M., GASPARRI G., GILIBERTI M., LINCIANO N., TAROLA I., *Il* trading *ad alta frequenza. Caratteristiche, effetti, domande di* policy, Documenti di discussione CONSOB n. 5, in *SSRN*, dicembre 2012.
- CALABRESI G., AL MUREDEN E., *Driverless cars. Intelligenza artificiale e futuro della mobilità*, Bologna, 2021.
- CALASSO R., *L'innominabile attuale*, Milano, 2017.
- CALCAGNO R., *Reato omissivo improprio e responsabilità contrattuale, tra "contatto sociale" e contratto: riflessioni sul principio di legalità*, in *Cass. pen.*, 2014, pp. 3559 ss.
- CALO R., *Artificial Intelligence Policy: a Primer and Roadmap*, in *SSRN*, 2017.
- CALO R., *Robotics and the Lessons of Cyberlaw*, in *California Law Review*, 2015, pp. 513 ss.
- CALO R., *Robots in American Law*, in *University of Washington School of Law Legal Studies Research Paper*, n. 4, 2016.
- CAMAIONI S., *Trasferimento e successione di posizioni di garanzia fra riserva di legge e autonomia privata*, in *Riv. it. dir. proc. pen.*, 2010, pp. 1628 ss.
- CAMILLIERI F., *Gli algoritmi predittivi alla luce dei principi delineati nella European Ethical Charter*, in *Gruppo di Pisa, Quaderno monografico*, n. 3, 2021, pp. 313 ss.
- CANEPA A., *L'imputazione soggettiva della colpa. Il reato colposo come punto cruciale nel rapporto tra illecito e colpevolezza*, Torino, 2011.
- CANESTRARI S., *La responsabilità del partecipe per il reato diverso da quello voluto e il principio di colpevolezza*, in *Studium Iuris*, 1996, pp. 1396 ss.

- CANZIO G., CERQUA L.D., LUPARIA L., *Diritto penale delle società, profili sostanziali e processuali*, Padova, 2014.
- CANZIO G., *Intelligenza artificiale, algoritmi e giustizia penale*, in *Sistema Penale*, 8.1.2021.
- CANZIO G., *Medicina e narrativa*, in *Riv. it. med. leg.*, 2014, fasc. 3, pp. 869 ss.
- CAPILLI G., *I criteri di interpretazione della responsabilità*, in ALPA G., *Diritto e intelligenza artificiale*, Pisa, 2020, pp. 457 ss.
- CAPPELLINI A., Machina delinquere non potest? *Brevi appunti su intelligenza artificiale e responsabilità penale*, in *Criminalia*, 2018, ora anche in *disCrimen* dal 27.3.2019.
- CAPPELLINI A., *Profili penalistici delle self-driving cars*, in *Dir. pen. cont. - Riv. Trim.* 2/2019, pp. 325 ss.
- CAPUTO M., *"Filo d'Arianna" o "Flauto magico"? Linee guida e checklist nel sistema della responsabilità per colpa medica*, in *dirittopenalecontemporaneo.it*, 16.7.2012.
- CARMONA A., *La responsabilità degli enti: alcune note sui reati presupposto*, in *Riv. trim. dir. pen. econ.*, 2003, pp. 995 ss.
- CARNEVALI U., *La responsabilità del produttore*, Milano, 1979.
- CAROCCIA F., *Soggettività giuridica dei robot?*, in ALPA G., *Diritto e intelligenza artificiale*, Pisa, 2020, pp. 213 ss.
- CARRER S., *Se l'amicus curiae è l'algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin*, in *Giurisprudenza Penale Web*, 24.4.2019.
- CARROZZA M.C., ODDO C., ORVIETO S., DI MININ A., MONTEMAGN G., *AI: profili tecnologici Automazione e Autonomia: dalla definizione alle possibili applicazioni dell'Intelligenza Artificiale*, in *BioLaw Journal*, 2019.
- CASONATO C., *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Diritto pubblico comparato ed europeo*, 2019, pp. 101 ss.
- CASONATO C., *Potenzialità e sfide dell'intelligenza artificiale*, in *BioLaw Journal*, 2019, pp. 177 ss.
- CASTALDO A., *La concretizzazione del «rischio giuridicamente rilevante»*, in *Riv. it. dir. proc. pen.*, 1995, pp. 1096 ss.
- CASTALDO A., *Responsabilità oggettiva e principio di colpevolezza*, in *Riv. it. dir. proc. pen.*, 1988, pp. 1119 ss.
- CASTALDO A.R., *L'imputazione oggettiva nel delitto colposo d'evento*, Napoli, 1989.
- CASTRONUOVO D., *L'evoluzione teorica della colpa penale tra dottrina e giurisprudenza*, in *Riv. it. dir. proc. pen.*, 2011, pp. 1594 ss.
- CASTRONUOVO D., *La colpa penale*, Milano, 2009.
- CASTRONUOVO D., *La responsabilità colposa nell'esercizio di attività produttive. Profili generali in tema di omicidio o lesioni per violazione delle discipline sulla sicurezza del lavoro e dei prodotti*, in CADOPPI A., CANESTRARI S., PAPA M., *I delitti contro la persona*, I, Torino, 2006, pp. 579 ss.
- CASTRONUOVO D., *Principio di precauzione e beni legati alla sicurezza. La logica precauzionale come fattore espansivo del "penale" nella giurisprudenza della Cassazione*, in *dirittopenalecontemporaneo.it*, 21.7.2011.
- CASTRONUOVO D., *Principio di precauzione e diritto penale. Paradigmi dell'incertezza nella struttura del reato*, Roma, 2012.

- CASTRONUOVO D., *Responsabilità da prodotto e struttura del fatto colposo*, in *Riv. it. dir. proc. pen.*, 2005, pp. 301 ss.
- CAVACEPPI C., *L'intelligenza artificiale applicata al diritto penale: criticità attuali e prospettive future*, in TADDEI ELMI G., CONTALDO A., *Intelligenza artificiale. Algoritmi giuridici*, Ius condendum *o "fantadiritto"?*, Pisa, 2020, pp. 97 ss.
- CAVALLO V., *La responsabilità obbiettiva nel diritto penale*, Napoli, 1937.
- CELOTTO A., *I robot possono avere diritti?*, in *BioLaw Journal*, 2019, pp. 91 ss.
- CENTONZE F., MANTOVANI M., *La responsabilità «penale» degli enti. Dieci proposte di riforma*, Bologna, 2016.
- CERQUETTI G., *La rappresentazione e la volontà dell'evento nel dolo*, Torino, 2004.
- CHIAO V., *Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice*, in *International Journal of Law in Context*, 2019, pp. 126 ss.
- CHIARLONI S., *Riflessioni minime su Intelligenza Artificiale e servizi giuridici*, in *Giur. it., Speciale 170 anni*, 2019, pp. 8 ss.
- CIANI G., *Autore mediato e reato proprio*, in *Cass. Pen.*, 1997, pp. 1001 ss.
- CIANI G., *Brevi considerazioni sulla responsabilità del concorrente per reato diverso da quello voluto*, in *Cass. pen.*, 1996, pp. 3644 ss.
- CINGOLANI R., ANDRESCIANI D., *Robots, macchine intelligenti e sistemi autonomi: analisi della situazione e delle prospettive*, in ALPA G., *Diritto e intelligenza artificiale*, Pisa, 2020, pp. 23 ss.
- CIVELLO G., *La "colpa eventuale" nella società del rischio. Epistemologia dell'incertezza e "verità soggettiva" della colpa*, Torino, 2013.
- COMTE A., *Cours de philosophie positive*, 1 et 2 leçons, 1830-1842.
- CONSORTE F., *Tutela penale e principio di precauzione. Profili attuali, problematicità, possibili sviluppi*, Torino, 2013.
- CONSULICH F., *Il nastro di Möbius. Intelligenza artificiale e imputazione penale nelle nuove forme di abuso del mercato*, in *Banca Borsa Titoli di credito*, 2018, pp. 195 ss.
- CONSULICH F., *Il principio di autonomia della responsabilità dell'ente. Prospettive di riforma dell'art. 8*, in *La responsabilità amministrativa delle società e degli enti*, 2018, pp. 197 ss.
- CONSULICH F., *Tutela del consumatore* (voce), in PALAZZO F., PALIERO C.E. (diretto da), *Commentario breve alle leggi penali complementari*, Padova, 2007, pp. 2967 ss.
- CONTISSA G., LAGIOIA F., SARTOR G., *The Ethical Knob: Ethically-Customisable Automated Vehicles and the Law*, in *Artificial Intelligence and Law*, 2017, pp. 365 ss.
- CONTISSA G., LASAGNI G., SARTOR G., *Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo*, in *Diritto di Internet*, 4/2019, pp. 619 ss.
- COOPER R., DAVIS M., VLIET B.V., *The Mysterious Ethics of High-Frequency Trading*, in *Business Ethics Quarterly*, gennaio 2016.

- CORASANITI G., *Intelligenza artificiale e diritto: il nuovo ruolo del giurista*, in RUFFOLO U., *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020, pp. 395 ss.
- CORN E., *Il principio di precauzione nel diritto penale. Studio sui limiti all'anticipazione della tutela penale*, Torino, 2013.
- CRICENTI G., *Il problema della colpa omissiva*, Padova, 2002.
- CUCCO C., *La partita del diritto penale nell'epoca dei "*drones-crimes*"*, in *Dir. pen. cont. - Riv. Trim.* 2/2019, pp. 304 ss.
- HASHIMOTO D.A. et al., *Artificial Intelligence in Surgery: Promises and Perils*, in *Annals of Surgery*, 2018, pp. 70 ss.
- D'AGOSTINO L., *Gli algoritmi predittivi per la commisurazione della pena*, in *Dir. pen. cont. - Riv. Trim.* 2/2019, pp. 354 ss.
- D'ALOIA A., *I diritti della persona alla prova dello* human enhancement, in RUFFOLO U. *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, Torino, 2020, pp. 85 ss.
- D'ALOIA A., *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale*, in *BioLaw Journal*, 2019, pp. 3 ss.
- D'AVACK L., *Per un uso umano dell'*enhancement, in RUFFOLO U. *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, Torino, 2020, pp. 79 ss.
- DALIA A.A., *Le cause sopravvenute interruttive del nesso causale*, Napoli, 1975.
- DANAHER J., *Robotic Rape and Robotic Child Sexual Abuse: Should They be Criminalized?*, in *Criminal Law and Philosophy*, 2017, pp. 71 ss.
- DE BERARDINIS V., *L'impiego delle nuove tecnologie in medicina*, in ALPA G., *Diritto e intelligenza artificiale*, Pisa, 2020, pp. 489 ss.
- DE FARIA COSTA J., *Contributo per una legittimazione della responsabilità penale delle persone giuridiche,* in *Riv. it. dir. proc. pen.*, 1993, pp. 1238 ss.
- DE FRANCESCO G., *Gli enti collettivi: soggetti dell'illecito o garanti dei precetti normativi?*, in *Dir. pen. proc.*, 2005, pp. 753 ss.
- DE FRANCESCO G., *L'imputazione della responsabilità penale in campo medico-chirurgico: un breve sguardo d'insieme*, in *Riv. it. med. leg.*, 2012, pp. 969 ss.
- DE FRANCESCO G., *La responsabilità degli enti: un nuovo modello di giustizia "punitiva"*, Torino, 2004.
- DE MAGLIE C., *La disciplina della responsabilità amministrativa delle persone giuridiche e delle associazioni*, in *Dir. pen. proc.*, 2001, pp. 1342 ss.
- DE MARISCO A., *Coscienza e volontà nella nozione del dolo*, Napoli, 1930.
- DE SIMONE G., *La responsabilità da reato degli enti: natura giuridica e criteri (oggettivi) d'imputazione*, in *dirittopenalecontemporaneo.it*, 28.10.2012.
- DE SIMONE G., *Persone giuridiche e responsabilità da reato. Profili storici, dogmatici e comparatistici*, Pisa, 2012.
- DE VERO G., *Corso di diritto penale. Parte generale*, Torino, 2020.
- DE VERO G., *I reati societari nella dinamica evolutiva della responsabilità ex crimine degli enti collettivi*, in *Riv. it. dir. proc. pen.*, 2003, pp. 720 ss.
- DE VERO G., *Il nesso causale e il diritto penale del rischio*, in *Riv. it. dir. e proc. pen.*, 2016, pp. 670 ss.
- DE VERO G., *Il progetto di modifica della responsabilità degli enti tra originarie e nuove aporie*, in *Dir. pen. proc.*, 10/2010, pp. 1137 ss.

- DE VERO G., *Il reo quale ente collettivo*, in DE VERO G. (a cura di), *La legge penale. Il reato. Il reo. La persona offesa*, Torino, 2015, pp. 523 ss.
- DE VERO G., *La responsabilità penale delle persone giuridiche*, Milano, 2008.
- DE VERO G., *Prospettive evolutive della responsabilità da reato degli enti collettivi*, in *La responsabilità amministrativa delle società e degli enti*, 2011, pp. 9 ss.
- DE VERO G., *Struttura e natura giuridica dell'illecito di ente collettivo dipendente da reato*, in *Riv. it. dir. proc. pen*., 2001, pp. 1126 ss.
- DELL'ANDRO R., *La fattispecie plurisoggettiva in diritto penale*, Milano, 1956.
- DELMASTRO M., NICITA A., *Big Data. Come stanno cambiando il nostro mondo*, Bologna, 2019.
- DEMURO G.P., *Il dolo*, II, *L'accertamento*, Milano, 2010.
- DI FLORIO M., *Il diritto penale che verrà. Brevi considerazioni sul possibile impiego dell'IA per prevenire il rischio di disastri colposi*, in *Archivio Penale*, 2021.
- DI GIOVINE O., *Il* judge-bot *e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione giurisprudenziale)*, in *Cass. Pen.*, 3/2020, pp. 951 ss.
- DI GIOVINE O., *La responsabilità penale del medico: dalle regole ai casi*, in *Riv. it. med. leg.*, 2013, pp. 78 ss.
- DI LANDRO A.R., *Dalle linee guida e dai protocolli all'individualizzazione della colpa penale nel settore sanitario. Misura oggettiva e soggettiva della malpractice*, Torino, 2012.
- DI MARTINO A., *Danno e rischio da prodotti. Appunti per la rilettura critica di un'esperienza giurisprudenziale italiana*, in BARTOLI R., *Responsabilità penale e rischio nelle attività mediche e d'impresa: (un dialogo con la giurisprudenza)*, Firenze, 2010, pp. 437 ss.
- DOLCINI E., *L'imputazione dell'evento aggravante, un contributo di diritto comparato*, in *Riv. it. dir. proc. pen.*, 1979, pp. 755 ss.
- DOLCINI E., *Responsabilità oggettiva e principio di colpevolezza. Qualche indicazione per l'interprete in attesa di un nuovo codice penale*, in *Riv. it. dir. proc. pen.*, 2000, pp. 863 ss.
- DOMENICI I., *Tecnologie sanitarie innovative: il diritto di fronte all'incertezza scientifica ed etica*, in PICIOCCHI C., FASAN M., REALE C.M. (a cura di), *Le (in)certezze del diritto. Atti delle giornate di studio (17-18 gennaio 2019)*, Collana Quaderni della Facoltà di Giurisprudenza, Università degli Studi di Trento, Vol. n. 49, 2021, pp. 295 ss.
- DONINI M., CASTRONUOVO D. (a cura di), *La riforma dei reati contro la salute pubblica. Sicurezza del lavoro, sicurezza alimentare, sicurezza dei prodotti*, Padova, 2007.
- DONINI M., *Il principio di offensività. Dalla penalistica italiana ai programmi europei*, in *Dir. pen. cont. - Riv. Trim.*, 4/2013, pp. 4 ss.
- DONINI M., *La causalità omissiva e l'imputazione "per l'aumento del rischio". Significato teorico e pratico delle tendenze attuali in tema di accertamenti eziologici probabilistici e decorsi causali ipotetici*, in *Riv. it. dir. proc. pen.*, 1999, pp. 32 ss.
- EDMONDS D., *Uccideresti l'uomo grasso? Il dilemma etico del male minore*, Milano, 2014.

- EUSEBI L., *Appunti per una pianificazione terapeutica condivisibile*, in *Riv. it. med. leg*., 2016, fasc. 3, pp. 1155 ss.
- EUSEBI L., *Il dolo come volontà*, Brescia, 1993.
- FASAN M., *Intelligenza artificiale e pluralismo: uso delle tecniche di profilazione nello spazio pubblico democratico*, in D'ALOIA A., *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, Milano, 2020, pp. 345 ss.
- FASAN M., *L'intelligenza artificiale nella dimensione giudiziaria. Primi profili sottili e spunti dall'esperienza francese per una disciplina dell'AI nel settore della giustizia*, in *Gruppo di Pisa, Quaderno monografico*, n. 3, 2021, pp. 325 ss.
- FASAN M., *Nuove tecnologie e (in)certe risposte del diritto*, in PICIOCCHI C., FASAN M., REALE C.M. (a cura di), *Le (in)certezze del diritto. Atti delle giornate di studio (17-18 gennaio 2019)*, Collana Quaderni della Facoltà di Giurisprudenza, Università degli Studi di Trento, Vol. n. 49, 2021, pp. 265 ss.
- FIANDACA G., *Il reato commissivo mediante omissione*, Milano, 1979.
- FIANDACA G., MUSCO E., *Diritto Penale. Parte generale*, Bologna, 2019.
- FIANDACA G., *Reati omissivi e responsabilità penale per omissione*, in *Il Foro italiano*, 1983, pp. 27 ss.
- FLORIDI L., *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Milano, 2017.
- FLORIDI L., SANDERS J.W., *On the morality of artificial agents*, in *Minds and Machine*, 2004, pp. 349 ss.
- FLORIDI L., *Soft Ethics and the Governance of the Digital*, in *Philos. Technol*., 2018.
- FLORIDI L., *What the Near Future of Artificial Intelligence Could Be*, in *Philos. Technol*., 2019.
- FOFFANI L., *Responsabilità per il prodotto e diritto comunitario: verso un nuovo diritto penale del rischio? Note comparatistiche sugli ordinamenti italiano e spagnolo*, in DONINI M., CASTRONUOVO D. (a cura di), *La riforma dei reati contro la salute pubblica. Sicurezza del lavoro, sicurezza alimentare, sicurezza dei prodotti*, Padova, 2007, pp. 145 ss.
- FORNASARI G., *Dilemma etico del male minore e* ticking bomb *scenario riflessioni penalistiche (e non) sulle strategie di legittimazione della tortura*, Napoli, 2020.
- FORNERO G., *Intelligenza artificiale e filosofia*, in ABBAGNANO N., *Storia della filosofia*, Vol. IV, *La filosofia contemporanea*, a cura di FORNERO G., RESTAINO F., ANTISERI D., Milano, 2013, pp. 536 ss.
- FORTI G., *"Accesso" alle informazioni sul rischio e responsabilità: una lettura del principio di precauzione*, in *Criminalia*, 2006, pp. 155 ss.
- FORTI G., *Colpa ed evento nel diritto penale*, Milano, 1990.
- FORTUNATO G., *Ancora sui rapporti tra il principio di affidamento ed équipe medica*, in *dirittopenalecontemporaneo.it*, 4.5.2017, pp. 31 ss.
- FREITAS P.M., ANDRADE F., NOVAIS P., *Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible*, in PAGALLO U., PALMIRANI M., CASANOVAS P., SARTOR G., VILLATA S., *AI Approaches to the Complexity of Legal Systems*, Berlin-Heidelberg, 2014, pp. 145 ss.
- FROSINI V., *Cibernetica, diritto e società*, Milano, 1973.

- GABBRIELLI M., *Dalla logica al deep learning: una breve riflessione sull'intelligenza artificiale*, in RUFFOLO U., *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, Torino, 2020, pp. 21 ss.
- GALIANO A., LEOGRANDE A., MASSARI S.F., MASSARO A., *I processi automatici di decisione: profili critici sui modelli di analisi e impatti nella relazione con i diritti individuali*, in *Rivista italiana di informatica e diritto*, fasc. 1, 2020, pp. 41 ss.
- GALLO M., *Il dolo oggetto e accertamento*, Milano, 1953.
- GALLO M., *Lineamenti di una teoria sul concorso di persone nel reato*, Milano, 1957.
- GARBEY M., LEE BASS B., BERCELI S., COLLET C., CERVERI P., *Computational Surgery and Dual Training: Computing, Robotics and Imaging*, New York, 2013.
- GARGANI A., *Imputazione del reato agli enti collettivi e responsabilità penale dell'intraneo: due piani irrelati?*, in *Dir. pen. proc.*, 2002, pp. 1061 ss.
- GARGANI A., *La responsabilità omissiva dei titolari di funzioni di protezione civile tra passato e futuro*, in *disCrimen*, 24.6.2019.
- GARGANI A., *Le posizioni di garanzia*, in *Giur. it.*, 2016, pp. 214 ss.
- GARGANI A., *Posizioni di garanzia nelle organizzazioni complesse: problemi e prospettive*, in *Riv. trim. dir. pen. econ.*, 2017, pp. 508 ss.
- GARGANI A., *Responsabilità collettiva da delitto colposo d'evento: i criteri di imputazione nel diritto vivente*, in *La Legislazione penale*, 11.1.2016.
- GARGANI A., *Ubi culpa, ibi omissio. La successione di garanti in attività inosservanti*, in *Ind. Pen.*, 2000, pp. 581 ss.
- GARUTI G., *Responsabilità degli enti per illeciti amministrativi dipendenti da reato*, Padova, 2002.
- GIALUZ M., *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei* risk assessment tools *tra Stati Uniti ed Europa*, in *dirittopenalecontemporaneo.it*, 29.5.2019.
- GIUNTA F., *Il diritto penale e le suggestioni del principio di precauzione*, in *Criminalia*, 2006, pp. 227 ss.
- GIUNTA F., *Illiceità e colpevolezza nella responsabilità colposa*, I, *La fattispecie*, Padova, 1993.
- GIUNTA F., *La legalità della colpa*, in *Criminalia*, 2008, pp. 149 ss., ora anche in *disCrimen*, 28.11.2018.
- GIUNTA F., *La posizione di garanzia nel contesto della fattispecie omissiva impropria*, in *Dir. Pen. Proc.*, 1999, pp. 620 ss.
- GLESS S., *AI in the courtroom: a comparative analysis of machine evidence in criminal trials*, in *Georgetown Journal of International Law*, vol. 51, 2019, pp. 197 ss.
- GLESS S., SILVERMAN E., WEIGEND T., *If robots cause harm, who is to blame? Self-driving cars and criminal liability*, in *New Criminal Law Review*, vol. 19, n. 3, 2016, pp. 412 ss.
- GOMES L., *Hidden obstacles for Google's self-driving cars*, in *MIT Technology Review*, 2014.
- GRASSO G., *Il reato omissivo improprio. La struttura obiettiva della fattispecie*, Milano, 1983.

- GRECO E. *Profili di responsabilità penale del controllore del traffico aereo. Gestione del rischio e imputazione dell'evento per colpa nei sistemi a interazione complessa*, Torino, 2021.
- GREENE J.D., *Our driverless dilemma. When should your car be willing to kill you?*, in *Science*, vol. 352, 24.6.2016, pp. 1514 ss.
- GRIMALDI L., *Giustizia predittiva e garanzie del giusto processo*, in *Gruppo di Pisa, Quaderno monografico*, n. 3, 2021, pp. 353 ss.
- GROSSO C.F., *Il principio di colpevolezza nello schema di delega legislativa per l'emanazione di un nuovo codice penale*, in *Cass. pen.*, 1995, pp. 3125 ss.
- GROSSO C.F., *Questioni aperte in tema di imputazione del fatto*, in *Riv. it. dir. proc. pen.*, 1993, pp. 21 ss.
- GUERRINI R., *La responsabilità da reato degli enti. Sanzioni e loro natura*, Milano, 2006.
- GULLO A., *La responsabilità del partecipe per il reato diverso da quello voluto tra* versari in re illicita *e principio di colpevolezza*, in *Riv. it. dir. proc. pen.*, 2000, pp. 1194 ss.
- GULLO A., *Nuove frontiere tecnologiche e sistema penale: alcune note introduttive*, in *Dir. pen. cont. - Riv. Trim.* 2/2019, pp. VI ss.
- HALLEVY G., *AI v. IP. Criminal Liability for IP Offences of AI Entities*, in *SSRN*, 2015.
- HALLEVY G., *Dangerous Robots – Artificial Intelligence vs. Human Intelligence*, in *SSRN*, 2018.
- HALLEVY G., *I, Robot – I, Criminal — When Science Fiction Becomes Reality: Legal Liability of AI Robots committing Criminal Offenses*, in *Syracuse Journal of Science and Technology Law*, 2010.
- HALLEVY G., *Liability for Crimes Involving Artificial Intelligence Systems*, Svizzera, 2015.
- HALLEVY G., *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, in *Akron Intellectual Property Journal*, 2010, pp. 171 ss.
- HALLEVY G., *Unmanned Vehicles: Subordination to Criminal Law under the Modern Concept of Criminal Liability*, in *SSRN*, 2012.
- HALLEVY G., *Virtual Criminal Responsibility*, in *SSRN*, 2011.
- HARISH V., MORGADO F., STERN A., DAS S., *Artificial Intelligence and Clinical Decision Making: The New Nature of Medical Uncertainty*, in *Academic Medicine*, 2020, pp. 31 ss.
- HATON J.P., *A brief introduction to artificial intelligence*, in *IFAC*, 2006, pp. 8 ss.
- IAGNEMMA C., *I "robot medici": profili problematici in tema di alleanza terapeutica e di responsabilità penale*, in *Corti supreme e salute*, 2/2020.
- IENCA M., *Intelligenza[2]. Per un'unione di intelligenza naturale e artificiale*, Torino, 2019.
- INGRASSIA A., *Il ruolo dell'ISP nel ciberspazio: cittadino, controllore o tutore dell'ordine? Risposte attuali e scenari futuribili di una responsabilità penale dei provider nell'ordinamento italiano*, in *dirittopenalecontemporaneo.it*, 8.11.2012.
- INSOLERA G., *Problemi di struttura del concorso di persone nel reato*, Milano, 1986.

- INSOLERA G., *Tentativo di una diversa lettura costituzionale dell'art. 116 c.p.*, in *Riv. it. dir. proc. pen.*, 1978, pp. 1489 ss.
- ITALIANO G., *Intelligenza artificiale, che errore lasciarla agli informatici*, in *agendadigitale.eu*, 11.6.2019.
- ITALIANO G., *Intelligenza Artificiale: passato, presente, futuro*, in PIZZETTI F., *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino, 2018, pp. 207 ss.
- KAPLAN J., *Intelligenza artificiale. Guida al futuro prossimo*, Roma, 2017.
- KEHL D., GUO P., KESSLER S., *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*, in *Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School*, 2017.
- KING T.C., AGGARWAL N., TADDEO M., FLORIDI L., *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions*, in *Science and Engineering Ethics*, 2020, pp. 89 ss.
- LA ROSA E., *Hechos ofensivos de bienes juridicos y inteligencia artificial: ¿una nueva frontera de la responsabilidad penal?*, in *El sistema juridico ante la digitalizacion. Estudios de derecho publico y criminologia*, 2021, pp. 313 ss.
- LA VATTIATA F.C., *Artificial Intelligence in Healthcare: Risk Assessment and Criminal Law*, in *Diritto Penale e Uomo*, 2.12.2020.
- LA VATTIATA F.C., *Brevi note "a caldo" sulla recente Proposta di Regolamento UE in tema di intelligenza artificiale*, in *Diritto Penale e Uomo*, 30.6.2021.
- LAGIOIA F., SARTOR G., *AI Systems Under Criminal Law: a Legal Analysis and a Regulatory Perspective*, in *Philos. Technol*., 2020, pp. 433 ss.
- LANCELLOTTI G., *La responsabilità della società per il reato dell'amministratore*, Torino, 2003.
- LANEY D., *3D Data Management: Controlling Data Volume, Velocity, and Variety*, in *META Group*, 6.2.2001.
- LATAGLIATA A.R., *I principi del concorso di persone nel reato*, Napoli, 1964.
- LAUKYTE M., *Artificial agents among us: Should we recognize them as agents proper?*, in *Ethics and Information Technology*, 19/2017.
- LAVACCHINI M., *La legittimazione dell'intervento penale tra principio di offensività e principio del danno (*harm principle*)*, in *disCrimen*, 2.8.2019.
- LEENES R., LUCIVERO F., *Laws on Robots, Laws by Robots, Laws in Robots: Regulating Robot Behaviour by Design*, in *Law, Innovation and Technology*, 28.2.2014, pp. 194 ss.
- LEONCINI I., *Obbligo di attivarsi, obbligo di garanzia, obbligo di sorveglianza*, Torino, 1999.
- LEVIS M., PERINI A., *La responsabilità amministrativa della società e degli enti*, Bologna, 2014.
- LIANG H., TSUI B., NI H., et al., *Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence*, in *Nature Medicine*, 25, 11.2.2019, pp. 433 ss.
- LOMBARDI F., *Il principio di affidamento nel trattamento sanitario d'équipe*, in *Giurisprudenza penale*, 2.7.2018.
- LORICCO R., *Autonomous Vehicles: Why we need them, but are unprepared for their arrival*, in *Quinnipiac LR*, vol. 36, 2018, pp. 297 ss.

- LOSANO M.G., *Giuscibernetica: macchine e modelli cibernetici nel diritto*, Torino, 1969.
- MACCHIA A., *Concorso anomalo: un tentativo (azzardato?) di ricostruzione della responsabilità per il fatto diverso da quello voluto*, in *Cass. pen.*, 2/2017, pp. 492 ss.
- MACRÌ E., FURLANETTO A., *I robot tra mito e realtà nell'interazione con le persone, negli ambienti sociali e negli ospedali. Un approccio tra* risk management *e diritto*, in *Riv. it. med. leg.*, fasc. 3, 2017, pp. 1045 ss.
- MADEO A., *Il dolo nella concezione "caleidoscopica" della giurisprudenza*, in *Riv. it. dir. proc. pen.*, 2010, pp. 837 ss.
- MAGRO M.B., *Biorobotica, robotica e diritto penale*, in PROVOLO D., RIONDATO S., YENISEY F., *Genetics, robotics, law, punishment*, Padova, 2014, pp. 499 ss.
- MAGRO M.B., *Decisione umana e decisione robotica. Un'ipotesi di responsabilità da procreazione robotica*, in *La Legislazione Penale*, 10.5.2020.
- MAGRO M.B., *Robot, cyborg e intelligenze artificiali*, in CADOPPI A., CANESTRARI S., MANNA A., PAPA M., *Trattato di diritto penale. Cybercrime*, Torino, 2019, pp. 1179 ss.
- MANES V., *I recenti tracciati della giurisprudenza costituzionale in materia di offensività e ragionevolezza*, in *Dir. pen. cont. - Riv. Trim.* 1/2012, pp. 99 ss.
- MANES V., *Il principio di offensività nel diritto penale. Canone di politica criminale, criterio ermeneutico, parametro di ragionevolezza*, Torino, 2005.
- MANES V., *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in RUFFOLO U., *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020, pp. 547 ss.
- MANIACI G., Harm principle *e* offence principle *secondo un'etica liberale*, in *Criminalia*, 30.1.2019.
- MANNA A., *La c.d. responsabilità amministrativa delle persone giuridiche: un primo sguardo d'insieme*, in *Riv. trim. dir. pen. econ.*, 2002, pp. 501 ss.
- MANNA A., *La c.d. responsabilità amministrativa delle persone giuridiche: il punto di vista del penalista*, in *Cass. pen.*, 2003, pp. 1101 ss.
- MANNA A., *Medicina difensiva e diritto penale*, Pisa, 2014.
- MANTOVANI E., *Intelligenza artificiale e discriminazione: quali prospettive? Il modello inglese del Data trust*, in *Gruppo di Pisa, Quaderno monografico*, n. 3, 2021, pp. 367 ss.
- MANTOVANI F., *Causalità, obbligo di garanzia e dolo nei reati omissivi*, in *Riv. it. dir. proc. pen.*, 2004, pp. 984 ss.
- MANTOVANI F., *Diritto Penale. Parte Generale*, Padova, 2020.
- MANTOVANI F., *Il principio di affidamento nel diritto penale*, in *Riv. it. dir. proc. pen.*, 2002, pp. 536 ss.
- MANTOVANI F., *L'obbligo di garanzia ricostruito alla luce dei principi di legalità, di solidarietà, di libertà e di responsabilità personale*, in *Riv. it. dir. proc. pen.*, 2001, pp. 337 ss.
- MANTOVANI F., *Responsabilità oggettiva espressa e responsabilità oggettiva occulta*, in *Riv. it. dir. proc. pen.*, 1981, pp. 456 ss.
- MANTOVANI M., *Alcune puntualizzazioni sul principio di affidamento*, in *Riv. it. dir. proc. pen.*, 1997, pp. 1051 ss.

- MANTOVANI M., *Il principio di affidamento nella teoria del reato colposo*, Milano, 1997.
- MANYIKA J., CHUI M., BUGHIN J., DOBBS R., BISSON P., MARRS A., *Disruptive technologies: Advances that will transform life, business, and the global economy*, in *McKinsey Global Institute Report*, 1.5.2013.
- MARAS M.H., SHAPIRO L.R., *Child Sex Dolls and Robots: More Than Just an Uncanny Valley*, in *Journal of Internet Law*, 2017, pp. 3 ss.
- MARINUCCI G., *Causalità reale e causalità ipotetica nell'omissione impropria*, in *Riv. it. dir. proc. pen.*, 2009, pp. 523 ss.
- MARINUCCI G., DOLCINI E., GATTA G.L., *Manuale di diritto penale. Parte generale*, Milano, 2022.
- MARINUCCI G., *Finalismo, responsabilità obiettiva, oggetto e struttura del dolo*, in *Riv. it. dir. proc. pen.*, 2003, pp. 363 ss.
- MARINUCCI G., *Innovazioni tecnologiche e scoperte scientifiche: costi e tempi di adeguamento delle regole di diligenza*, in *Riv. it. dir. proc. pen.*, 2005, pp. 29 ss.
- MARINUCCI G., *La colpa per inosservanza di leggi*, Milano, 1965.
- MARINUCCI G., *La responsabilità penale delle persone giuridiche. Uno schizzo storico-dogmatico*, in *Riv. it. dir. proc. pen.*, 2007, pp. 445 ss.
- MASSARO A., *La colpa nei reati omissivi impropri*, Roma, 2011.
- MASSARO A., *La responsabilità colposa per omesso impedimento di un fatto illecito altrui*, Napoli, 2013.
- MASSARO A., *Principio di affidamento e "obbligo di vigilanza" sull'operato altrui: riflessioni in tema di attività medico-chirurgica in* équipe, in *Cass. pen.*, 2011, pp. 3857 ss.
- MASSARO A., *Principio di precauzione e diritto penale:* nihil novi sub sole*? Funzioni e limiti del principio di precauzione de iure condito e condendo*, in *dirittopenalecontemporaneo.it*, 9.5.2011.
- MASUCCI M., *"Fatto" e "valore" nella definizione del dolo*, Torino, 2004.
- MAUGERI A.M., *L'uso di algoritmi predittivi per accertare la pericolosità sociale: una sfida tra* evidence based practices *e tutela dei diritti fondamentali*, in *Archivio Penale*, 1/2021.
- MAZZACUVA N., *Alcune riflessioni su intelligenza artificiale e diritto penale sostanziale*, in RUFFOLO U., *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, Torino, 2020, pp. 287 ss.
- MAZZUCATO C., VISCONTI A., *Dalla medicina narrativa alla giustizia riparativa in ambito sanitario: un progetto "integrato" di prevenzione delle pratiche difensive e di risposta alla colpa medica*, in *Riv. it. med. leg.*, fasc. 3, 2014, pp. 847 ss.
- MCCARTHY J., HAYES P., *Some Philosophical Problems From the Standpoint of Artificial Intelligence*, in *Machine Intelligence*, 1969.
- MCCARTHY J., MINSKY M., ROCHESTER N., SHANNON C, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, 31.8.1955.
- MCCARTHY J., *What is artificial intelligence?*, Stanford University, 12.11.2007.
- MEZZETTI L., *Introduzione*, in RUFFOLO U., *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, Torino, 2020, pp. 1 ss.

- MICHELETTI D., *La normatività della colpa medica nella giurisprudenza della Cassazione*, in AA.VV., *Medicina e diritto penale*, a cura di CANESTRARI S., GIUNTA F., GUERRINI R., PADOVANI T., Pisa, 2009, pp. 247 ss.
- MILITELLO V., *La colpevolezza nell'omissione: il dolo e la colpa del fatto omissivo*, in *Cass. pen.*, 1998, pp. 979 ss.
- MILITELLO V., *Rischio e responsabilità penale*, Milano, 1988.
- MOBILIO G., *L'intelligenza artificiale e i rischi di una* "disruption" *della regolamentazione giuridica*, in *BioLaw Journal*, 2020, pp. 401 ss.
- MONGILLO V., *La responsabilità penale tra individuo ed ente collettivo*, 2018, Torino.
- MONTANARI A., *Questioni di tecnoetica in intelligenza artificiale, robotica e bionica*, in AA. VV., *Etica, informatica, diritto*, a cura di P. MORO, Milano, 2008, pp. 33 ss.
- MORO P., *Biorobotica e diritti fondamentali. Problemi e limiti dell'intelligenza artificiale*, in PROVOLO D., RIONDATO S., YENISEY F., *Genetics, robotics, law, punishment*, Padova, 2014, pp. 533 ss.
- MORO P., *Libertà del robot? Sull'etica delle macchine intelligenti*, in BRIGHI R., ZULLO S., *Filosofia del diritto e nuove tecnologie. Prospettive di ricerca tra teoria e pratica*, Roma, 2015, pp. 525 ss.
- MORO P., *Macchine come noi. Natura e limiti della soggettività robotica*, in RUFFOLO U., *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020, pp. 45 ss.
- MORSELLI E., *Note critiche sulla normativa del concorso di persone nel reato*, in *Riv. it. dir. proc. pen.*, 1983, pp. 422 ss.
- NAPPI A., *Condotta omissiva e colpa per omissione: la causalità tra diritto e processo*, in *Cass. Pen.*, 2004, pp. 4296 ss.
- OCCHIUZZI B., *Algoritmi predittivi: alcune premesse metodologiche*, in *Dir. pen. cont. - Riv. Trim.* 2/2019, pp. 392 ss.
- PADOVANI T., *Le ipotesi speciali di concorso nel reato*, Milano, 1973.
- PAGALLO U., *Intelligenza Artificiale e diritto. Linee guida per un oculato intervento normativo*, in *Sistemi intelligenti*, 2017, pp. 615 ss.
- PAGALLO U., *Profili tecnico-informatici e filosofici*, in CADOPPI A., CANESTRARI S., MANNA A., PAPA M., *Trattato di diritto penale. Cybercrime*, Torino, 2019, pp. 3 ss.
- PAGALLO U., *Saggio sui robot e il diritto penale*, in VINCIGUERRA S., DASSANO F. (a cura di), *Scritti in memoria di Giuliano Marini*, Napoli, 2010, pp. 595 ss.
- PAGALLO U., *Vital, Sophia, and Co. – The Quest for the Legal Personhood of Robots*, in *Information*, 10.9.2018.
- PAGLIARO A., *Colpevolezza e responsabilità obiettiva: aspetti di politica criminale ed elaborazione dogmatica*, in *Riv. it. dir. proc. pen.*, 1988, pp. 387 ss.
- PAGLIARO A., *Diversi titoli di responsabilità per uno stesso fatto concorsuale*, *Riv. it. dir. proc. pen.*, 1994, pp. 3 ss.
- PAGLIARO A., *La responsabilità del partecipe per il reato diverso da quello voluto*, Milano, 1966.
- PALAZZO F., *Il principio di determinatezza nel diritto penale: la fattispecie*, Padova, 1979.

- PALAZZO F., *Offensività e ragionevolezza nel controllo di costituzionalità sul contenuto delle leggi penali*, in *Riv. it. dir. proc. pen*., 1998, pp. 350 ss.
- PALIERO C.E., *L'autunno del patriarca. Rinnovamento o trasmutazione del diritto penale dei codici*, in *Riv. it. dir. proc. pen.*, 1994, pp. 1220 ss.
- PALIERO C.E., *La causalità dell'omissione: formule concettuali e paradigmi prasseologici*, in *Riv. it. med. leg.*, 1992, pp. 821 ss.
- PALIERO C.E., *La responsabilità della persona giuridica nell'ordinamento italiano: profili sistematici*, in PALAZZO F., *Societas puniri potest. La responsabilità da reato degli enti collettivi. Atti del Convegno organizzato dalla Facoltà di giurisprudenza e dal Dipartimento di diritto comparato e penale dell'Università di Firenze (15-16 marzo 2002)*, Padova, 2003, pp. 17 ss.
- PALIERO C.E., *La società punita: del* come, *del* perché, *e del* per cosa, in *Riv. it. dir. proc. pen.*, 2008, pp. 1516 ss.
- PALIERO C.E., *Responsabilità degli enti e principio di colpevolezza al vaglio della Cassazione: occasione mancata o definitivo* de profundis*?*, in *Le Società*, 2014, pp. 474 ss.
- PALMERINI E., *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, in *Responsabilità civile e previdenza*, 6/2016, pp. 1816 ss.
- PALMISANO M., *L'abuso di mercato nell'era delle nuove tecnologie. Trading algoritmico e principio di personalità dell'illecito penale*, in *Dir. pen. cont. - Riv. Trim.* 2/2019, pp. 129 ss.
- PANNAIN R., *Sull'art. 116 del c.p.*, in *Archivio Penale*, 1965.
- PAONESSA C., *Obbligo di impedire l'evento e fisiognomica del potere impeditivo*, in *Criminalia*, 2012, pp. 641 ss., ora anche in *disCrimen*, 4.2.2019.
- PAPA M., *Future crimes: intelligenza artificiale e rinnovamento del diritto penale*, in *Criminalia*, 2019, ora anche in *disCrimen* dal 4.3.2020.
- PARODI C., SELLAROLI V., *Sistema penale e intelligenza artificiale: molte speranze e qualche equivoco*, in *dirittopenalecontemporaneo.it*, fasc. 6/2019, pp. 47 ss.
- PASCULLI M.A., *La responsabilità "da reato" degli enti collettivi nell'ordinamento italiano. Profili dogmatici ed applicativi*, Bari, 2005.
- PATERNITI C., *La responsabilità obiettiva nel diritto penale. Struttura fondamento prospettive*, Milano, 1978.
- PECORARO ALBANI A., *Il dolo*, Napoli, 1955.
- PEDRAZZI C., *Il concorso di persone nel reato*, Palermo, 1952.
- PEDRAZZI C., *Tramonto del dolo?*, in *Riv. it. dir. proc. pen.*, 2000, pp. 1265 ss.
- PELISSERO M., *Il concorso nel reato proprio*, Milano, 2004.
- PELLISSERO M., *L'estensione della responsabilità degli enti i reati colposi. Una riflessione sui rapporti tra parte generale e parte speciale del d. lgs. 231/2001*, in *Scritti in onore di Alfonso M. Stile*, Napoli, 2013, pp. 1199 ss.
- PERIN A., *Standardizzazione, automazione e responsabilità medica. Dalle recenti riforme alla definizione di un modello d'imputazione solidaristico e liberale*, in *BioLaw Journal*, 1/2019.
- PERINI C., *Il concetto di rischio nel diritto penale moderno*, Milano, 2010, pp. 207 ss.

- PICOTTI L., *Diritto penale e tecnologie informatiche: una visione d'insieme*, in CADOPPI A., CANESTRARI S., MANNA A., PAPA M., *Trattato di diritto penale. Cybercrime*, Torino, 2019, pp. 33 ss.
- PIERDONATI M., *Dolo e accertamento nelle fattispecie penali c.d. «pregnanti»*, Napoli, 2012.
- PIERGALLINI C., *Attività produttive e imputazione per colpa: prove tecniche di «diritto penale del rischio»*, in *Riv. it. dir. proc. pen.*, 1997, pp. 1473 ss.
- PIERGALLINI C., *Attività produttive, decisioni in stato di incertezza e diritto penale*, in DONINI M., PAVARINI M. (a cura di), *Sicurezza e diritto penale*, Bologna, 2011, pp. 327 ss.
- PIERGALLINI C., *Danno da prodotto e responsabilità penale*, in *Studium iuris*, 2006, pp. 299 ss.
- PIERGALLINI C., *Danno da prodotto e responsabilità penale. Profili dommatici e politico-criminali*, Milano, 2004.
- PIERGALLINI C., *Intelligenza artificiale: da "mezzo" ad "autore" del reato?*, in *Riv. it. dir. proc. pen.*, 2020, pp. 1745 ss.
- PIERGALLINI C., *La disciplina della responsabilità amministrativa delle persone giuridiche e delle associazioni*, in *Dir. pen. proc.*, 2001, pp. 1342 ss.
- PIERGALLINI C., *La responsabilità del produttore: avamposto o Sackgasse del diritto penale*, in *Riv. it. dir. proc. pen.*, 1996, pp. 352 ss.
- PIERGALLINI C., *La responsabilità del produttore: una nuova frontiera del diritto penale?*, in *Dir. pen. e proc.*, 2007, pp. 1125 ss.
- PIETROPAOLI S., *Habeas Data. I diritti umani alla prova dei big data*, in FARO S., EDOARDO T., PERUGINELLI G., *Dati e algoritmi. Diritto e diritti nella società digitale*, Bologna, 2020, pp. 97 ss.
- PIOLETTI U., *Contributo allo studio del delitto colposo*, Padova, 1990.
- PIRAS P., CARBONI A., *Linee guida e colpa specifica del medico*, in AA.VV., *Medicina e diritto penale*, a cura di CANESTRARI S., GIUNTA F., GUERRINI R., PADOVANI T., Pisa, 2009, pp. 285 ss.
- PIRAS P., LUBINU G.P., *L'attività medica plurisoggettiva fra affidamento e controllo reciproco*, in CANESTRARI S., GIUNTA F., GUERRINI R., PADOVANI T. (a cura di), *Medicina e diritto penale*, Pisa, 2009, pp. 301 ss.
- PIRAS P., *Quando la causa sopravvenuta è sufficiente a determinare l'evento*, in *Dir. pen. proc.*, 1997, pp. 961 ss.
- PIZZETTI F., *Protezione dei dati personali in Italia tra GDPR e codice novellato*, Torino, 2021.
- PIZZETTI F., *La protezione dei dati personali e la sfida dell'Intelligenza Artificiale*, in PIZZETTI F., *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino, 2018, pp. 5 ss.
- PIZZETTI F.G., *The Robot Sophia as a "New Citizen" of Saudi Arabia: what about granting legal personhood, "citizenship" and eventually dignity to non-human entities with artificial intelligence?*, in *Notizie di Politeia*, 2019, pp. 63 ss.
- PRAKKEN H., *On the problem of making autonomous vehicles conform to traffic law*, in *Artificial Intelligence and Law*, 25, 2017, pp. 341 ss.
- PULITANÒ D., *I confini del dolo. Una riflessione sulla moralità del diritto penale*, in *Riv. it. dir. proc. pen.*, 2013, pp. 22 ss.

- PULITANÒ D., *La responsabilità "da reato" degli enti: i criteri d'imputazione*, in *Riv. it. dir. proc. pen.*, 2002, pp. 415 ss.
- QUATTROCOLO S., *An introduction to AI and criminal justice in Europe*, in *Revista Brasileira de Direito Processual Penal*, 2019, pp. 1519 ss.
- QUATTROCOLO S., ANGLANO C., CANONICO M., GUAZZONE M., *Technical Solutions for Legal Challenges: Equality of Arms in Criminal Proceedings*, in *Global Jurist*, 2020.
- QUATTROCOLO S., *Artificial intelligence, computational modelling and criminal proceedings*, Svizzera, 2020.
- QUATTROCOLO S., *Equità del processo penale e* automated evidence *alla luce della Convenzione Europea dei Diritti dell'Uomo*, in *Revista ítalo-española de Derecho Procesal*, 2019, pp. 107 ss.
- QUATTROCOLO S., *Equo processo penale e sfide della società algoritmica*, in *BioLaw Journal*, 2019, pp. 135 ss.
- QUATTROCOLO S., *Forecasting the future while investigating the past. The use of computational models in pre-trial detention decisions*, in *Revista Brasileira de Direito Processual Penal*, 2021, pp. 1859 ss.
- QUATTROCOLO S., *Intelligenza artificiale e giustizia: nella cornice della Carta Etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche*, in *La Legislazione Penale*, 22.3.2018.
- QUATTROCOLO S., *Quesiti nuovi e soluzioni antiche? Consolidati paradigmi normativi vs rischi e paure della giustizia digitale "predittiva"*, in *Cass. Pen.*, 2019, pp. 1748 ss.
- QUINTERO OLIVARES G., *La robótica ante el derecho penal: el vacío de respuesta jurídica a las desviaciones incontroladas*, in *Revista Electronica de Estudios Penales y de la Seguridad*, 2017.
- RAFFAELE S., *Essenza e confini del dolo*, Milano, 2018.
- RAMPONI L., *Concause antecedenti e principio di affidamento: fra causalità attiva ed omissiva*, in *Cass. pen.*, 2008, pp. 566 ss.
- RICCIO G., *Ragionando su intelligenza artificiale e processo penale*, in *Archivio Penale*, 3/2019.
- RICCIO S., *Il reato colposo*, Milano, 1952.
- RICCIO S., *L'autore mediato*, Napoli, 1939.
- RICOLFI M., *Il futuro della proprietà intellettuale nella società algoritmica*, in *Giur. it., Speciale 170 anni*, 2019, pp. 10 ss.
- RIONDATO S., *Robot: talune implicazioni di diritto penale*, in MORO P., SARRA C., *Tecnodiritto. Temi e problemi di informatica e robotica giuridica*, Milano, 2017, pp. 85 ss.
- RIONDATO S., *Robotica e diritto penale (robot, ibridi, chimere, "animali tecnologici")*, in PROVOLO D., RIONDATO S., YENISEY F., *Genetics, robotics, law, punishment*, Padova, 2014, pp. 599 ss.
- RISICATO L., *Combinazione e interferenza di forme di manifestazione del reato*, Milano, 2001.
- RISICATO L., *Gli elementi normativi della fattispecie. Profili generali e problemi applicativi*, Milano, 2004.
- RISICATO L., *Il concorso colposo tra vecchie e nuove incertezze*, in *Riv. it. dir. proc. pen.*, 1998, pp. 132 ss.

- RISICATO L., *Il nuovo statuto penale della colpa medica: un discutibile progresso nella valutazione della responsabilità del personale sanitario*, in *La Legislazione Penale*, 5.6.2017.
- RISICATO L., *L'attività medica di équipe tra affidamento ed obblighi di controllo reciproco. L'obbligo di vigilare come regola cautelare*, Torino, 2013.
- RISICATO L., *La metamorfosi della colpa medica nell'era della pandemia*, in *disCrimen*, 25.5.2020.
- RISICATO L., *La partecipazione mediante omissione a reato commissivo*, in *Riv. it. dir. proc. pen.*, 1995, pp. 1267 ss.
- RISICATO L., *Rileggendo Cesare Pedrazzi*, Il concorso di persone nel reato, *Palermo, 1952*, in *Criminalia*, 2020, pp. 37 ss., ora anche in *disCrimen*, 23.2.2021.
- RIVERDITI M., *La responsabilità degli enti: un crocevia tra repressione e specialprevenzione. Circolarità ed innovazione dei modelli sanzionatori*, Napoli, 2009.
- RODI F., *Gli interventi dell'Unione Europea in materia di intelligenza artificiale e robotica: problemi e prospettive*, in ALPA G., *Diritto e intelligenza artificiale*, Pisa, 2020, pp. 187 ss.
- RODOTÀ S., *Dal soggetto alla persona*, Napoli, 2007.
- RODOTÀ S., *Elaboratori elettronici e controllo sociale*, Bologna, 1973.
- ROMANO G., *Diritto, robotica e teoria dei giochi: riflessioni su una sinergia*, in ALPA G., *Diritto e intelligenza artificiale*, Pisa, 2020, pp. 103 ss.
- ROMANO M., Societas delinquere non potest *(nel ricordo di Franco Bricola)*, in *Riv. it. dir. proc. pen.*, 1995, pp. 1031 ss.
- RONCO M., *Il principio di tipicità della fattispecie penale nell'ordinamento vigente*, Torino, 1979.
- ROSSI A., *La responsabilità degli enti: i soggetti responsabili ed i modelli organizzativi*, in BARTOLI R., *Responsabilità penale e rischio nelle attività mediche e d'impresa: (un dialogo con la giurisprudenza)*, Firenze, 2010, pp. 393 ss.
- ROTOLO G., *Profili di responsabilità medica alla "luce" della medicina narrativa*, in *Riv. it. med. leg.*, fasc. 3, 2014, pp. 873 ss.
- ROVATTI R., *Il processo di apprendimento algoritmico e le applicazioni nel settore legale*, in RUFFOLO U., *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, Torino, 2020, pp. 31 ss.
- RUFFOLO U., AL MUREDEN E., *Autonomous vehicles e responsabilità nel nostro sistema ed in quello statunitense*, in *Giur. it.*, 2019, pp. 1704 ss.
- RUFFOLO U., AMIDEI A., *Intelligenza Artificiale e diritti della persona: le frontiere del "transumanesimo"*, in *Giur. it.*, 2019, pp. 1658 ss.
- RUFFOLO U., AMIDEI A., *Intelligenza artificiale, biotecnologie e potenziamento: verso nuovi diritti della persona?*, in RUFFOLO U., *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, Torino, 2020, pp. 101 ss.
- RUFFOLO U., AMIDEI A., *Intelligenza Artificiale,* human enhancement *e diritti della persona*, in RUFFOLO U., *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020, pp. 179 ss.
- RUFFOLO U., *Il problema della "personalità elettronica"*, in *Journal of Ethics and Legal Technologies*, 2020, pp. 75 ss.

- RUFFOLO U., *Intelligenza Artificiale ed* automotive*: le responsabilità da veicoli* self-driving *e* driverless, in RUFFOLO U., *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020, pp. 153 ss.
- RUFFOLO U., *Intelligenza Artificiale, machine learning e responsabilità da algoritmo*, in *Giur. it.*, 2019, pp. 1689 ss.
- RUFFOLO U., *Le responsabilità da intelligenza artificiale nel settore medico e farmaceutico*, in RUFFOLO U., *Intelligenza artificiale e responsabilità*, Milano, 2017, pp. 53 ss.
- RUFFOLO U., *Le responsabilità da produzione, proprietà e "conduzione" di veicoli autonomi*, RUFFOLO U., *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, Torino, 2020, pp. 163 ss.
- RUFFOLO U., Machina delinquere potest? *Responsabilità ed "illeciti" (anche penali?) della "persona elettronica" e tutele per gli agenti* software *autonomi*, in RUFFOLO U., *XXVI lezioni di Diritto dell'Intelligenza Artificiale,* Torino, 2020, pp. 295 ss.
- RUFFOLO U., *Per i fondamenti di un diritto della robotica self-learning; dalla* machinery *produttiva all'auto* driverless*: verso una "responsabilità da algoritmo"?*, in RUFFOLO U., *Intelligenza artificiale e responsabilità*, Milano, 2017, pp. 1 ss.
- RUFFOLO U., *Self-driving car, auto driverless e responsabilità*, in RUFFOLO U., *Intelligenza artificiale e responsabilità*, Milano, 2017, pp. 31 ss.
- RUGA RIVA C., *Dolo e colpa nei reati ambientali Considerazioni su precauzione, dolo eventuale ed errore*, in *dirittopenalecontemporaneo.it*, 19.1.2015.
- RUGA RIVA C., *Principio di precauzione e diritto penale. Genesi e contenuto della colpa in contesti di incertezza scientifica*, in *Studi in onore di Giorgio Marinucci*, a cura di DOLCINI E., PALIERO C.E., Milano, 2006, pp. 1743 ss.
- RUGGIERO G., *Contributo allo studio della capacità penale. Lo "statuto" della persona fisica e degli enti*, Torino, 2007.
- RUPALI M., AMIT P., *A Review Paper on General Concepts of "Artificial Intelligence and Machine Learning"*, in *International Advanced Research Journal in Science, Engineering and Technology*, 2017, pp. 79 ss.
- RUSSEL S., NORVIG P., *Artificial Intelligence. A modern approach*, US, 2020.
- SALAZAR C., *Umano, troppo umano...o no? Robot, androidi e cyborg nel "mondo del diritto" (prime notazioni)*, in *BioLaw Journal*, 2014, pp. 255 ss.
- SALVADORI I., *Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale*, in *Riv. it. dir. proc. pen.*, 2021, pp. 83 ss.
- SALVEMME I., *Il ruolo del principio di precauzione nel "nuovo" diritto penale dell'ambiente*, in *dirittopenalecontemporaneo.it*, 25.5.2018.
- SAMMARCO G., *Il concetto di autore e partecipe del reato nella più recente dottrina tedesca*, in *Riv. it. dir. proc. pen.*, 1979, pp. 1023 ss.
- SANTI F., *La responsabilità delle società e degli enti. Modelli di esonero delle imprese. D.Lgs. 8/6/2001, n. 231. D.M. 26/6/2003, n. 201*, Milano, 2004.
- SANTOSUOSSO A., BOSCARATO C., CAROLEO F., *Robot e diritto: una prima ricognizione*, in *Nuova Giurisprudenza Civile Commentata*, 2012.
- SARTOR G., *Intelligenza artificiale e diritto. Un'introduzione*, Milano, 1996.
- SCALZINI S., *Alcune questioni a proposito di Algoritmi, Dati, Etica e Ricerca*, in *Riv. it. med. leg.*, fasc. 1, 2019, pp. 169 ss.

- SCHANK R.C., *What's IA, Anyway?*, in *IA Magazine*, 1987.
- SCHWAB K., *La quarta rivoluzione industriale*, Milano, 2016.
- SCORDAMAGLIA I., *Il diritto penale della sicurezza del lavoro tra i principi di prevenzione e di precauzione*, in *dirittopenalecontemporaneo.it*, 23.11.2012.
- SEARLE J.R., *Minds, brains and programs*, in *Behavioral and Brain Sciences*, 1980, pp. 417 ss.
- SELVAGGI N., *L'interesse dell'ente collettivo: quale criterio di ascrizione della responsabilità da reato*, Napoli, 2006.
- SEMINARA S., *Tecniche normative e concorso di persone nel reato*, Milano, 1987.
- SEVERINO P., *Intelligenza artificiale e diritto penale*, in U. RUFFOLO, *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020, pp. 531 ss.
- SGUBBI F., *Il diritto penale totale*, Bologna, 2019.
- SGUBBI F., *Responsabilità penale per omesso impedimento dell'evento*, Padova, 1975.
- SHORTLIFFE E.H., SEPÚLVEDA M.J., *Clinical Decision Support in the Era of Artificial Intelligence*, in *JAMA*, 5.11.2018, pp. 2199-2200.
- SIGNORATO S., *Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo*, in *Rivista di diritto processuale*, 2/2020, pp. 605 ss.
- SIMAAN N., YASIN M.R., WANG L., *Medical Technologies and Challenges of Robot assisted Minimally Invasive Intervention and Diagnostics*, in *Ann. Rev. Control, Robotics, and Autonomous Systems*, fasc. 1, 2018, pp. 465 ss.
- SIMMLER M., MARKWALDER N., *Guilty robots – Rethinking the nature of culpability and legal personhood in an age of artificial intelligence*, in *Criminal Law Forum*, 4.12.2018.
- SIMONCINI A., *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal*, 2019, pp. 63 ss.
- SIMONINI G.F., *L'intelligenza artificiale guida le nostre vetture*, Modena, 2018.
- SINISCALCO M., voce *Autore mediato*, in *Enc. dir.*, IV, Milano, 1959, pp. 445 ss.
- SOLUM L.B., *Artificially Intelligent Law*, in *BioLaw Journal*, 2019, pp. 53 ss.
- SOLUM L.B., *Legal Personhood for Artificial Intelligences*, in *North Carolina Law Review*, 70, 1992, pp. 1231 ss.
- SPASARI M., *L'omissione nella teoria della fattispecie penale*, Milano, 1957.
- SPILLER E., If data is the new atoms… *Le incertezze sul concetto di dato personale al tempo dei big data*, in PICIOCCHI C., FASAN M., REALE C.M. (a cura di), *Le (in)certezze del diritto. Atti delle giornate di studio (17-18 gennaio 2019)*, Collana Quaderni della Facoltà di Giurisprudenza, Università degli Studi di Trento, Vol. n. 49, 2021, pp. 273 ss.
- SPINA A., *La medicina degli algoritmi: Intelligenza Artificiale, medicina digitale e regolazione dei dati personali*, in PIZZETTI F., *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino, 2018, pp. 319 ss.
- STELLA F., *Criminalità di impresa: nuovi modelli di intervento*, in *Riv. it. dir. proc. pen.*, 1999, pp. 1254 ss.
- STILE A.M., *Responsabilità oggettiva e giudizio di colpevolezza*, Napoli, 1989.
- STOCCO R., *Alla ricerca di una dimensione costituzionale dell'art. 116 c.p.*, in *Cass. pen.*, 1990, pp. 35 ss.

- STORTONI L., *Angoscia tecnologica ed esorcismo penale*, in *Riv. it. dir. proc. pen.*, 2004, pp. 71 ss.
- STRADELLA E., *La regolazione della Robotica e dell'Intelligenza artificiale: il dibattito, le proposte, le prospettive. Alcuni spunti di riflessione*, in *MediaLaws*, 10.3.2019, pp. 73 ss.
- SURDEN H., *Artificial Intelligence and Law: An Overview*, in *Georgia State University Law Review*, 2019, pp. 1305 ss.
- SURDEN H., *Machine Learning and the Law*, in *Washington Law Review*, vol. 89, n. 1, 2014, pp. 87 ss.
- SURDEN H., WILLIAMS M.A., *Technological opacity, predictability, and self-driving cars*, in *Cardozo Law Review*, 2016, pp. 121 ss.
- TAFANI D., *Sulla moralità artificiale. Le decisioni delle macchine tra etica e diritto*, in *Rivista di filosofia*, 2020, pp. 81 ss.
- TAGLIARINI F., *I delitti aggravati dall'evento*, Padova, 1979.
- TAMBURRINI G., *Etica delle macchine. Dilemmi morali per robotica e intelligenza artificiale*, Roma, 2020.
- TERRIZZI L.A., *Le linee guida in funzione espansiva del diritto penale: quando l'*Unglück *si trasforma in* Unrecht, in *Dir. pen. cont. - Riv. Trim.*, fasc. 7-8/2019, pp. 107 ss.
- TEUBNER G., *Ibridi ed attanti. Attori collettivi ed enti non umani nella società e nel diritto*, Milano, 2015.
- TEUBNER G., *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi*, Napoli, 2019.
- TITOMANLIO R., *Il principio di precauzione fra ordinamento europeo e ordinamento italiano*, Torino, 2018.
- TOFFALORI C., *Algoritmi*, Bologna, 2015.
- TRAVERSI A., *Intelligenza artificiale applicata alla giustizia: ci sarà un giudice* robot*?*, in *Questione Giustizia*, 10.4.2019.
- TREVISI C., *La regolamentazione in materia di intelligenza artificiale, robot, automazione: a che punto siamo*, in *MediaLaws*, 25.6.2018.
- TROYER L., *I nuovi reati ambientali "abusivi": quando la rinuncia alla legalità penale diviene un illusorio instrumentum regni*, in *Criminalia*, 2015, pp. 329 ss.
- TUMMINELLO L., *Sicurezza alimentare e diritto penale: vecchi e nuovi paradigmi tra prevenzione e precauzione*, in *dirittopenalecontemporaneo.it*, 15.10.2013.
- TURANO A., *Robotica e roboetica: questioni e prospettive nazionali ed europee*, in ALPA G., *Diritto e intelligenza artificiale*, Pisa, 2020, pp. 125 ss.
- TURING A.M., *Computing machinery and intelligence*, in *Mind*, 1950, pp. 433 ss.
- TUZET G., *L'algoritmo come pastore del giudice? Diritto, tecnologie, prova scientifica*, in *MediaLaws*, 16.3.2020.
- UBERTIS G., *Intelligenza artificiale, giustizia penale, controllo umano significativo*, in *Sistema Penale*, 11.11.2020.
- VALLINI A., *"Cause sopravvenute da sole sufficienti" e nessi tra condotte. Per una collocazione dell'art. 41, comma 2, c.p. nel quadro teorico della causalità "scientifica"*, in *dirittopenalecontemporaneo.it*, 11.7.2012.
- VANNINI O., *I reati commissivi mediante omissione*, Roma, 1916.

- VARSHNEY H., KHAN R. A., KHAN U., VERMA R., *Approaches of Artificial Intelligence and Machine Learning in Smart Cities: Critical Review*, in *IOP Conf. Ser.: Mater. Sci. Eng.*, 2021.
- VENANZONI A., *Intersezioni costituzionali - Internet e Intelligenze Artificiali tra ordine spontaneo, natura delle cose digitale e garanzia dei diritti fondamentali*, in *Forum di Quaderni Costituzionali*, 27.4.2018.
- VENEZIANI P., *Regole cautelari "proprie" e "improprie" nella prospettiva delle fattispecie colpose causalmente orientate*, Padova, 2003.
- VERUGGIO G., *Euron Roboethics Roadmap*, 2007.
- VESPIGNANI A., *L'algoritmo e l'oracolo. Come la scienza predice il futuro e ci aiuta a cambiarlo*, Milano, 2019.
- VIRZÌ A. et al., *Medicina narrativa: cos'è?*, in *Medicina Narrativa*, fasc. 1, 2011, pp. 9 ss.
- WATERS B., *Citizen Sophia: It's (Past) Time to Legislate Robotics Regulation*, in *Georgetown Law Technology Review*, 2017.
- YANG G.Z. et al., *Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy*, in *Science Robotics*, 2017.
- YANG X., WANG Y., BYRNE R., SCHNEIDER G., YANG S., *Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery*, in *Chem. Rev.*, 2019, pp. 10520 ss.
- YATES J., *Paura e società del rischio. Un'intervista a Ulrich Beck*, in *Lo Sguardo – Rivista di filosofia*, 2016, pp. 209 ss.
- ZARA G., *Tra il probabile e il certo. La valutazione del rischio di violenza e di recidiva criminale*, in *dirittopenalecontemporaneo.it*, 20.5.2016.
- ZHANG C., LU Y., *Study on artificial intelligence: The state of the art and future prospects*, in *Journal of Industrial Information Integration*, 2021.
- ZIRULIA S., *Esposizione a sostanze tossiche e responsabilità penale*, Milano, 2018.
- ZORNOZA A., LAUKYTE M., *Robotica e diritto: riflessioni critiche sull'ultima iniziativa di regolamentazione in Europa*, in *Contr. Impr. Eur.*, 2/2016, pp. 808 ss.

SITOGRAFIA

- https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
- https://www.eurai.org/
- https://www.tesla.com/it_IT/AI
- https://oecd.ai/en/ai-principles
- https://www.lexico.com/definition/post-truth
- http://www.roboethics.org/index_file/Roboethics%20Roadmap%20Rel.1.2.pdf
- https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf
- https://ethicsinaction.ieee.org/wp-content/uploads/ead1e.pdf
- https://www.mfa.gov.cn/ce/cefi/eng/kxjs/P020171025789108009001.pdf
- http://www.robotics-openletter.eu/
- https://news.sky.com/story/tesla-accident-leaves-one-dead-and-20-injured-in-paris-prompting-taxi-firm-to-suspend-use-of-model-12496220#
- https://www.teslaclub.it/la-responsabilita-dell-incidente-mortale-con-il-pilota-automatico-tesla.html
- https://www.repubblica.it/tecnologia/2019/11/21/news/uber_auto_a_guida_autonoma_responsabilita_umana_e_del_software_per_l_incidente_mortale-241593528/
- https://www.hdmotori.it/tesla/articoli/n516697/tesla-model-x-problemi-autopilot-incidente-2018/
- https://www.ranker.com/list/self-driving-car-accidents/eric-vega
- https://www.sae.org/standards/content/j3016_202104/
- https://www.moralmachine.net