

Studi di diritto penale costituzionale ed europeo

Diretta da R. Bartoli, S. Manacorda, V. Manes
M. Pelissero, C. Piergallini, L. Riscato
A. Vallini, F. Viganò

Sezione I
Monografie

Massimiliano Lanzi

Self-driving cars **e responsabilità penale**

La gestione del “rischio stradale”
nell’era dell’intelligenza artificiale



Giappichelli

INTRODUZIONE

DAL “BOCCONE AVVELENATO” A UNA NUOVA RIVOLUZIONE SCIENTIFICA DELLA MODERNITÀ

Molti sono i “grandi” del passato che hanno lasciato un segno tangibile e duraturo sul corso della storia, contribuendo a plasmare il mondo per come successivamente conosciuto (e prima che qualcun altro, inevitabilmente, intervenga per mutarlo di nuovo). Proprio per questo, del resto, sono definiti “grandi”.

Pochi, tuttavia, sono coloro ai quali – almeno nella prospettiva di chi viene dopo – possono essere ricondotte non una bensì plurime impronte sull’evoluzione dei tempi. Uno di questi, per chi è vissuto tra il XX e il XXI secolo, è senza dubbio Alan Mathison Turing.

Anzitutto, egli è stato, in ragione del ruolo chiave ricoperto nel programma di decrittazione delle comunicazioni dell’Asse, tra coloro che – anche per come ricordato dalla cinematografia contemporanea¹ – hanno offerto il contributo individuale più rilevante ai fini del successo del mondo libero sulla barbarie nazista².

¹ Il riferimento è alla pellicola del 2014 *The Imitation Game*, regia di Morten Tyldum.

² Vedi a tale proposito J. KEEGAN, *La seconda guerra mondiale*, BUR, 2015. In questa prospettiva, nella copiosa bibliografia che si è occupata dell’avventura dei *code breakers* di *Bletchley Parks* nel corso del secondo conflitto mondiale, cfr. T. DOWNING, *Churchill’s war lab – Code Breakers, Boffins and Innovators: The Mavericks Churchill Led to Victory*, Little Brown, 2010, pp. 161-162, ove si ricostruisce l’ideazione e la successiva realizzazione da parte di Turing di *Colossus*, «*the world’s first operational computer*», in grado di elaborare le decine di milioni di computazioni necessarie per la decodifica delle più complesse comunicazioni tedesche. Un risultato, peraltro, foriero di rilevanti conseguenze anche per le vicende italiane in quel terribile conflitto, capace soprattutto di tenere in scacco la Regia Marina per l’intera durata della guerra nel Mediterraneo. Cfr. J. WINTON, *Ultra at Sea: How Breaking the Nazi Code Affected Allied Naval Strategy During World War II*, W. Morrow ed., 1988; A. PETACCO, *Le battaglie navali nel mediterraneo nella Seconda guerra mondiale*, Mondadori, 1995. Vedi tuttavia, *contra*, E. CERNUSCHI, «*Ultra*» – *La fine di un mito*, Mursia, 2014, il

Il vincente progetto di decrittazione svolto a *Bletchery Park*, tuttavia, è stato solo un episodio del ben più vasto percorso di ricerca compiuto dal brillante accademico britannico, il quale è riconosciuto altresì, in ambito scientifico, come uno dei padri dell'informatica e dell'intelligenza artificiale³. Di qui, la seconda, e altrettanto decisiva, impronta di quest'uomo sulla realtà del mondo di oggi.

In uno storico articolo pubblicato nel 1950⁴, infatti, Alan Turing esponeva le ragioni per le quali ben si poteva immaginare una macchina alla stregua di un soggetto “pensante” e come ciò – nelle previsioni dell'autore – sarebbe entrato nella comune consapevolezza entro la fine del secolo⁵. Seguendo un coinvolgente procedimento argomentativo, Turing sosteneva la propria teoria affermando, tuttavia, di non essere in grado, in positivo, di dimostrarla, proponendosi piuttosto di confutare i diversi argomenti che gli “scettici” del pensiero artificiale – in allora come oggi, del resto – utilizzavano per negare valore pensante alle macchine⁶.

quale ridimensiona, in parte, l'importanza delle decrittazioni britanniche rispetto agli accadimenti nel teatro mediterraneo.

³ Vedi a tale proposito gli interessanti contributi in S. BARRY COOPER, J. VAN LEEUWEN (a cura di), *Alan Turing: His Work and Impact*, Elsevier Science, 2019, *passim*.

⁴ Ci si riferisce alla celebre pubblicazione: A. TURING, *Computer Machinery and Intelligence*, in *Mind*, 1950, pp. 433-460.

⁵ Sul punto, A. TURING, *Computer Machinery and Intelligence*, cit., si rappresenta come appunto «*at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted*». Il “test di Turing”, descritto nella suddetta pubblicazione scientifica, consiste nel comunicare contemporaneamente con un computer e con un essere umano, senza sapere chi è chi, formulando domande e persino facendo giochi. Se al termine della conversazione l'agente non è in grado di dire chi sia l'uomo e chi la macchina, quest'ultima ha “passato il test” e noi dovremmo trattarla come se possedesse davvero una mente.

⁶ Diversi sono gli ambiti su cui si avventurava Turing in questo percorso di confutazione: (1) l'obiezione “teleologica”, per la quale il pensiero sarebbe un precipitato dell'immortalità divina dell'animo umano («*thinking is a function of man's immortal soul*»); (2) l'obiezione della “testa nella sabbia”, per la quale lo sviluppo di macchine pensanti avrebbe delle conseguenze totalmente negative, per cui occorrerebbe limitarsi a sperare che ciò non avvenga mai; (3) l'obiezione matematica, per la quale vi sarebbe evidenza matematica, appunto, dei limiti delle capacità di calcolo di una macchina; (4) l'obiezione (o “l'argomento”) della coscienza, per il quale una genuina capacità pensante comprenderebbe non solo la capacità di scrivere qualcosa, ma altresì la consapevolezza di averla scritta, nonché la capacità di esprimere una vasta gamma di emozioni e stati d'animo che sono appannaggio esclusivo dell'uomo; (5) l'obiezione delle “varie disabilità”, per la quale ciascuno può pensare ad almeno una cosa che, secondo la propria esperienza, una macchina non potrebbe mai essere in grado di fare; (6) l'obiezione ch'egli definiva “di Lady Lovelace” («*Lady Lovelace's objection*»), per la quale una macchina non avrebbe la propensione a creare nulla, quanto solamente di ordina-

Una storia densa di significato, quella di Alan Turing: eroe di guerra e maestro nelle scienze, oggi ricordato, altresì, come vittima-simbolo delle discriminazioni e dell'intolleranza del suo tempo⁷. Una storia che – almeno nell'immaginario collettivo⁸ – sopravviverebbe nel simbolo della “mela morsicata”, tra i più iconici del progresso social-tecnologico della contemporaneità e, in fondo, ennesimo lascito del grande matematico inglese.

Diverse decadi dopo la morte di Alan Turing, lo scenario che si presenta all'osservatore è quello di un mondo in via di profondo cambiamento nella direzione dell'intelligenza artificiale. Macchine dotate di grande capacità di calcolo che, alimentate in maniera pressoché automatica con enormi flussi di dati, svolgono, grazie a sofisticati algoritmi, compiti di fondamentale e crescente importanza sul piano economico e sociale; tanto da comportare, nel tempo, una sostanziale ridefinizione nella diffusione e allocazione del potere⁹.

re, comportando in questo modo l'incapacità della macchina di sorprenderci («*a machine can never “take us by surprise”*»); (7) l'obiezione della *continuità* del sistema nervoso umano, il quale funzionerebbe in modo radicalmente differente da una macchina avente invece natura *discreta*; (8) l'obiezione per la quale non sarebbe possibile fissare ex ante il comportamento che una persona dovrebbe tenere in ogni possibile situazione che gli possa capitare nel mondo reale – e si tratta di una questione della massima rilevanza nel campo dei veicoli a guida autonoma, con riferimento ai quali l'intelligenza artificiale dovrebbe proprio elaborare di continuo scelte di comportamento ritagliate, come vedremo, sul codice della strada e su un generale onere di diligenza umano; (9) l'obiezione per la quale la realtà umana è condizionata da una dimensione di percezioni extrasensoriali, che pure ad oggi sfuggono a ogni legge scientifica di riferimento.

⁷ Come noto, Alan Turing venne condannato nel 1952 per *Gross Indecency* per avere intrattenuto una relazione omosessuale (caso *Regina v. Turing and Murray*) e morì di lì a breve, nel 1954, probabilmente suicida in ragione del profondo disagio causato dalla c.d. castrazione chimica alla quale era stato sottoposto in sostituzione della pena detentiva. Il ricordo della triste vicenda di Turing fu trainante per l'approvazione nel 2017 in Gran Bretagna di una legge (*Policing and Crime Act 2017*), con la quale è stata concessa l'amnistia – nel caso di Turing, postuma – a tutti i soggetti ammoniti o condannati per il compimento di atti omosessuali; tale legge infatti è chiamata, a livello informale, “*Alan Turing Law*”. Per una ricostruzione delle vicende anche personali di Turing vedi, tra le molte biografie disponibili, A. HODGES, *Alan Turing: the Enigma*, Vintage Books, 2012.

⁸ L'ipotesi secondo la quale il marchio di Apple Computer sia un tributo ad Alan Turing – alla luce della tesi per la quale lo stesso si sarebbe tolto la vita addentando una mela riempita di cianuro – non è mai stata confermata né smentita, anche se più elementi portano a ritenere che si tratti invero di una (bella) “leggenda metropolitana”: vedi *Unraveling the tale behind the Apple logo*, 7 ottobre 2011, in www.cnn.com.

⁹ Vedi K. CRAWFORD, *Atlas of AI*, Yale University Press., 2021, *Introduction*, in cui si evidenzia come l'impiego massivo delle intelligenze artificiali tenda a “catturare” il pianeta in una «*computationally legible form*» (p. 11), interrogandosi dunque su quali siano le conseguenze sociali e materiali delle Intelligenze Artificiali in una serie di settori sensibili dell'esperienza politica, quali l'educazione, la sanità, il mondo del lavoro, le comunicazioni, la giusti-

Complessi *software* “frugano” nei conti correnti per valutare il merito creditizio delle persone; raccolgono – in prospettiva, anche attraverso rilevatori biometrici e *scanner* di riconoscimento facciale – preferenze e gusti dei consumatori per sottoporre proposte commerciali mirate e per condizionarne, altresì, desideri e aspettative; o, ancora, valutano la produttività dei lavoratori; svolgono operazioni finanziarie a forte connotazione speculativa.

Si tratta quindi, in prima battuta, di strumenti *non-embodied* che operano in una dimensione dematerializzata. A questi, tuttavia, si sono aggiunti, nei tempi più recenti, sistemi di intelligenza artificiale *embodied*, dotati cioè di un corpo che attua movimenti e produce lavoro fisico¹⁰. Sono apparsi nel panorama “artificiale”, così, *robot* impiegati nel settore medico-chirurgico per interventi ad elevatissima precisione; droni armati, capaci di selezionare e ingaggiare, in modo automatico, i propri bersagli; e, per quanto di specifico interesse in questo lavoro, veicoli autonomi o automatici di tipo terrestre, navale e aereo¹¹.

Molti commentatori, a tale proposito, evidenziano come lo sviluppo delle intelligenze artificiali sia qualitativamente diverso rispetto a qualsiasi altro progresso tecnologico registrato fino ad oggi. Le macchine “intelligenti”, infatti, non si limiterebbero a fare *meglio* ciò che l'uomo o anche altre macchine tradizionali sono in grado di fare; le stesse non costituiscono, in altre parole, strumenti migliori rispetto ai precedenti.

zia. L'intelligenza artificiale, secondo altri autori, contribuirebbe al consolidamento di un «potere immenso, con caratteristiche che lo rendono particolarmente sfuggente rispetto alle tradizionali forme di controllo e limitazione; un potere accresciuto dalla possibilità di trattare, proprio attraverso l'intelligenza artificiale, una massa di dati incalcolabile e altrimenti concretamente ingestibile» (così C. CASONATO, *Potenzialità e sfide dell'intelligenza artificiale*, in *BioLaw Journal-Rivista di BioDiritto*, 2019, 1, p. 177).

¹⁰Cfr., in merito allo sviluppo della *embodied IA*, R. CALO, *Robotics and the Lessons of Cyberlaw*, in *California Law Review*, 2015, 103, p. 532 ss., in cui rappresenta le tre principali caratteristiche dei moderni *robot* che comportano sfide importanti sul piano della disciplina giuridica: «*embodiment*» (cioè l'incorporazione materiale del *software*), «*emergence*» (che l'A. definisce come il comportamento utile imprevedibile della macchina) e «*social valence*» (richiamando a tale proposito quelle teorie per le quali i *robot*, in ragione della loro capacità di “sentire” e agire nel mondo, sul piano ontologico si troverebbero “a metà strada” tra gli oggetti e gli umani); cfr. altresì, su tale tematica, P.R. DAUGHERTY, H.J. WILSON, *Human + Machine: Reimagining Work in the Age of AI*, Harvard Business School, 2018.

¹¹Cfr. a tale riguardo la definizione di «veicolo autonomo» offerta in G.F. SIMONINI, *L'intelligenza artificiale guida le nostre vetture*, Mucchi, 2018, p. 13, per cui «per “veicoli autonomi” si intende, in una massima estensione della definizione, un veicolo capace, attraverso la propria strumentazione, governata da un'intelligenza artificiale (IA), di trasferirsi da un punto di partenza ad uno di arrivo, fissati dal suo utilizzatore, senza alcun aiuto da parte di questo (inteso come persona umana o infrastruttura) a bordo del mezzo [...]».

Esse sono piuttosto un “mondo”, le cui potenzialità devono essere ancora apprese appieno ma che, già adesso, si presenta come rivoluzionario in senso sociale, culturale e, affermano alcuni, persino evolutivistico¹²; e di cui sono evidenti, agli esperti, non solo i vantaggi ma altresì alcuni significativi rischi¹³. Il mondo digitale, in altre parole, rappresenterebbe un vero e proprio cambio di paradigma «al fondo del quale potrebbe esserci una nuova idea (e una nuova esperienza) dell’umanità»¹⁴.

Da questo punto di vista, sono di tutta evidenza i grandi passi in avanti compiuti, proprio grazie ai sistemi di intelligenza artificiale e di *machine learning*, dalla scienza robotica. Questa nasceva, infatti, in tempi recenti – negli anni ’20 – e riguardava strumenti meccanici concepiti per svolgere specifici compiti e lavorazioni¹⁵. I moderni sistemi robotici tendono verso capacità generali, avvicinandosi sempre più, sul

¹²In questi termini le considerazioni di C. ACCOTO, *Il mondo ex machina: cinque brevi lezioni di filosofia dell’automazione*, Egea, 2019, pp. 6-7, il quale si riferisce all’automazione «non come meccanismo neo-meccanico, ma come dinamismo neo-ecologico», alla luce del quale «il mondo è avviato [...] ad una trasformazione profonda e irreversibile delle proprie fondamenta ontologiche e ontogenetiche».

¹³Vedi a questo proposito il considerando (15) della bozza di Regolamento sull’armonizzazione in materia di intelligenza artificiale (c.d. *AI Act*), COM(2021) 206 final, 21 aprile 2021, per cui «l’intelligenza artificiale presenta, accanto a molti utilizzi benefici, la possibilità di essere utilizzata impropriamente e di fornire strumenti nuovi e potenti per pratiche di manipolazione, sfruttamento e controllo sociale», in aperto contrasto con i valori dell’Unione, relativi al rispetto «della dignità umana, della libertà, dell’uguaglianza, della democrazia e dello Stato di diritto e dei diritti fondamentali dell’Unione».

¹⁴Così A. D’ALOIA, *Il diritto verso il “mondo nuovo”. Le sfide dell’intelligenza artificiale*, in A. D’ALOIA (a cura di), *Intelligenza artificiale e diritto*, FrancoAngeli, 2020, p. 15, il quale cita, a tale proposito, il *Draft Ethics Guidelines for Trustworthy AI*, predisposto dall’*High Level Expert Group on Artificial Intelligence* della Commissione Europea del 18 dicembre 2018, per il quale «*Artificial Intelligence is one of the most transformative forces of our time, and is bound to alter the fabric of society*». Da sottolineare, a tale riguardo, anche le considerazioni di C. BURCHARD, *L’intelligenza algoritmica come fine del diritto penale? Sulla trasformazione algoritmica della società*, in *Riv. it. dir. proc. pen.*, 2019, 4, p. 1908 ss., il quale invita a non cadere nella tentazione di facili banalizzazioni sociologico-criminali dell’IA (tra cui anche una paventata carenza di diversità qualitativa tra lo sviluppo dell’IA e le altre forme di progresso tecnologico registrate in passato), le quali – come in fondo, riteniamo, tutti i processi di banalizzazione di questioni, all’opposto, complesse – suonano «accattivanti e rassicuranti».

¹⁵La paternità del termine «*robot*» è attribuita al drammaturgo ceco Karel Capek, il quale lo utilizza nel suo dramma in tre atti «*R.U.R. (Rossumovi univerzátní roboti)*» pubblicato nel 1920 e messo in scena la prima volta al Teatro Nazionale di Praga nel 1925. Il neologismo deriva da «*robota*», che in lingua ceca significa “lavoro forzato”. Vedi R. CIGOLANI, D. ANDRESCIANI, *Robot, macchine intelligenti e sistemi autonomi*, in G. ALPA (a cura di), *Diritto e intelligenza artificiale*, Pacini Giuridica, 2020, p. 24.

piano qualitativo della capacità di elaborazione dei dati e delle potenzialità comportamentali, agli organismi biologici complessi¹⁶.

Si impone così agli osservatori di tale *trend* di sviluppo una realtà robotica i cui riferimenti culturali e sociali sono al confine con la *fiction*, anche quella letteraria e cinematografica. Se si sta parlando, del resto, di una “rivoluzione” tecnologica e scientifica, cioè di qualcosa di nuovo che irrompe sulla scena promettendo di spazzare via categorie e concetti noti¹⁷, non può stupire che i primi riferimenti provengano proprio da coloro che si occupano di “pensare” il mondo di domani, senza vincoli e limiti dettati da ciò che la scienza e la società oggi consentono.

Caso emblematico, sotto questo punto di vista, è rappresentato dalle “leggi di Asimov”¹⁸. Da soggetto *cult* per gli appassionati di fantascienza, le famose leggi della robotica sono diventate un esplicito riferimento normativo contemporaneo, degno di trovare riconoscimento in uno dei “considerando” della Risoluzione del Parlamento Europeo del 16 febbraio 2017, in tema di «Norme di diritto civile sulla robotica»¹⁹.

¹⁶I moderni studi sulle neuroscienze, del resto, rappresentano (i) come ogni organismo sia un algoritmo, compreso *Homo sapiens*, il quale è un insieme di algoritmi organici modellati da milioni di anni di evoluzione; (ii) che i calcoli algoritmici non dipendano dai materiali con cui sono costruiti i calcolatori e che pertanto (iii) non c'è ragione di pensare che gli algoritmi organici possano fare cose che gli algoritmi non organici non siano in grado di fare (vedi Y.N. HARARI, *Homo Deus. Breve storia del futuro*, Bompiani, 2015, p. 389). È di tutta evidenza che maggiore è la convinzione che la coscienza umana sia limitata a un insieme di calcoli algoritmici realizzati mediante impulsi biochimici, minore è la possibilità di differenziare gli organismi organici da quelli non organici, una volta che questi avranno raggiunto una analoga capacità e sofisticatezza di calcolo.

¹⁷Sulla portata epocale dello sviluppo delle neuroscienze e della ricostruzione dei *pattern* del pensiero umano, da una parte, e della capacità di raccolta, elaborazione e gestione di *meta data* da parte di algoritmi non organici, dall'altra, vedi diffusamente K. KELLY, *Quello che vuole la tecnologia*, Codice, 2011, ove si traccia la prospettiva del c.d. datismo, per il quale l'universo consiste in flussi di dati e che il valore di ciascun fenomeno o entità è determinato dal suo contributo all'elaborazione degli stessi.

¹⁸Le famose “leggi di Asimov” vengono riportate in apertura alla fortunata raccolta di racconti di fantascienza intitolata «*I, Robot*», pubblicata per la prima volta negli Stati Uniti nel 1950. Esse si rintracciano, nella narrativa dell'A., in un fantastico «*Handbook of Robotics, 56th Edition, 2058 A.D.*», e sono le seguenti: (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 un 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. Quanto all'edizione italiana, cfr. I. ASIMOV, *Io, Robot*, Oscar Mondadori, 2021.

¹⁹Il riferimento, in particolare, è al considerando T del suddetto documento, ove si rappresenta che «considerando che le leggi di Asimov devono essere considerate come rivolte ai

Scontato dire che tale rivoluzione “epocale” sollevi questioni umanistiche di altissimo rilievo, sul piano filosofico e giuridico²⁰. Nella misura in cui gli interessi fondamentali della persona possono essere coinvolti dall'intervento di agenti “artificiali”, vengono alla luce delicate criticità applicative in tema di tutela dei diritti, di allocazione delle responsabilità e di etica della scienza²¹.

Con riferimento, in particolare, alla scienza penale, affiorano inevitabilmente questioni di grande rilievo che pongono problematiche in parte del tutto nuove, in parte preesistenti, ma che le nuove tecnologie enfatizzano in maniera significativa e potenzialmente “dirompente”.

Sul piano dei *nova*, si manifesta l'esigenza di confrontarsi con agenti artificiali dotati di autonome capacità di azione e decisionali e – considerati i sistemi propri di auto-apprendimento – in rapporto di almeno parziale imprevedibilità rispetto al “creatore” umano e alle sue determinazioni.

Ecco, allora, che, in questa prospettiva, prende forma l'eventualità di ripensare l'antropocentrismo tipico dello *ius criminale*²². Allo stesso tempo, l'affermazione delle intelligenze artificiali – almeno fino a che non ne sarà disciplinato o “scoperto” un completo statuto giuridico – non fa che aumentare quella *grammatica dell'incertezza* tipica della post-modernità e dell'attuale “società del rischio”²³, rispetto alla quale

progettisti, ai fabbricanti e agli utilizzatori di robot, compresi i robot con capacità di autonomia e di autoapprendimento integrate, dal momento che tali leggi non possono essere convertite in codice macchina».

²⁰Confronta a tale proposito A. D'ALOIA, *Il diritto verso il “mondo nuovo”. Le sfide dell'intelligenza artificiale*, cit., pp. 64-65, ove si rappresenta come vi sia bisogno «di un lavoro progressivo di sperimentazione e implementazione (sociale, economica, culturale, etica, giuridica) di questa nuova contestualità rappresentata dalla coesistenza tra esseri umani e sistemi agenti artificiali».

²¹Cfr. L. D'AVACK, *La rivoluzione tecnologica e la nuova era digitale: problemi etici*, in U. RUFFOLO (a cura di), *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Giuffrè, 2020, p. 6, ove si sottolinea come il potere immenso riconducibile alle intelligenze artificiali metta a rischio la stessa centralità dell'uomo, il quale anzi rischia di vedersi declassato «da attore principale a comparsa nello scenario della vita civile».

²²Vedi V. MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrasia*, in U. RUFFOLO (a cura di), *Intelligenza artificiale. Il diritto, i diritti e l'etica*, cit., p. 547 ss. e in *DisCrimen*, 15 maggio 2020, p. 2, per il quale «il diritto penale è pensato ed edificato sull'uomo, sul rimprovero personale e colpevole, sul grado di responsabilità e di rimproverabilità per una azione umana».

²³Vedi a tale proposito, pur nella ampissima produzione scientifica, umanistica e giuridica, che individua nell'incertezza dei rapporti e degli strumenti di regolazione sociale uno dei caratteri più qualificanti del mondo post-moderno, G. ALPA, *La certezza del diritto nell'età dell'incertezza*, ESI, 2006; P. GROSSI, *L'invenzione del diritto*, Laterza, 2017, p. 90 ss., il

il diritto penale – articolato com'è, sul piano fondativo, per confrontarsi con *fatti* e non con mere *possibilità* – già appare vetusto e poco attrezzato²⁴; e ciò, in un contesto in cui alcune sue aree di intervento fondamentali sono e saranno a breve chiamate a sforzi performativi ancora maggiori.

Emblematica, in questo senso, l'area in cui opera la colpa penale: già protagonista di una vera e propria “esplosione” dovuta all'aumento quantitativo e qualitativo di attività umane rischiose e regolamentate, si trova ad affrontare un'ulteriore espansione conseguente all'affermazione di queste nuove forme di meccanizzazione e di massificazione dei processi sociali ed economici²⁵.

E si tratta a ben vedere di criticità che, con il passare del tempo, non potranno mai scemare o venire meno: al contrario, esse saranno acuite dalla continua evoluzione delle tecnologie e delle potenzialità di impiego delle IA.

La rivoluzione digitale, infatti, è una via a senso unico, che in molti considerano ormai irreversibile o – mutuando una terminologia affine all'ambito tecnico che stiamo trattando – non disattivabile²⁶. L'evoluzione del *machine learning*, ovvero di macchine in grado di apprendere ricorrendo a dati aggiuntivi e a nuovi *pattern* derivanti anche dalla propria esperienza, prendendo quindi scelte di comportamento che

quale rappresenta come il giurista contemporaneo si debba confrontare con un «tempo di disseminate incertezze, che soffriamo sulla nostra pelle e che siamo pronti a confrontare con le ferme certezze della modernità giuridica. La quale era un'epoca di certezze» (p. 91).

²⁴ Cfr., a tale riguardo, C. PIERGALLINI, *Attività produttive, decisioni in stato di incertezza e diritto penale*, in M. DONINI, M. PAVARINI (a cura di), *Sicurezza e diritto penale*, Bononia University Press, 2011, p. 327 ss.

²⁵ In merito all'aumento del campo di intervento della responsabilità penale colposa – originariamente considerata un'eccezione rispetto alla centralità dell'imputazione dolosa – vedi G. MARINUCCI, *La responsabilità colposa: teoria e prassi*, in M. DONINI, R. ORLANDI (a cura di), *Reato colposo e modelli di responsabilità*, Bononia University Press, 2013, p. 47 ss., ove l'A. evidenziava – prendendo le mosse dalle osservazioni di Engisch, il quale, nella classica monografia del 1930 *Untersuchungen Uber Vorsatz und Fahrlässigkeit im Strafrecht*, definiva i reati colposi una «felice rarità» – che «solo quando si impiegano forze umane le conseguenze di piccole imprudenze o negligenze sono del tutto insignificanti. Da quando, invece, l'uomo ha moltiplicato le proprie forze con le forze della tecnica e della natura, la funzione di controllo e di freno svolta dalle «forze mentali incoscienti» [cui faceva riferimento Engisch, *nda*] è stata pressoché vanificata. Anche la minima negligenza o imprudenza può infatti scatenare eventi lesivi a volte disastrosi».

²⁶ Vedi, in questo senso, L. ALEXANDRE, *La guerra delle intelligenze. Intelligenza artificiale “contro” intelligenza umana*, EDT, 2018, p. 295, per il quale «l'IA ormai non è più un'opzione disattivabile, un interruttore che avremo ancora la possibilità di spegnere. È diventata indispensabile».

non si limitano ad attualizzare la volontà originaria del programmatore umano²⁷, rende ancora più critico lo scenario e impone di rispondere a numerosi interrogativi che, ad oggi, trovano con fatica adeguati riscontri.

Come deve modellarsi l’ordinamento rispetto ai mutamenti “epocali” sopra descritti? È sufficiente adattare categorie e istituti esistenti a tecnologie più complesse e sofisticate? O forse occorre considerare nuove strade e soluzioni, “rifondando” l’ordinamento secondo coordinate (almeno in parte) qualitativamente diverse rispetto al passato?

Il settore dei “veicoli autonomi” sembra, quindi, rappresentare un ambito in cui si percepisce, con maggiore urgenza, la necessità di fornire risposta a tali quesiti, perché l’automazione dei trasporti – già nel breve-medio termine – coinvolgerà direttamente miliardi di persone in tutto il mondo. E il diritto penale rientra, di certo, tra le materie che maggiormente necessitano di un profondo impegno in questo senso, in virtù del crescente interesse che i veicoli autonomi hanno destato agli occhi della scienza criminale, dalla *cybersecurity* alla sicurezza stradale.

Inoltre, sul piano contenutistico, lo sviluppo dei sistemi di intelligenza artificiale entra in tensione con alcune categorie fondamentali della materia. Il progressivo allontanamento della macchina dal controllo umano, infatti, mette in crisi i tradizionali criteri di ascrizione e allocazione delle responsabilità. Parimenti, la caratteristica di costante aggiornamento e affinamento delle capacità di elaborazione – tipica delle intelligenze artificiali – rende difficile, se non impossibile, l’individuazione *ex ante* di regole di comportamento umane utili ai fini dell’accertamento della colpa; ciò comporta un *vulnus* non solo in termini

²⁷ Si richiama, a tale proposito, H. SURDEN, *Machine Learning and Law*, in *Washington Law Review*, 2014, 89, p. 87 ss., ad avviso del quale «*machine learning involves computer algorithms that have the ability to “learn” or improve in performance over time on some task*»; ID., *Artificial Intelligence and Law: an Overview*, in *Georgia State University Law Review*, 2019, vol. 35, pp. 1311-1316, ove si spiega il funzionamento dei sistemi *machine learning* facendo l’esempio dei *software* di funzionamento della posta indesiderata (*spam*), ampiamente diffusi in tutti i programmi di posta elettronica. Il *software* stesso, infatti, elabora dei criteri per identificare, su base probabilistica, le mail *spam*, sulla base di dati nuovi e sopravvenuti rispetto alla programmazione iniziale: ad esempio, elaborando i contenuti delle mail che l’utente sceglie di catalogare, nel tempo, come *spam*. Il programma è in grado di riconoscere eventuali parole o frasi ricorrenti nelle mail scartate, e poter quindi adattarsi a queste preferenze dell’utente e catalogare autonomamente come *spam* future *email* aventi contenuto affine: ciò dimostrerebbe «*how software can learn a useful pattern on its own without having a programmer explicitly program that pattern head of time*». Ciò tuttavia, precisa l’A., non a fronte di un ragionamento analogo a quello dell’utente umano e inerente al concetto e al significato di quelle parole, quanto su un mero calcolo euristico di natura probabilistica.

di personalità e colpevolezza ma, altresì, sul piano del rispetto della legalità.

Problemi ai quali i legislatori, sia nazionali sia sovranazionali, non hanno fornito, ancora oggi, alcuna appagante spiegazione, ma che necessitano di una pronta soluzione affinché le intelligenze artificiali possano essere sviluppate e utilizzate in modo socialmente e giuridicamente sostenibile, così da comportate quei concreti benefici effetti che la scienza si aspetta e preannuncia.

PARTE I

**ASSETTO E DINAMICHE EVOLUTIVE
DEI VEICOLI AUTONOMI: PROFILI
DEFINITORI E ITINERARI GEOGRAFICI**

CAPITOLO I

INTELLIGENZA ARTIFICIALE E MOBILITÀ

SOMMARIO: 1. La scienza giuridica ai tempi dell'intelligenza artificiale. – 1.1. Il problema definitorio dell'intelligenza artificiale. – 1.2. La normativa in tema di *artificial intelligence*, tra lacune legislative e *soft law*. – 1.3. “Precarietà” dell'indagine scientifica sull'intelligenza artificiale. – 2. I benefici sociali offerti dai veicoli autonomi. – 2.1. *Self-driving cars* e sicurezza. – 2.2. Mezzi intelligenti ed efficienza dei trasporti: verso lo *zero net emission*. – 2.3. Una nuova mobilità *smart*. – 3. Coordinate tecnologiche in tema di *self-driving cars*: un'automazione “per gradi”. – 4. Etica e precauzione nell'era dei veicoli autonomi.

1. *La scienza giuridica ai tempi dell'intelligenza artificiale.*

1.1. *Il problema definitorio dell'intelligenza artificiale.*

Prima di avviare un'indagine sulle possibili ripercussioni penalistiche derivanti dall'impiego dei sistemi “intelligenti” di guida autonoma, appare utile evidenziare come la ricerca giuridica, nel più ampio settore dell'intelligenza artificiale, sia afflitta da un “peccato originale” che ne condiziona, in maniera significativa, lo sviluppo.

Giova rilevare, infatti, nell'elaborazione di una teoria organica dei profili giuridici legati all'utilizzo di “macchine pensanti”, la carenza di una definizione univoca di “intelligenza artificiale”.

A tale proposito, i primi e più risalenti esercizi definitori in materia ne privilegiavano una prospettiva “antropocentrica”, tale per cui possono definirsi “intelligenti” solo quelle macchine capaci di compiere attività che, qualora messe in atto da esseri umani, sarebbero qualificabili – appunto – come “intelligenti”¹. In altri termini, si riconosce “in-

¹ In questi termini, su tutti, si esprimeva nel 1956 lo scienziato John McCarthy, introducendo un *workshop* svoltosi negli Stati Uniti, al Dartmouth College nel New Hampshire,

telligenza” agli strumenti in grado di svolgere attività di prerogativa esclusiva dell’essere umano².

Tracce di questa visione possono rinvenirsi nella nozione offerta dalla Commissione europea per l’efficienza della giustizia (CEPEJ), istituita presso il Consiglio d’Europa, la quale descrive l’intelligenza artificiale come un «insieme di metodi scientifici, teorie e tecniche finalizzate a riprodurre, mediante le macchine, le capacità cognitive degli esseri umani»³.

Tuttavia – secondo una impostazione diversa – si tratterebbe questa di una categorizzazione alquanto imprecisa, dal momento che, ad oggi, non esiste un metodo affidabile e universalmente riconosciuto, in grado di “pesare” la stessa intelligenza umana. L’essenza dell’intelligenza artificiale risiederebbe, piuttosto, nella «capacità di fare generalizzazioni appropriate in modo tempestivo e su una base dati limitata»⁴.

In tempi più recenti, un utile tentativo definitorio delle macchine a intelligenza artificiale proviene dal *High-Level Expert Group on Artificial Intelligence*, nominato dalla Commissione Europea, il quale ravvisa la caratteristica fondamentale dei sistemi “intelligenti” nella loro capacità, attraverso l’utilizzo di processi simbolici o di modelli matematici,

divenuto poi celebre quale uno dei momenti fondanti della materia: si legge in particolare, nella dichiarazione di apertura del congresso scientifico, come l’obiettivo fosse «to find how to make machines use language, from abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves [...]. For the present purpose the artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving». Sul punto, si rinvia a J. MOOR, *The Dartmouth College Artificial Intelligence Conference: the Next Fifty Years*, in *AI Magazine*, 2006, 27, 4, pp. 87-89; J. MCCARTHY, M.L. MINSKY, N. ROCHESTER, C. SHANNON, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, August 31, 1955, in *AI Magazine*, 27(4), 2006, pp. 12-14. Per una disamina della figura di John McCarthy, cfr. M. CHILDS, *John McCarthy: Computer scientist known as the father of AI*, in *www.independent.co.uk*, 1° novembre 2011.

² Invero, secondo F. AMIGONI, V. SCHIAFFONATI, M. SOMALVICO, Voce *Intelligenza Artificiale*, in *Enciclopedia della Scienza e della Tecnica Treccani*, 2008, «l’IA studia i fondamenti teorici, le metodologie e le tecniche che consentono di progettare sistemi hardware e sistemi di programmi software atti a fornire all’elaboratore elettronico prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell’intelligenza umana».

³ Cfr. CEPEJ, *Carta etica europea sull’utilizzo dell’intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, Appendice III, Strasburgo, 4-5 dicembre 2018, p. 47, consultabile sul sito istituzionale della CEPEJ, www.coe.int/en/web/cepej.

⁴ Così, J. KAPLAN, *Intelligenza artificiale. Guida al futuro prossimo*, Luiss University Press, 2017, p. 21, il quale aggiunge, a tale proposito, che «tanto più vasto è il campo di applicazione, tanto più rapidamente vengono tratte le conclusioni con informazioni minime, tanto più intelligente è il comportamento osservato».

di analizzare il contesto, adattandovi le “azioni” corrispondenti⁵.

Quale che sia la prospettiva accolta, però, le plurime definizioni di AI contengono un nucleo concettuale comune: la capacità dei sistemi ad “intelligenza artificiale” di immagazzinare ed elaborare una grandissima quantità di dati (siano essi immessi dall’uomo o siano essi raccolti direttamente dalla macchina) consente loro di assumere comportamenti dotati di un livello gradatamente alto di “autonomia” rispetto al controllo umano esercitabile a monte. Il che rende le macchine “intelligenti” sempre meno *strumenti* nelle mani di un programmatore o controllore umano, e sempre più *agenti* autonomi e distinti da quest’ultimi⁶.

Un’ulteriore prospettiva comune ai numerosi tentativi definitivi è rappresentata, altresì, dalla nota distinzione tra *standard* “forte” e *standard* “debole” di AI⁷.

La visione “forte” (*strong AI*) riconosce, in capo alle macchine, una

⁵ Così, il testo di *A definition of AI: Main capabilities and Disciplines*, Brussels, 8 aprile 2019: «*software (and possibly hardware) systems designated by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions*». Vedi V. FRANCESCHELLI, *Sull’intelligenza artificiale (IA o AI)*, in *Riv. dir. ind.*, 2023, 1, § 2, il quale sottolinea come, attribuendo all’IA sole capacità *analitiche* le si riconoscano, in sostanza, un terzo delle funzioni cognitive dell’intelligenza umana; che a queste unisce, infatti, una componente *creativa* e una *pratica*.

⁶ La letteratura di riferimento, a tale proposito, è vastissima. Si richiamano tra tutti, in questa sede, C. CASONATO, *Per una intelligenza artificiale costituzionalmente orientata*, in A. D’ALOIA (a cura di), *Intelligenza artificiale e diritto*, FrancoAngeli, 2020, p. 133 ss., il quale individua come tratti comuni delle varie definizioni di IA la capacità della macchina di (i) processare grandi quantità di dati, di (ii) percepire e cogliere alcune caratteristiche della realtà nella quale è inserita, di (iii) prendere decisioni in forma variabilmente autonoma e imprevedibile, e infine di (iv) modificare il proprio funzionamento al fine di migliorare le proprie prestazioni, imparando da sé (vero e proprio *machine learning*).

⁷ Tale distinzione viene fatta risalire, tradizionalmente, a J.R. SEARLE, *Minds, Brains, and Programs*, in *The Behavioral and brain sciences*, n. 3, 1980, pp. 417-457, per cui «*I find it useful to distinguish what I will call “strong” AI from “weak” or “cautious” AI (Artificial Intelligence). According to weak AI, the principal value of the computer in the study of the mind is that it gives us a very powerful tool. For example, it enables us to formulate and test hypotheses in a more rigorous and precise fashion. But according to strong AI, the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to understand and have other cognitive states. In strong AI, because the programmed computer has cognitive states, the programs are not mere tools that enable us to test psychological explanations; rather, the programs are themselves the explanations*».

capacità cognitiva idonea a «contestualizzare problemi specializzati di varia natura in maniera completamente autonoma»⁸. Si tratta, in altri termini, di una forma di intelligenza artificiale in grado di esibire la maggior parte delle capacità cognitive analoghe – o addirittura superiori – a quelle umane, nella prospettiva di realizzare futuri sistemi di *superintelligenza*, diretti a soppiantare l'uomo e divenire, se non adeguatamente regolati e indirizzati, un rischio esistenziale per lo stesso⁹.

La visione “debole”, al contrario, riconosce alla IA la sola natura di simulazione della mente umana, assumendo funzionalità cognitive dell'uomo senza raggiungere, tuttavia, le reali capacità intellettuali tipiche di quest'ultimo¹⁰. L'intelligenza artificiale debole persegue, infatti, l'obiettivo – più modesto – di svolgere singoli compiti, magari con un altissimo grado di precisione ed efficienza, ma pur sempre limitati rispetto all'ampio spettro delle diverse competenze umane.

Come intuibile, le suddette concezioni di intelligenza artificiale pongono, sul piano giuridico, problematiche differenti.

⁸ Così, la definizione di intelligenza artificiale “forte” offerta da Commissione europea per l'efficienza della giustizia (CEPEJ), *Carta etica per l'uso dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, 3 dicembre 2018, in <https://rm.coe.int>, p. 47, alla quale viene contrapposta la nozione di intelligenza artificiale “debole” o “moderata”, caratterizzata da «alte prestazioni nel loro ambito di addestramento». Le IA, nel loro complesso, vengono definite in tale sede quale «insieme dei metodi scientifici, teorie e tecniche finalizzate a riprodurre mediante le macchine le capacità cognitive degli esseri umani».

Nella letteratura di riferimento, lo standard forte è categorizzato da A. NEWELL, H.A. SIMON, *GPS. A Program that Simulates Human Thought*, in E.A. FEIGENBAUM, J. FELDMAN (a cura di), *Computers and Thought*, AAAI Press, 1995, p. 279.

⁹ Vedi G. SARTOR, F. LAGIOIA, *Le decisioni algoritmiche tra etica e diritto*, in U. RUFFOLO (a cura di), *L'Intelligenza artificiale. Il diritto, i diritti e l'etica*, cit., p. 66 ss., in cui gli Autori richiamano il dibattito in tema di *Superintelligence* in cui, secondo una parte degli scienziati e tecnocrati – tra cui grandi nomi quali Stephen Hawking e Bill Gates – l'indispensabile condizione per la creazione di super intelligenze è che queste «siano amichevoli verso la specie umana, rispettandone i valori». Gli stessi Autori citati del resto ritengono che i rischi legati all'emersione di una intelligenza generale artificiale siano «un problema molto serio che porrà grandi sfide in futuro» (p. 66). Sono attribuite a Stephen Hawking, in particolare, le considerazioni per le quali la capacità di auto-sviluppo delle macchine superintelligenti soppianterebbe le limitate possibilità dell'uomo, frenato dai lenti ritmi dell'evoluzione biologica (v. R. CELLAN-JONES, *Stephen Hawking warns artificial intelligence could end mankind*, in www.bbc.com, 2 dicembre 2014).

¹⁰ Cfr. J. KAPLAN, *Intelligenza artificiale*, cit., p. 97 ss., il quale – pure critico in merito alla possibilità di differenziare in maniera così netta le due “dimensioni”, forte e debole, dell'IA – rappresenta lo standard debole alla stregua di una simulazione, in cui le macchine, pur non essendo realmente “intelligenti”, sarebbero semplicemente capaci di agire «come se» lo fossero» (p. 99).

Per un verso, infatti, l'intelligenza artificiale “forte” – i cui tempi di concreto sviluppo appaiono comunque ad oggi ancora massimamente incerti¹¹ – suscita intuibili incertezze, in ragione sia della distanza, progressivamente di maggiore portata, tra gli *input* iniziali del programmatore umano e il comportamento della macchina; sia dell'esigenza di modulare su quest'ultima il concetto tradizionale di “soggettività giuridica”¹², in relazione al quale iniziano a schiudersi suadenti – quanto traballanti e, in ultima analisi, improponibili¹³ – prospettive punitive¹⁴.

¹¹In merito alle concrete prospettive di sviluppo delle superintelligenze, e alle relative conseguenze di grande portata sulla civiltà umana vedi, diffusivamente, N. BOSTROM, *Superintelligence. Tendenze, pericoli, strategie*, Bollati Boringhieri, 2018 (vers. orig. *Superintelligence*, Oxford University Press, 2014) il quale rappresenta che «se un giorno costruiremo cervelli artificiali superiori al cervello umano in quanto a intelligenza generale, questa nuova superintelligenza potrebbe diventare molto potente. Inoltre, così come oggi il destino dei gorilla dipende da noi esseri umani più che dai gorilla stessi, il destino della nostra specie dipenderebbe dalle azioni della superintelligenza artificiale». Quanto ai tempi di realizzazione di una simile tecnologia, in ogni caso, occorre essere cauti: dall'invenzione dei computer, negli anni '40 del '900, si pensava che l'avvento delle superintelligenze avrebbe richiesto una ventina d'anni; «da allora, la data di arrivo prevista si è allontanata al ritmo di un anno ogni anno; così oggi i futurologi che studiano la possibilità di un'intelligenza generale artificiale spesso credono ancora che le macchine intelligenti arriveranno tra una ventina d'anni».

¹²È il caso della *Federal Court of Australia* la quale, con sentenza del 30 luglio 2021, ha ritenuto valida la registrazione di un brevetto per invenzione che designa come inventore un'intelligenza artificiale (DABUS) e come titolare il suo proprietario (caso *Thaler v Commissioner of Patents* [2021] FCA 879). Per la Corte, l'Ufficio aveva errato nel respingere la domanda di brevetto dell'attore (dott. Tahler) in quanto la legge brevetti australiana non offre una definizione di “inventore” per la quale lo stesso debba necessariamente consistere in un essere umano, esigendo solo che *il titolare* della privativa sia un soggetto dotato di capacità giuridica. Nella pronuncia si enfatizza, in particolare, come l'intelligenza artificiale – specie quella *machine learning* – consista in collegamenti neuronali artificiali che riproducono in gran parte, nelle loro dinamiche di funzionamento, quelli umani, e come nel caso di specie il DABUS fosse una forma di «*neurocomputing that allows a machine to generate new concepts which are encoded as chained associative memories within the artificial neuronal networks*» (§ 35 della sentenza, motivazione di Justice Beach).

¹³Vedi su tale tema, più diffusivamente, *infra* Pt. II, Cap. II, § 4.7.

¹⁴In questo senso, l'interessante ricostruzione offerta da C. PIERGALLINI, *Intelligenza artificiale: da “mezzo” ad “autore” del reato?*, in R. ACQUAROLI, E. FRONZA, A. GAMBERINI (a cura di), *La giustizia penale tra ragione e prevaricazione*, Aracne, 2021, p. 235 ss., il quale apre al possibile riconoscimento, in futuro, di una soggettiva giuridica «nuova e tecnologica» alla macchina, al fine di soddisfare eventuali istanze risarcitorie e anche nella prospettiva della costituzione di un autonomo sistema punitivo. Riconoscimento che pure, secondo l'A., non potrà comportare la punibilità della macchina, intesa come l'inclusione di queste entro il perimetro del diritto penale in senso stretto. Tanto sarebbe impedito, in particolare, dalla circostanza che, per quanto elaborata sia la capacità di elaborazione del *robot* e per quanto avanzata sia la sua capacità di auto-apprendimento, «le decisioni che partorisce sono una conseguenza immediata e diretta riferibile all'algoritmo che le governa, che, però, è sprovi-

Per altro verso, l'intelligenza artificiale “debole”, essendo già una realtà in molti settori, impone di trovare *adesso* le risposte a numerosi interrogativi giuridici, etici e politici. Anche siffatto utilizzo dell'intelligenza artificiale solleva, infatti, rilevanti questioni in termini di disoccupazione, diseguaglianza, esclusione sociale e discriminazioni¹⁵. La stessa comunità degli esperti del settore, tra cui alcuni nomi di grande e sicuro impatto mediatico, del resto, non ha mancato di lanciare significativi *warnings* circa i rischi che lo sviluppo delle IA comportano per l'umanità, proponendo quindi di fermarne lo sviluppo fino a quando non si sarà tracciata una adeguata cornice di sicurezza¹⁶.

Sul versante penalistico, poi, l'impiego massivo di sistemi ad auto apprendimento stimola delicati interrogativi in merito all'attribuibilità delle responsabilità per eventuali comportamenti lesivi della macchina “intelligente”, sia sul piano causale, alla luce delle numerose inferenze che si accumulano e sovrappongono rispetto ai più qualificati e isolabili interventi umani – sia sotto il profilo della colpevolezza, a cagione dell'imprevedibilità e inevitabilità delle decisioni autonome della macchina.

sto della libertà di autodeterminazione che contraddistingue l'esperienza degli esseri umani» (p. 240). Il *robot*, insomma, agirebbe sempre e comunque perché *obbligato* a farlo dagli algoritmi che lo governano, difettando in radice la possibilità di cogliere alcun reale coefficiente di colpevolezza nelle sue azioni.

¹⁵ Rispetto alle quali occorrerà verificare in che modo la volontà di un approccio *human centered*, attento cioè alla tutela dell'uomo e dei suoi diritti fondamentali, riuscirà ad avere la meglio rispetto a scopi «non allineati alle esigenze umanistiche, ma piuttosto concernenti il controllo e la manipolazione degli individui, la concentrazione delle opportunità e della ricchezza, l'uso della forza» (così G. SARTOR, F. LAGIOIA, *Le decisioni algoritmiche tra etica e diritto*, cit., pp. 67-68).

¹⁶ Vedi *Elon Musk: l'intelligenza artificiale è un rischio per l'umanità*, in *Corriere della Sera*, 29 marzo 2023. Oltre mille tra ricercatori e autori esperti di IA – tra cui nomi noti al dibattito pubblico quali, oltre ad Elon Musk, Steve Wozniak, co-fondatore di Apple, e lo storico israeliano Yval Noah Harari – hanno sottoscritto una lettera aperta («*Pause Giant AI Experiments: An Open Letter*», pubblicata *online* il 22 marzo 2023 sul sito del *Future of Life Institute*) chiedendo una pausa di sei mesi nello sviluppo di nuovi sistemi più potenti dell'attuale (al momento in cui si scrive) Gpt-4 («*We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4*»). Nel documento, in particolare, si stigmatizza come negli ultimi mesi i centri di sviluppo delle *Giant AI* abbiano dato vita a una corsa al potenziamento di sempre più potenti intelligenze artificiali «*that no one – not even their creators – can understand, predict, or reliably control*».

1.2. *La normativa in tema di artificial intelligence, tra lacune legislative e soft law.*

L'assenza di una definizione univoca di AI non è, tuttavia, l'unico ostacolo che incontra l'elaborazione di una teoria organica¹⁷ sui profili giuridici legati all'impiego di moderni sistemi di intelligenza artificiale.

La tendenza del diritto, tipica degli ordinamenti moderni, a normare ogni singolo aspetto dei rapporti sociali deve necessariamente fare i conti con la fisiologica rapidità di sviluppo dell'IA¹⁸.

Come intuibile, la celerità con cui si modifica il settore in considerazione impone altrettanta rapidità nei tempi di regolazione giuridica, i quali, tuttavia, si rivelano inconciliabili con le tempistiche proprie dell'ordinario dibattito democratico e del successivo *iter* politico di regolamentazione¹⁹. Le procedure legislative tradizionali, infatti, difficilmente consentirebbero di raggiungere gli obiettivi regolatori perseguiti da un settore in continuo mutamento, correndo piuttosto il rischio di intervenire solo in seguito all'avvenuta emersione di tensioni o problematiche cui porre rimedio²⁰, e di rivestire, così, un ruolo meramente sussidiario nella materia *de qua*.

Non solo.

Accanto all'esigenza di apprestare immediate soluzioni normative, lo sviluppo e le applicazioni dell'IA rendono necessaria, altresì, la pre-

¹⁷Fatta salva naturalmente la considerazione che, con l'aumentare della specializzazione delle macchine "intelligenti" e della eterogeneità dei loro campi di intervento, anche per queste potranno ipotizzarsi delle discipline speciali, differenziate e ritagliate sulle relative specificità.

¹⁸Vedi a tale proposito A. D'ALOIA, *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale*, in A. D'ALOIA (a cura di), *Intelligenza artificiale e diritto*, cit., p. 22 ss., per cui «il diritto è semplicemente travolto, sfidato da queste cose, e non è solo una difesa d'ufficio di questo focus».

¹⁹Questa, del resto, secondo le moderne scienze sociali, sarebbe una delle ragioni del declino della politica nel XXI secolo. Sommerso da una mole di dati gigantesca che non riesce più ad elaborare, «il governo», è stato scritto, «è diventato mera amministrazione. Gestisce il Paese, ma non ne è più alla guida» (cfr. Y.N. HARARI, *Homo deus*, cit., p. 460).

²⁰Cfr., in questo senso, le considerazioni in G. ROMANO, *Diritto, robotica e teoria dei giochi: riflessioni su una sinergia*, in G. ALPA (a cura di), *Diritto e intelligenza artificiale*, Pacini Giuridica, 2020, p. 103, il quale evidenzia anche il carattere per certi versi fisiologico di siffatta dinamica, per cui «un lasso di tempo è sempre necessario affinché emergano problematiche relazionali, patologie di nuovo ordine o lacune imprevedibili, essenziali a discernere zone d'ombra interne all'ordinamento in cui poter far poi operare la discrezionalità del giudice e, eventualmente, la mano del legislatore».

visione di regole flessibili, capaci di adattarsi più facilmente alla natura dinamica del settore in esame.

E così, l'inquadramento giuridico dell'IA appare, ad oggi, espresso, in maniera preponderante – e non solo nel contesto italiano o europeo²¹ – da strumenti di *soft law*, corrispondenti a documenti quali *guidelines, statements*, raccomandazioni e studi preparatori²².

Se si volge lo sguardo, in particolare, alla produzione della Commissione europea, è possibile riscontrare l'avvenuta pubblicazione, nel 2020, di un pacchetto di documenti di grande respiro sulle prospettive dell'Unione in tema di intelligenza artificiale²³.

Si tratta di un «libro bianco» che descrive l'approccio europeo al tema dell'intelligenza artificiale («*White Paper on Artificial Intelligence – A European Approach to Excellence and Trust*» COM(2020) 65), accompagnato da una Relazione in merito a sicurezza e responsabilità, sempre in tema di intelligenza artificiale («*Report on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things and Robotics*» (COM(2020) 64). A ciò si affiancano due comunicazioni: la prima denominata «*A European Data Strategy*», volta a promuovere l'economia agile basata sui dati, e la seconda intitolata «*Shaping Europe's Digital Future*», diretta a tracciare alcuni obiettivi chiave ((i) *Technology that works for people*; (ii) *A fair and competitive Economy*; (iii) *An Open, Democratic and Sustainable Society*) e le relative azioni di implementazione.

Si tratta, all'evidenza, di strumenti che, in quanto riconducibili alla categoria della *soft law*, appaiono privi di significativa precettività e, come tali, dotati di un mero indirizzo programmatico e politico²⁴. Ciò

²¹ Ci si riferisce, anzitutto, all'esperienza degli Stati Uniti ove, pur a fronte di un vantaggio culturale e tecnologico in materia dei veicoli autonomi, permane la carenza di specifici testi precettivi. Un primo atto ricognitivo delle sfide regolamentari di tali tecnologie, infatti, è individuabile nel *Preliminary Statement of Policy Concerning Automated Vehicles*, pubblicato nel 2013 dalla *National Highway Traffic Safety Administration* (NHTSA), autorità federale statunitense con poteri regolatori in tema di sicurezza stradale.

²² Per un'approfondita disamina delle prospettive regolatorie in tema di intelligenza artificiale, si rinvia a L. PARONA, *Prospettive europee e internazionali di regolazione dell'intelligenza artificiale tra principi etici, soft law e self-regulation*, in *Riv. reg. merc.*, 2020, 1, p. 70 ss.; E. STRADELLA, *La regolazione della robotica e dell'intelligenza artificiale: il dibattito, le proposte, le prospettive. Alcuni spunti di riflessione*, in *www.medialaws.it*, p. 73.

²³ I documenti elencati nel prosieguo del paragrafo sono liberamente consultabili sul sito internet della Commissione europea, <http://ec.europa.eu/commission>.

²⁴ Per una panoramica dell'approccio europeo alle IA cfr. M. ZANICHELLI, *Ecosistemi, opacità, autonomia: le sfide dell'intelligenza artificiale in alcune proposte recenti della Commissione europea*, in A. D'ALOIA (a cura di), *Intelligenza artificiale e diritto*, cit., p. 67 ss., la

nonostante, nei suddetti documenti si toccano problemi di grande rilievo per le tematiche in argomento, primo fra tutti l'opacità del processo decisionale e di funzionamento dell'intelligenza artificiale²⁵.

Ed è muovendo da queste basi che, nell'aprile del 2021, la Commissione europea ha finalmente pubblicato una proposta per l'adozione di un Regolamento europeo in tema di IA («*Proposal for a Regulation of the European Parliament and of the Council*» (COM(2021) 206 final)), ove si rappresenta, in particolare, la volontà politica di introdurre il minor numero possibile di norme prescrittive, alla ricerca di un efficace punto di equilibrio tra la volontà di prevenire i rischi connessi allo sviluppo delle IA, da un lato, e la necessità di non ostacolare il progresso di tali benefiche e competitive tecnologie, dall'altro.

Neppure il Parlamento europeo, dal canto suo, è rimasto inattivo sul fronte dell'IA.

Tra i documenti più recenti meritano di essere ricordate tre importanti risoluzioni, pubblicate il 20 ottobre 2020, finalizzate anch'esse a guidare, con la struttura propria delle *soft law* e delle disposizioni di indirizzo, una futura più puntuale normazione su queste tematiche.

Si tratta, in particolare, delle risoluzioni (i) sugli aspetti etici dell'intelligenza artificiale, della robotica e delle tecnologie connesse («*Fra-*

quale pone ben in evidenza quale sia il problema di fondo al quale guarda il dibattito su queste tematiche, ovverosia «la struttura sempre più sofisticata degli algoritmi basati sull'apprendimento automatico, che tende a rendere pressoché inaccessibile all'essere umano la logica sottesa al funzionamento dei sistemi» (p. 69).

²⁵ Sulla considerazione che «*moral responsibility, in whatever sense, cannot be allocated or shifted to "autonomous" technology*» (cfr. il documento del *European Group on Ethics in Science and New Technologies*, «*Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems*», 2018, p. 10). Cfr. M. ZANICHELLI, *Ecosistemi, opacità, autonomia: le sfide dell'intelligenza artificiale in alcune proposte recenti della Commissione europea*, cit., per la quale se da un lato, l'Europa programma lo stanziamento di risorse adeguate per non restare indietro nella corsa tecnologica all'implementazione delle intelligenze artificiali, dall'altro lato, essa intende qualificare la propria presenza in questi settori attraverso «l'adozione di un quadro etico-giuridico chiaro e riconoscibile, capace di affrontare l'IA in *primis* in termini di sostenibilità, di vantaggi per la società, di attendibilità e trasparenza dei processi decisionali» (p. 71). Prenderebbe corpo, in questo modo, l'ambizioso obiettivo di istituire un ecosistema digitale al cui centro porre sempre e solo l'uomo, il quale ne è quindi il controllore e insieme il fine, informando cioè l'azione della macchina a un contenuto etico di particolare spessore e qualità, espressivo di un «*meaningfull human control*» sui robot.

Vedi, a proposito del *black box* dell'attività algoritmica e delle sue ripercussioni in chiave di trasparenza anche nell'ambito dell'impiego in ambito giudiziario, O. DI GIOVINE, *Il "judge-bot" e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione giurisprudenziale)*, in *Cass. pen.*, 2020, p. 960 ss.; 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 *Arch. pen. web*, 17 maggio 2019, p. 19 ss.

mework of ethical aspects of artificial intelligence, robotics and related technologies» 2020/2012 INL), (ii) sui diritti di proprietà intellettuale per lo sviluppo delle tecnologie basate sull'intelligenza artificiale («*Intellectual property rights for the development of artificial intelligence technologies»* 2020/2015 INI) e – di maggiore interesse per la presente trattazione – (iii) sul regime della responsabilità civile in tema di intelligenza artificiale («*Civil liability regime for artificial intelligence»* 2020/2014 INL).

A queste si sono accompagnati *report* e documenti di studio sull'impiego delle tecnologie di intelligenza artificiale in alcuni settori particolarmente delicati, tra i quali – per quanto qui di interesse – il *draft report* per l'impiego delle IA nel diritto penale e il loro utilizzo da parte della polizia e dell'autorità giudiziaria nei procedimenti penali²⁶.

Il panorama così descritto documenta, nell'ambito giuridico-regolatorio, una frammentarietà di linee guida tendenzialmente settoriali e la carenza, ad oggi, di interventi di carattere organico; anche se, come si vedrà nel corso del lavoro, non mancano decisi passi in avanti nel senso di una maggiore precettività sul tema in argomento: dalla proposizione nell'aprile 2021 di una mozione per un Regolamento che stabilisca regole armonizzate sull'intelligenza artificiale, alla proposta per l'adozione di una nuova Direttiva in materia di responsabilità per danno da in-

²⁶ Il riferimento è al *Draft Report* intitolato «*Artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters»* (2020/2016 INI), in cui sono variamente declinate le tematiche – di immediata e stringente attualità – inerenti ai cosiddetti “algoritmi predittivi” e all'opportunità di sostituire o affiancare il giudice con tali strumenti “intelligenti” in alcuni passaggi particolarmente delicati dell'attività giurisdizionale: non solo nell'accertamento, ma altresì nella scelta e nella commisurazione della risposta sanzionatoria. In tema di algoritmi predittivi ed esercizio della giurisdizione, vedi V. MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in U. RUFFOLO (a cura di), *Intelligenza artificiale – il diritto, i diritti e l'etica*, cit., in cui si pone, in particolare, l'accento sui possibili attriti tra una giurisdizione “algoritmica” e meccanica rispetto a molti statuti fondamentali a presidio dell'ordinamento penale, dal principio di uguaglianza (p. 557) a quello di personalità della responsabilità penale e della sanzione (p. 559). Interessanti profili più strettamente processual-penalistici inerenti al medesimo tema dell'utilizzo di sistemi di IA nell'attività giurisdizionale in S. QUATTROCOLO, *Equo processo e sfide della società algoritmica*, in A. D'ALOIA (a cura di), *Intelligenza artificiale e diritto*, cit., p. 267 ss., dove si enfatizza la possibile violazione, tra gli altri, del diritto alla “parità delle armi” riconosciuto dall'art. 6 CEDU, alla luce della difficoltà (se non della stretta impossibilità) di confutare, contro argomentando, decisioni prese da “macchine” il cui percorso di valutazione è alquanto opaco e di cui sia impossibile per le parti accedere al “codice sorgente” (pp. 274-275). Così anche L. LUPARIA, G. FIORELLI, *Diritto probatorio e giudizi criminali ai tempi dell'intelligenza artificiale*, in R. GIORDANO, A. PANZAROLA, A. POLICE, S. PREZIOSI, M. PROTO (a cura di), *Il diritto nell'era digitale. Persona, Mercato, Amministrazione, Giustizia*, Giuffrè, 2022.

telligenza da artificiale, avanzata nell'ottobre del 2022 dalla Commissione Europea²⁷.

Le Istituzioni sono apparse, negli ultimi anni, maggiormente orientate a fornire agli operatori economici le linee guida per lo sviluppo di queste tecnologie; il cui effettivo inquadramento, inteso anche come statuto di funzionamento, rischia così di restare appannaggio dei soli tecnici, esperti di settore.

Il programma “intelligente”, infatti, «è standardizzato a grandi linee ma non è definitivo»²⁸, in quanto soggetto a continue evoluzioni e aggiornamenti che, a differenza di quanto accade per i prodotti dell'industria “tradizionale”, sfuggono a un pieno controllo regolamentare, così come a un'efficace azione preventiva di omologazione da attuarsi secondo criteri normativi e previamente fissati per via legislativa. Ed è naturale che, specie nella prospettiva del diritto penale, più sono “liquidi” i criteri e gli statuti di funzionamento tecnologico, minore è la possibilità di individuare regole e *standard* di comportamento la cui violazione possa essere legittimamente censurata e, nel caso, sanzionata.

Queste considerazioni, naturalmente, condizionano le prospettive di rischio penale riferibili alla generalità degli attori presenti sul palcoscenico del nuovo “ecosistema” delle intelligenze artificiali, giacché incidono, anzitutto, sulla censurabilità dell'attività di implementazione dei programmatori, riguardando, altresì, i produttori e i distributori di macchine “intelligenti”, i quali entrano in un settore inesplorato dalla disciplina della responsabilità da prodotto e toccano, infine, i consumatori/utenti finali, che devono interagire con una tecnologia sfuggente alle ordinarie coordinate della omologazione e, quindi, della controllabilità del rischio.

Ecco che, in ultima analisi, l'assenza di riferimenti normativi o di *standard* di condotta adeguatamente predeterminati, pone in seria crisi le categorie della prevedibilità ed evitabilità *ex ante* e, con esse, la certezza del diritto²⁹.

²⁷ In merito a questi profili e alle prospettive di intervento normativo europeo in tema di intelligenza artificiale vedi *infra* Pt. I, Cap. II, § 1.1.

²⁸ Vedi, in questi termini, M.G. LOSANO, *Verso l'auto a guida autonoma in Italia*, in *Dir. inf.*, 2019, p. 425, il quale rappresenta che «mentre l'automobile omologata è prodotta “in serie”, i prodotti informatici vengono invece venduti senza specifiche garanzie di totale conformità a un modello originario, perché si tratta di prodotti in continua evoluzione». Si avrà modo di approfondire la rilevanza di queste considerazioni in tema di responsabilità da prodotto e dei relativi criteri di allocazione.

²⁹ Sul tema della “certezza del diritto” la letteratura è sterminata. Si rinvia, per tutti, a F. LOPEZ DE OÑATE, *La certezza del diritto*, Giuffrè, 1968; P. CALAMANDREI, *La certezza del*

1.3. “Precarietà” dell’indagine scientifica sull’intelligenza artificiale.

Infine, è d’uopo evidenziare un ultimo importante aspetto, che condiziona l’indagine scientifica giuridica in materia di intelligenza artificiale.

Ogni evoluzione tecnologica non è mai un punto di arrivo, bensì un mero passaggio all’interno di un più ampio processo di sviluppo, di cui è difficile enucleare l’inizio e, in ogni caso, impossibile prevedere una “fine”. E ciò, seppur vero per ogni tipo di avanzamento scientifico, nel campo delle IA assume una rilevanza ancor più pregnante.

Una parte considerevole, infatti, dei lavori in tema di intelligenze artificiali pone in luce il carattere estremamente precario delle indagini nel settore in considerazione, giacché il progresso di simili strumenti tecnologici è talmente rapido da risultare arduo – se non impraticabile – immaginare cosa accadrà nel breve-medio corso e, ancor di più, nel lungo periodo³⁰.

Le intelligenze artificiali, infatti, dotate di sorprendenti capacità di calcolo algoritmico e di rapidità di esecuzione, si evolvono a una velocità maggiore rispetto a quella dell’apprendimento umano; al punto che, come visto, la comunità tecnologica non ha mancato di lanciare segnali di allarme, veri e propri *warnings*, proponendo una “pausa” nel loro sviluppo, in attesa di attendibili riscontri quanto alla loro controllabilità e sostenibilità³¹.

Tali considerazioni, che pure possono parere a tratti più proiezioni da *science fiction* che non realistiche prospettive di sviluppo tecnologi-

diritto e le responsabilità della dottrina, ivi, p. 167 ss.; M. CORSALE, *La certezza del diritto*, Giuffrè, 1970; L. GIANFORMAGGIO, *Certezza del diritto*, in EAD. (a cura di), *Studi sulla giustificazione giuridica*, Giappichelli, 1986, p. 157 ss.; EAD., voce *Certezza del diritto*, in *Dig. disc. priv.*, II, Utet, 1988, p. 274 ss.; C. LUZZATI, *L’interprete e il legislatore. Saggio sulla certezza del diritto*, Giuffrè, 1999.

³⁰Cfr. così, nella bibliografia penalistica più recente, C. PIERGALLINI, *Intelligenza artificiale: da “mezzo” ad “autore” del reato?*, cit., p. 242, ove sottolinea come la propria tesi per cui *machina artificialis delinquere et puniri non potest* valga «allo stato», volendo rimarcare in questo senso che «non ci è dato sapere cosa potrebbe accadere nell’apocalittico scenario prefigurato da Hawking», rispetto alle imprevedibili potenzialità evolutive di future superintelligenze artificiali ‘forti’ o ‘general’.

³¹Il riferimento è, come si è già ricordato in precedente nota, alla “lettera aperta”, pubblicata nel marzo 2023 da oltre mille ricercatori e personaggi di primo piano del settore delle IA, i quali lamentano che le intelligenze artificiali costituiscano un rischio per l’umanità e proponendo quindi una pausa di sei mesi al loro sviluppo. Vedi *Elon Musk: l’intelligenza artificiale è un rischio per l’umanità*, in *Corriere*, 29 marzo 2023.

co, portano a valutare la possibilità (o forse addirittura, la probabilità) che, già nel giro di qualche anno, molte delle coordinate tecnologiche e commerciali a cui si fa riferimento in sede di trattazione giuridica sull'IA potrebbero essere considerate superate e non più attuali.

Certo, le problematiche saranno affrontate facendo ricorso a categorie e istituti che non perderanno naturalmente di attualità al mutare delle tecnologie; il contesto, tuttavia, non potrà che essere differente rispetto al presente, e non è possibile oggi avere piena contezza del grado e dell'intensità di tali difformità.

In questa prospettiva, il diritto è posto innanzi a una sfida ancora più impegnativa.

Se, da un lato, come visto, esso (in)segue l'evoluzione tecnologica, offrendo strumenti di risoluzione solo quando i contrasti sono già emersi (e quando la politica, dal canto suo, riesce a "tenere il passo"), dall'altro lato, tuttavia, appare auspicabile che il diritto abbia la capacità di *anticipare* il processo tecnologico ed elaborare istituti giuridici idonei a governarlo sul nascere, prima cioè che si verifichino attriti o danni potenzialmente di grande portata³².

2. I benefici sociali offerti dai veicoli autonomi.

Nel vasto panorama dello sviluppo delle intelligenze artificiali – e delle relative problematiche definitorie e giuridiche – una posizione preminente è rivestita dai veicoli che utilizzano, con livelli crescenti di implementazione, sistemi di guida autonoma o automatica.

Nel prosieguo del presente lavoro ci si riferirà principalmente, in questo senso, ai «veicoli autonomi» (o «A.V.» per «*Autonomous Vehicles*»), in linea con la classificazione ampiamente in uso nella letteratura internazionale – e specie anglosassone – da qualche anno attenta e prolifica su tali tematiche³³. Si tratta naturalmente di una categoria molto

³²Così le considerazioni in G. PROIETTI, *La responsabilità nell'intelligenza artificiale e nella robotica*, Giuffrè, 2020, p. 231, il quale rappresenta «una realtà tecnologica futura in cui l'intelligenza artificiale e la robotica, con ogni probabilità, non saranno limitati a specifici domini, paventerà la necessità di rintracciare nuovi istituti giuridici da coniare già prima della sua nascita, in considerazione del dato che non può essere certa l'identificazione di un orizzonte temporale della sua genesi, e non ci si potrà permettere un complicato studio a posteriori».

³³Per un'ampia ricognizione dei profili tecnologici in argomento e degli ambiti legali maggiormente coinvolti in tali cambiamenti vedi D. GLANCY, *Autonomous and Automated*