UNIVERSITA' DEGLI STUDI DI SALERNO

DIPARTIMENTO DI SCIENZE GIURIDICHE

(SCUOLA DI GIURISPRUDENZA)

DOTTORATO DI RICERCA

IN

SCIENZE GIURIDICHE

Curriculum Giuspubblicitisco

XXXV° ciclo

Tesi di dottorato

Diritto Penale e Intelligenza artificiale

Tutor

Ch.mo Prof. Andrea R. Castaldo

Coordinatore

Ch.mo Prof. Geminello Preterossi

Candidato

Dott. Giovanni de Bernardo

INTRODUZIONE

CAPITOLO I

INTELLIGENZA ARTIFICIALE E DIRITTO

1. Premessa.
2. Informatizzazione e intelligenza artificiale
    2.1. Nascita e sviluppo dell'intelligenza artificiale
    2.2. Intelligenza artificiale forte e intelligenza artificiale debole (Strong AI VS. Weak AI)
    2.3. Principali sistemi di apprendimento: *machine-learning* e *deep learning*
3. I nuovi orizzonti della tecnologia informatica e della robotica
4. Il contesto normativo in materia di intelligenza artificiale e robotica
    4.1. Il quadro normativo europeo in materia di robotica
    4.1.1. Il cerchio interno: la 'direttiva macchine'
    4.1.2. Il cerchio maggiore ed il cerchio esterno: le disposizioni sulla sicurezza generale dei prodotti e sulla vendita dei beni di consumo
    4.1.3. Settori specifici VS. assenza di regolazione
    4.1.4. I progetti specifici relativi alla rilevanza dei robot per il diritto
5. L'informatica del diritto
6. Modelli di responsabilità in ambito penale per self-learning robot: campo di indagine
7. Intelligenza artificiale e diritto civile
8. Intelligenza artificiale e diritto penale
9. Proseguimento dell'indagine

CAPITOLO II

IA E LAW ENFORCEMENT

1. Smart city e predictive policing;
2. AI e Law Enforcement;
3. La polizia predittiva: ratio ed impiego;
    3.1. Sistemi di individuazione degli hotspots;
    3.2. Sistemi di crime linking;
4. Il valore delle informazioni raccolte tramite i tools;
5. Le criticità;

INTRODUZIONE

«Il progresso irrompe, non chiede permesso. E nel contesto attuale disegnare questo nuovo rapporto tra esseri umani e macchine non è per niente facile. Anche perché le tecnologie digitali hanno una velocità impressionante. Le tecnologie di ieri, come ad esempio la TV, la radio, l'elettricità, l'automobile hanno impiegato più di 50 anni per raggiungere i 50 milioni di utenti. Ci hanno concesso tutto il tempo per abituarci alle loro innovazioni, per avere nuove regole sul loro utilizzo, e per organizzare le nostre vite e le nostre società di conseguenza. Oggi, le tecnologie digitali irrompono molto più velocemente, e non ci danno affatto il tempo per organizzarci e per abituarci alle loro dirompenti innovazioni. Un esempio evidente di questa velocità viene dalle reti sociali: Twitter ha impiegato meno di 3 anni per raggiungere i 50 milioni di utenti; Facebook e Instagram meno di 2 anni. Anche se il record della velocità è quello di Pokemon Go, che è riuscito a raggiungere i 50 milioni di download in soli 19 giorni!»[1].

Anche il diritto penale si deve, quindi, attrezzare per tenere il passo di questa rapidissima evoluzione tecnologica, per non rischiare di soccombere di fronte a quello che si preannuncia essere un nuovo, sconvolgente «shock da modernità»[2], che comporterà problemi «analoghi a quelli che hanno contraddistinto altre "transizioni" tecnologiche: verificare l'idoneità delle norme esistenti ad applicarsi alle nuove tecnologie, così da valutare se sia opportuno, per i legislatori, coniare delle regole ad hoc, nuove, ovvero persistere, non senza possibili forzature avallate, magari, sul piano giurisprudenziale, nell'applicazione delle norme preesistenti»[3].

Se fino a pochi anni fa il principale problema di tutti gli scienziati coinvolti nella ricerca relativa all'Intelligenza Artificiale era quello di poter dimostrare la realistica possibilità di utilizzare sistemi intelligenti per usi comuni, oggi che questo obiettivo è ampiamente raggiunto ci si chiede quale possa essere il futuro per l'Intelligenza Artificiale. Soprattutto con riferimento all'applicazione dei sistemi intelligenti ed al conseguente impatto sul tessuto sociale ed economico. Diventa pertanto centrale investigare il nuovo rapporto tra

---

[1] G.F. Italiano, *Intelligenza artificiale, che errore lasciarla agli informatici*, in *Agendadigitale.eu*, 11 giugno 2019.

[2] L'espressione è di F. Stella, *Giustizia e modernità. La protezione dell'innocente e la tutela delle vittime*, Giuffré, 2003, pp. 292 ss., e con essa il grande Maestro intendeva indicare l'affannosa rincorsa del diritto penale alle evoluzioni tecnologiche che si sono succedute nei decenni passati e che, a quanto pare, interverranno anche – e forse ancor più rapidamente – nei decenni futuri.

[3] M. Bassini, L. Liguori, O. Pollicino, *Sistemi di Intelligenza Artificiale, responsabilità e* accountability. *Verso nuovi paradigmi?*, in F. Pizzetti (a cura di), *Intelligenza artificiale,* cit., p. 334.

uomo e macchina, quanto a morale e all'etica lavorativa e al corretto utilizzo delle macchine nel rispetto dell'uomo. Probabilmente la direzione che si prenderà non è ancora ben delineata, ma potrà portare a una nuova rivoluzione culturale e industriale.

Quattro sono i possibili percorsi di indagine che caratterizzerebbero il futuro penalistico dell'intelligenza artificiale: a) le attività di law enforcement e, in particolare, di polizia predittiva; b) il possibile impiego di algoritmi decisionali per risolvere vertenze penali, così da operare una sorta di sostituzione, o per lo meno di affiancamento, del giudice-uomo col giudice-macchina; c) la valutazione della pericolosità criminale affidata ad algoritmi predittivi, capaci di attingere e rielaborare quantità enormi di dati al fine di far emergere relazioni, coincidenze, correlazioni, che consentano di profilare una persona e prevederne i successivi comportamenti, anche di rilevanza penale; d) infine, lo studio di un sistema di IA quale autore di reato.

Sono tutti percorsi di straordinario interesse. E taluni di essi risultano ormai al centro del dibattito penalistico. Rimandando alla copiosa bibliografia che si è ormai sviluppata su quei temi, sono da segnalare due ulteriori prospettive di indagine concernenti l'impatto che le tecnologie informatiche e le connesse forme di intelligenza artificiale stanno avendo e potranno avere sul diritto penale.

A) La prima prospettiva prende le mosse dall'attuale processo di digitalizzazione e informatizzazione del materiale normativo: quali effetti tale processo potrà produrre sulla natura stessa delle norme giuridiche, in particolare di quelle penali, e soprattutto sulla tecnica legislativa?

B) La seconda prospettiva di indagine riguarda il possibile contributo dell'intelligenza artificiale alla formulazione della fattispecie incriminatrice.

C) La terza prospettiva si fonda sulla possibilità di alcuni software di accedere a dati riferibili ad abitudini di vita, "instabilità residenziale", tendenze sessuali, gusti commerciali, "modo di utilizzo del tempo libero", ovvero a fatti e caratteristiche in grado di comportare una valutazione prognostica sulla pericolosità, i cui confini e prospettive sono da approfondire attraverso la ricerca, quanto meno in ordine alla compatibilità con lo Stato di diritto.

D) La più delicata – sul piano filosofico ed etico, prima ancora che giuridico – è senz'altro quella concernente la possibilità di concepire le entità intelligenti come autori di reato. I sistemi di Intelligenza Artificiale di ultima generazione sono infatti dotati di un grado di autonomia dall'uomo tale da mettere in crisi il modello tradizionale della responsabilità di quest'ultimo per i fatti di reato verificatisi a causa del comportamento dell'entità di

Intelligenza Artificiale. Meno problematica è invece la questione, postasi soprattutto in previsione di un ulteriore progresso tecnologico, se sistemi di Intelligenza Artificiale possano essere assunti ad autonomi beni giuridici meritevoli di tutela penale e, per questa via, a vere e proprie vittime del reato, secondo una prospettiva sperimentata di recente con riferimento a forme di vita "non umane".

CAPITOLO I
INTELLIGENZA ARTIFICIALE E DIRITTO

1. Premessa; 2. Informatizzazione e intelligenza artificiale; 2.1. Nascita e sviluppo dell'intelligenza artificiale; 2.2. Intelligenza artificiale forte e intelligenza artificiale debole (Strong AI VS. Weak AI); 2.3. Principali sistemi di apprendimento: machine-learning e deep learning; 3. I nuovi orizzonti della tecnologia informatica e della robotica; 4. Il contesto normativo in materia di intelligenza artificiale e robotica; 4.1. Il quadro normativo europeo in materia di robotica; 4.1.1. Il cerchio interno: la 'direttiva macchine'; 4.1.2. Il cerchio maggiore ed il cerchio esterno: le disposizioni sulla sicurezza generale dei prodotti e sulla vendita dei beni di consumo; 4.1.3. Settori specifici VS. assenza di regolazione; 4.1.4. I progetti specifici relativi alla rilevanza dei robot per il diritto; 5. L'informatica del dirito; 6. Modelli di responsabilità in ambito penale per self-learning robot: campo di indagine; 7. Intelligenza artificiale e diritto civile; 8. Intelligenza artificiale e diritto penale; 9. Proseguimento dell'indagine

1. Premessa

«I propose to consider the question, "Can machine think?"»[4], così nell'ottobre del 1950 Alan Turing esordiva in uno dei primi contributi scientifici in materia di intelligenza artificiale[5]. Negli stessi anni in cui veniva costruita la prima rete neurale artificiale[6], l'illustre matematico stava approfondendo la possibilità teorica di introdurre una macchina pensante e proponeva un test, destinato ad influenzare gli sviluppi successivi, per verificare in termini operativi quando possa ritenersi che una macchina tenga un comportamento intelligente.

Da lì, l'idea che una macchina potesse arrivare ad elaborare un ragionamento simile a quello di un umano non risultava più del tutto peregrina.

---

[4] A. TURING, *Computing Machinery and Intelligence*, in *Mind*, 1950, p. 433.

[5] Il primo studio generalmente riconosciuto come appartenente al settore della intelligenza artificiale è quello di Warren McCulloch e Walter Pitts che, nel 1943, proposero un modello di neuroni artificiali cfr. S. J. RUSSEL, P. NORVIG, *Artificial Intelligence – A modern approach*, *Third Edition*, 2010, p. 16.

[6] *Ivi*. Si tratta della SNARC costruita da M. Minsky e D. Edmonds nel 1951.

Seguire l'evoluzione della robotica è, dunque, imprescindibile per comprendere la sfida culturale, etica, giuridica e politica che la quarta rivoluzione industriale si porta con sé. Significa, paradossalmente, fare un viaggio nel futuro dell'umanità, tentandone di capire l'essenza, guardando appunto il non-umano, eppure essere intelligente.

Il crescente impiego dell'intelligenza artificiale (nel prosieguo, anche "IA" o "AI") all'interno dei settori produttivi e dell'amministrazione pubblica ha contraddistinto i primi decenni del nuovo secolo. Lo sviluppo tecnologico ha consentito l'utilizzo sempre maggiore di elaboratori elettronici in grado di eseguire compiti di supporto e talvolta in sostituzione dell'attività umana[7].

Un passaggio fondamentale ha riguardato la possibilità di raccogliere, utilizzare e processare masse di dati mediante i quali elaborare informazioni idonee a consentire un funzionamento intelligente da parte delle macchine stesse. Attraverso questo percorso si può concepire un processo decisionale automatico affidato a degli elaboratori che, sulla base dei dati inseriti, consente l'individuazione di una specifica soluzione ad un determinato problema. Queste nuove applicazioni hanno favorito l'introduzione dell'intelligenza artificiale in molti ambiti delle attività produttive e della pubblica amministrazione. In particolare, è aumentato l'impiego di algoritmi, sequenze di istruzioni informatiche ben definite che sono impiegate per risolvere una serie di problemi o per eseguire un determinato calcolo[8].

L'algoritmo è quindi una delle diverse applicazioni dell'intelligenza artificiale e comprende un'ampia gamma di strumenti[9], che, a loro volta, influenzano una varietà di

---

[7] In tema di nuove tecnologie in generale ed intelligenza artificiale, a carattere introduttivo, si veda S. Rodotà, *Elaboratori elettronici e controllo sociale*, Bologna, 1973; F. Pasquale, *The black Box Society: The Secret algorithms that Control Money And Information*, Cambridge, 2015, 5 ss.; S. Barocas - A. D. Selbst, *Big Data's Disparate Impact*, in 104 *Calif. L. Rev.* 671, 674 N.10, 2016; I. Ajunwa, *Algorithms At Work: Productivity Monitoring Platforms And Wearable Technology As The New Data-Centric Research Agenda For Employment And Labor Law*, in 63 *St. Louis U. L.J.* 2019; I. Ajunwa, *Genetic Testing Meets Big Data: Tort And Contract Law issues*, in 75 *Ohio St. L.J.* 1225, 2014; D. K. Citron - F. Pasquale, *The Scored Society: Due Process For Automated Predictions,* in 89 *Wash. L. Rev.* 1, 2014; K. Crawford - J. Schultz, *Big Data And Due Process: Toward A Framework To Redress Predictive Privacy Harms*, in 55 *B.C. L. Rev.* 93, 2014; L. Edwards-M. Veale, *Slave To The Algorithm? Why A 'Right to an explanation' Is Probably Not The Remedy You Are Looking For,* in 16 *Duke L. & Tech. Rev.* 18, 2017; G. Resta-V. Zeno Zencovich (a cura di), *La protezione transnazionale dei dati personali,* Roma, 2016; G. Resta, *Diritti esclusivi e nuovi beni immateriali*, Torino, 2011; G. Pitruzzella, *Big Data, Competition and Privacy: A Look from the Antitrust Perspective*, in *Concorrenza e Mercato*, 2016, 15 ss.; F. Pizzetti, *Privacy e diritto europeo nella protezione dei dati personali*, Torino, 2016; G. Pascuzzi, *Il diritto nell'era digitale. Tecnologie informatiche e regole privatistiche*, Bologna, 2002. A. D. Selbst-S. Barocas, *The Intuitive Appeal Of Explainable Machines*, in 87 *Fordham L. Rev.* 2018.

[8] S. Barocas-S. Hood-M. Ziewitz, *Governing Algorithms: A Provocation Piece, Governing Algorithms*, 29 Mar. 29, 2013

[9] G. Capuzzo, *"Do Algorithms dream about Electric Sheep?" Percorsi di studio in tema di discriminazione e processi decisori algoritmici tra le due sponde dell'Atlantico"* in *Saggi - Focus: innovazione, diritto e tecnologia: temi per il presente e il futuro*, RDM 3-2020.

operazioni. Si tratta di un tipo di decisione che viene presa miliardi di volte all'anno in svariati settori come quello creditizio, concorsuale o diagnostico in ambito medico[10]. Questa tipologia di intelligenze artificiali è sviluppata utilizzando metodi di apprendimento automatico ("machine learning"). Si tratta di tipologie di algoritmi che costruiscono funzioni di previsione sulla base di una serie di dati allo scopo di realizzare tali predizioni; una volta consolidata tale funzione, l'algoritmo riceve un particolare "input" (come le caratteristiche di un candidato ad una particolare posizione lavorativa) e predice alcuni risultati (come la sua performance in specifici ambiti).

L'utilità di queste macchine è tale che sono molteplici ormai gli ambiti in cui viene sfruttata e le sue implicazioni sociali sono del tutto evidenti. Molte delle decisioni che sono prese all'interno della società si basano o sono influenzate dagli esiti di queste elaborazioni algoritmiche. Per tali ragioni, i giuristi hanno cominciato ad interessarsi alle ripercussioni nel mondo giuridico di queste nuove tecnologie.

Il loro impiego è rilevante per il diritto[11], sia perché le regole dell'ordinamento si estendono al funzionamento dell'algoritmo quando è impiegato in relazione a particolari attività umane, sia perché il diritto stesso ha cominciato ad impiegare algoritmi nell'ambito della sua applicazione.

L'intelligenza artificiale è già presente in molti aspetti della nostra vita quotidiana. È alla base di tutte le ricerche su Internet e di tutte le app; è in ogni richiesta fatta al GPS, in ogni videogame o film d'animazione, in ogni banca e compagnia di assicurazione, in ogni ospedale, in ogni drone e in ogni auto a guida autonoma, e in futuro – questa la previsione di una delle massime esperte della materia – «ce la ritroveremo dappertutto»[12]: e, ovviamente, anche in ambiti che hanno immediata rilevanza per il diritto penale. In effetti, tra gli esperti della materia «l'opinione comune è che oggi stiamo vivendo in un mondo sempre più dominato dall'Intelligenza Artificiale [...] grazie alla proliferazione di tecniche di Intelligenza Artificiale che riescono a imparare molto velocemente ed efficacemente, come gli algoritmi di machine learning, le tecniche di mining e i sistemi

[10] A. Mantelero, *I Big Data nel quadro della disciplina europea della tutela dei dati personali*, in *Il Corriere giuridico - Speciali Digitali 2018*, 2018, 46 ss.; V. Morabito, *Big Data and Analytics. Strategic and Organizational Impacts,* New York, 2015, 23 ss; P. T. Kim, *Auditing Algorithms for Discrimination*, in 166 *U. Pa. L. Rev. Online* 189, 2017; Id., *Data-Driven Discrimination At Work,* in 58 *Wm. & Mary L. Rev.* 857, 2017; P. Kim-S. Scott, *Discrimination In Online Employment Recruiting*, in 63 *St. Louis U. L.J.*, 2019; C. A. Sullivan, *Employing*, 2018 (Seton Hall Public Law Research Paper); J. A. Kroll et al., *Accountable Algorithms*, in 165 *U. Pa. L. Rev.* 633, 2017

[11] Su questo punto il rimando è a S. Rodotà, *Elaboratori elettronici e controllo sociale*, cit.; Id. *Tecnologie e diritti*, Bologna, 1995

[12] M.A. Boden, *Intelligenza artificiale*, in J. l-Khalili (a cura di), *Il futuro che verrà*, Bollati Boringhieri, 2018, p. 133

predittivi, che sembrano promettere un livello senza precedenti, e forse anche un po' spaventoso, di IA nella nostra vita e nelle nostre società»[13].

Se poi volessimo lanciarci in previsioni di lungo termine, potremmo citare Stephen Hawking, ad avviso del quale «nell'arco dei prossimi cento anni, l'intelligenza dei computer supererà quella degli esseri umani»[14]. E che non si tratti di pura fantascienza, è confermato dal fatto che un'analoga previsione è contenuta anche nei Considerando della Risoluzione del Parlamento europeo sulla robotica del 16 febbraio 2017[15], dove si afferma che «è possibile che a lungo termine l'intelligenza artificiale superi la capacità intellettuale umana».

È, quindi, facile intuire che le possibili implicazioni anche di rilevanza penale, derivanti dall'impiego delle tecnologie di IA, potrebbero essere in un prossimo futuro assai numerose e significative[16].

2. Informatizzazione e intelligenza artificiale

La disciplina dell'Intelligenza Artificiale (IA) nasce in seno all'informatica finendo poi per andare ad influenzare e pervadere moltissime altre discipline sia scientifiche che non scientifiche quali ad esempio le scienze biomediche, l'economia, la sociologia, le scienze cognitive, la giurisprudenza solo per citarne alcune. Sebbene lo sviluppo ed il progetto di sistemi intelligenti richiedano competenze in ambito informatico e matematico-statistico anche molto avanzate, la consapevolezza dell'impatto di tali sistemi e delle tecnologie collegate sta ormai diventando estremamente importante anche nel campo delle scienze umane, ed il diritto non fa certamente eccezione.

---

[13] G.F. Italiano, *Intelligenza artificiale: passato, presente, futuro*, in F. Pizzetti (a cura di), *Intelligenza artificiale, protezione dei dati personali e regolazione*, Giappicchelli, 2018, p. 216. Per analoghe considerazioni, v. J. Kaplan, *Intelligenza artificiale. Guida al futuro prossimo*, Luiss University Press, II ed., 2018, pp. 81 ss., e pp. 193 ss.; L. Floridi, *What the Near Future of Artificial Intelligence Could Be*, in *Philosophy & Technology,* fasc. 32, 2019, pp. 3 ss. (*online* al link https://doi.org/10.1007/s13347-019-00345-yp); M.C. Carrozza, *I Robot e noi*, Il Mulino, 2017, pp. 3 ss. Per una panoramica sul mondo dei sistemi di IA, dalle sue origini a oggi, raccontata da alcuni esperti della materia, si vedano i saggi raccolti in D. Heaven (a cura di), *Macchine che pensano. La nuova era dell'intelligenza artificiale*, Edizioni Dedalo, 2018.

[14] Intervento di S. Hawking durante la Conferenza *Zeitgeist*, Londra, maggio 2015 (citazione riportata da Redazione, *Do You Trust This Computer?*, in *questa rivista*, 15 maggio 2019; v. la notizia anche su Newsweek (L. Walker, *Stephen Hawking warns artificial intelligence could end humanity*, 14 maggio 2015).

[15] Risoluzione del Parlamento europeo del 16 febbraio 2017, recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica (2015/2103(INL))

[16] F. Basile, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in Diritto Penale e Uomo, 2019

La rapida evoluzione dell'Intelligenza Artificiale e le sempre più numerose applicazioni che essa trova nei diversi settori della vita quotidiana impongono una profonda riflessione sulle sue implicazioni in ambito giuridico. Qui interessano, in particolare, quelle riguardanti il diritto penale (sostanziale), dove lo straordinario sviluppo dell'Intelligenza Artificiale dell'ultimo decennio ha sollevato questioni assai delicate.

Per poter comprendere quali tecnologie oggi vengano utilizzate nell'ambito del diritto è necessario capire in primo luogo che cosa si intenda per Intelligenza Artificiale (IA).

Non è stato possibile a tutt'oggi trovare una definizione univoca di IA, alla base di questa difficoltà generale troviamo la difficoltà stessa di definire l'intelligenza umana. Tuttavia, sono emersi, tra le molte definizioni che sono state date nel corso degli anni, dei fattori comuni che portano ad individuare l'IA come quell'insieme di programmi informatici o macchine capaci di comportamenti che riterremmo intelligenti se messi in atto da esseri umani[17].

Prima di riuscire a definire l'IA occorrerebbe trovare una definizione universale di intelligenza per valutare poi se questa possa adattarsi al mondo dei computer, ma tale definizione non esiste[18]. Per migliaia di anni, ricercatori, filosofi e scienziati hanno provato a capire come gli esseri umani riescano a pensare, e come è possibile per un essere intelligente, interagire con il complesso mondo che lo circonda, percepire, capire e predire eventi e modificare l'ambiente in cui vive a seconda dei suoi desideri e delle sue necessità e come riesca a comunicare con altri essere intelligenti[19].

C'è talmente tanta incertezza che l'Oxford Companion to the Mind, in esordio della sua trattazione della voce 'intelligenza' riferisce che «sono disponibili innumerevoli test per misurare l'intelligenza ma nessuno sa con sicurezza che cosa sia l'intelligenza, e addirittura nessuno sa con sicurezza che cosa misurino i test disponibili»[20]. Un gruppo di cinquantadue professori e ricercatori di settori correlati allo studio dell'intelligenza nelle maggiori università statunitensi di Londra hanno cercato di fornire un'enunciazione condivisa pubblicata sotto il nome di 'Mainstream Science on Intelligence'[21] dove

[17] J. KAPLAN, *Intelligenza artificiale, guida al futuro prossimo*, Roma 2017 p. 15

[18] G. CIACCI, G. BUONUOMO, Profili di informatica giuridica, Milano 2018 p.75

[19] G. CONTISSA, *Information technology for the law*, Torino 2017 p. 141

[20] R.L. GREGORY, "intelligence", in: The Oxford Companion to the Mind, a cura di R.L. Gregory, Oxford: Oxford University Press, 1987, pag.375 testo originale: "Innumerable tests are available for measuring intelligence, yet no one is quite sure of what intelligence is, or even of just what is that the available tests are measuring".

[21] Pubblicata in origine sul Wall Street Journal del 13 dicembre 1994, sul web http://wwwl.udel.edu/educ/gottfredson/reprints/1994WSJmainstream.pdf consultato il 12 novembre 2017, e poi da L.S. GOTTFREDSON, Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography, in Intelligence, Volume 24, Issue 1 , 1997, p. 13-23, di cui si può leggere una sintesi in questo link http://central.gutenberg.org/articles/mainstream_scince_on_intelligence.

l'intelligenza viene descritta come la capacità di ragionare, pianificare, risolvere problemi, pensare in maniera astratta, comprendere idee complesse, apprendere rapidamente e apprendere dall'esperienza.
L'intelligenza è considerata come un'abilità che non riguarda solo l'apprendimento dai libri o l'astuzia nei test, ma è una capacità più ampia e profonda di capire tutte le cose con cui entriamo in qualche modo in contatto consentendo all'uomo di attribuire un significato alle cose che lo circondano e capire come debba comportarsi di conseguenza. Nonostante questa definizione piuttosto completa, non sono mancate critiche da parte della comunità scientifica.
Per quanto riguarda il campo dell'IA si cerca di andare ancora oltre, infatti si vuole cercare non solo di capire, ma anche di costruire delle entità intelligenti, cioè delle entità che sono in grado di ragionare, imparare, adattarsi, comunicare e riuscire a risolvere compiti che normalmente richiederebbero l'intelligenza umana.

2.1. Nascita e sviluppo dell'intelligenza artificiale

L'espressione "Artificial Intelligence" nasce nell'ambito di una serie di workshops promossi, nel 1956, al Dartmouth College in New Hampshire, da un gruppo di accademici, capitanato da John McCarthy, mossi dall'aspirazione di poter insegnare alle macchine a migliorarsi autonomamente e a risolvere problemi riservati all'uomo.
Eppure, già in epoca precedente le innovazioni tecnologiche erano state di notevole impatto.
In un'antichità sorprendentemente remota l'uomo aveva cominciato a realizzare macchine in grado di eseguire operazioni impossibili per chi non fosse uno scienziato.
Il primo esempio è la "Macchina di Anticitera" risalente al 150/100 a.C., un antico calcolatore frutto della cultura ellenica che, grazie a ingranaggi di precisione, consentiva di determinare il sorgere del sole, le fasi lunari, gli equinozi e perfino le date dei giochi olimpici[22]. Nei secoli seguenti e nelle principali culture del mondo molti sono gli esempi di macchine con meccanismi sempre più complicati. Lo stesso Leonardo da Vinci aveva progettato un automa cavaliere in grado di eseguire movimenti molto simili a quelli umani grazie ad una serie di cavi e pulegge con meccanismi che controllavano separatamente la

[22] V.CASADEI, Dagli automi alla moderna robotica: un'insolenza che dura 2000 anni, https://insolenzadir2d2.it/dagli-automi-alla-moderna-robotica-uninsolenza-dura-2000-anni/ , Novembre 2019

parte superiore e quella inferiore del Cavaliere. Nel XVII secolo con la rivoluzione scientifica[23] ha un nuovo impulso la scienza moderna[24]. In questo periodo infatti vengono costruite macchine in grado di effettuare calcoli automatici.

La più nota è la "Pascalina" costruita da Blaise Pascal[25], una macchina utilizzata per compiere addizioni e sottrazioni, perfezionata successivamente da Gottfried Wilhelm von Leibniz[26] che costruì una macchina calcolatrice in grado di eseguire anche moltiplicazioni. Tutte e due queste macchine però erano calcolatrici automatiche ma non elettroniche. Funzionavano grazie ad ingranaggi e ruote dentate senza elettricità e non erano programmabili, potevano infatti eseguire di volta in volta solo una delle operazioni rientranti nella loro competenza (una sola addizione, sottrazione, moltiplicazione) senza poter compiere autonomamente un'intera combinazione di tali operazioni sulla base di indicazioni fornite in anticipo. Solamente con Charles Babbage venne concepito il motore analitico ("analitical engine"), cioè una macchina dotata di una competenza generale, capace di eseguire qualsiasi tipo di calcolo e soprattutto programmabile attraverso istruzioni riportate su schede perforate.

Successivamente, si cominciava, infatti, ancor prima di utilizzare la parola "intelligenza artificiale", a parlare di cibernetica, nel cui ambitovi sono stati validi esempi di scoperte rivoluzionarie a partire dall'aritmometro di V.T: Odhner[27], fino alla creazione dei primi elaboratori elettronici.

Nella seconda metà del Novecento si fa spazio, dunque, il concetto di "diritto artificiale", il quale si affianca a quello di diritto naturale. Il ponte di passaggio che ha consentito di collegare la cibernetica alla giurisprudenza, è stato offerto dall'uso della logica simbolica nel terreno culturale degli strumenti giuridici. Infatti, è grazie alla presentazione logica dei termini di una questione, che il giurista può rivolgersi a un elaboratore elettronico, per servirsene ai fini di un calcolo logico.

---

[23] Anche se convenzionalmente l'anno di inizio della rivoluzione scientifica è considerato il 1543 con la pubblicazione del De revolutionibus orbium coelestium di Niccolò Copernico

[24] P. FABBRI, Cos'è l'intelligenza artificiale e quali sono le sue applicazioni attuali e future, https://www.zerounoweb.it/analytics/cognitive-computing/cosa-intelligenza-artificiale/ consultato il novembre del 2019

[25] Blaise Pascal (1623-1662) è stato un matematico, fisico, filosofo e teologo francese. Bambino prodigio, fu istruito dal padre. I primi lavori di Pascal sono relativi alle scienze naturali e alle scienze applicate. Contribuì in modo significativo alla costruzione di calcolatori meccanici e allo studio dei fluidi.

[26] Gottfried Wilhelm von Leibniz (1646-1716) è stato un matematico, filosofo, scienziato, logico, teologo, glottoteta, diplomatico, giurista, storico, magistrato tedesco di origine soraba

[27] V.Knapp, *L'applicabilità della cibernetica al diritto*, Giulio Einaudi editore, Torino 1978.
L'aritmometro di Odhner era una delle prime macchine da calcolo: consentiva di effettuare calcoli con velocità tripla rispetto a quella dell'abaco, strumento prediletto nei calcoli quotidiani.

I due grandi "laboratori" della nuova direzione di pensiero, dunque, furono l'Unione Sovietica e gli Stati Uniti d'America, i quali, sono stati in grado di raggruppare un notevole corpo di scienziati ricercatori, adiuvati, grazie allo sviluppo di detti paesi, da notevoli mezzi finanziari[28].

Partendo da questo, la recezione della cibernetica nell'Unione Sovietica si articola in 3 fasi:

1) Negli anni immediatamente precedenti e seguenti la morte di Stalin, la cibernetica venne attaccata e dichiarata "la scienza degli oscurantisti", causando un pericolo sia per i paesi socialisti che capitalisti, ritenendo assurdo ed erroneo mettere sul medesimo piano un elaboratore elettronico ed il cervello umano.

2) Nella seconda fase (1954-58) si registra un atteggiamento di maggior liberalità nei riguardi del dibattito scientifico, avendo come punto iniziale la conferenza "Che cos'è la cibernetica?" tenuta il 19 novembre 1954 all'Accademia delle scienze sociali del Comitato centrale del PCUS.

3) La terza fase, dal 1958 ad oggi, vede l'affermazione della cibernetica nei più diversi settori, essa è una realtà che non può essere ignorata.

Alla fine della seconda guerra mondiale, l'Europa centrale supera l'Unione Sovietica nella produzione di elaboratori, i quali trovano una distribuzione sempre maggiore. Il loro utilizzo ha destato ed incentivato il dibattito teorico da cui Viktor Knapp ha tratto spunto per la sua ricerca che ha inserito nel suo libro "L'applicabilità della cibernetica al diritto". Egli creò un gruppo di lavoro dedicato ai problemi dell'applicazione della cibernetica, ispirato dai consigli del suo collega alla Accademia delle scienze a Praga, Arnost Kolman[29].

Quando ancora non si parlava di Intelligenza artificiale, cominciava a farsi spazio nel dibattito teorico la parola "cibernetica", quale scienza che studia i meccanismi di interazione negli esseri viventi e la loro riproducibilità sulle macchine. Il primo ad aver posto un rapporto di connessone tra la cibernetica e il diritto è stato il matematico Norbert Wiener, il quale affermò: "i problemi giuridici sono per loro natura problemi di comunicazione e di cibernetica, e cioè sono problemi relativi al regolato e ripetibile governo di situazioni critiche"[30].

---

[28] V.Frosini, *Cibernetica: Diritto e Società*, Ed. di Comunità (Collana "diritto e cultura moderna", n6), Milano 1968.

[29] V.Knapp, *L'applicabilità della cibernetica al diritto*, Giulio Einaudi editore, Torino 1978.

[30] N.Wiener, *The Human Use of the Human Beings,* trad. it. di D. Persiani con il titolo *Introduzione alla cibernetica,* Boringhieri, Torino 1953.

Fu proprio Wiener, a coniare per la prima volta il termine "cibernetica", il quale affermò che: "La cibernetica è la scienza del controllo e della comunicazione della società". Egli si valse del modello cibernetico come di uno schema interpretativo da applicare anche alla comprensione di strutture sociali.
Nonostante ciò, esso non si identifica con l'intero Management sociale, la cibernetica è quindi una scienza che affronta soltanto alcuni aspetti relativi alla guida della società.
Per la nascita ufficiale del campo di ricerca della intelligenza artificiale occorreva tuttavia aspettare l'estate del 1956, quando i brillanti studiosi John McCarhty, Marvin Minsky, Claude Shannon e Nathaniel Rochester invitarono una serie di ricercatori di teoria degli automi, reti neurali e studio della intelligenza a partecipare a un workshop di due mesi presso la città di Dartmouth[31].
Proprio in occasione dell'evento, chiamato "Dartmouth Summer Research Project on Artificial Intelligence", è utilizzata per la prima volta in modo scientificamente significativo l'espressione "Artificial Intelligence".
Si legge infatti nella dichiarazione di intenti redatta dagli autori:

«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»[32].

In seguito, i primi anni di sviluppo della nuova disciplina si contraddistinsero per importanti successi[33].

[31] S. J. RUSSEL, P. NORVIG, *Artificial Intelligence*, cit., p. 17; G.F. ITALIANO, *Intelligenza Artificiale: passato, presente, futuro*, in PIZZETTI F. (a cura di), *Intelligenza artificiale, protezione dei dati personali e regolazione*, Giappichelli, 2018, p. 208 ss.
[32] J. MCCARTHY, M.L. MINSKY, N. ROCHESTER, C. E. SHANNON, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, in *AI Magazine*, 2006, p. 12.
[33] S. J. RUSSEL, P. NORVIG, *Artificial Intelligence*, cit., p. 18

L'approccio secondo cui erano progettati i primi algoritmi consisteva nella tecnica di esplorare passo dopo passo un determinato spazio di ricerca. In modo non dissimile da chi cerca la strada per uscire da un labirinto, gli algoritmi 'tentavano' i diversi percorsi possibili e, in caso di errore, tornavano indietro per intraprendere nuove strade e così fino a quando avessero raggiunto la soluzione[34].

Sulla base di tale impostazione furono costruiti programmi capaci di risolvere teoremi matematici[35], di giocare a dama[36], così come si approfondì lo studio delle reti neurali artificiali.

Rilevanti scoperte furono fatte anche nell'ambito del linguaggio naturale. Tra il 1964 ed il 1966, ad esempio, Joseph Weizenbaum creò ELIZA, un sistema in grado di simulare la conversazione con un essere umano, sistema che secondo alcuni studiosi costituirebbe la prima macchina capace di superare il test proposto da Turing.

Il Test di Turing nasce come un criterio per determinare se una macchina sia in grado di pensare come un essere umano. Tale criterio è stato suggerito da Alan Turing nell'articolo "Computing machinery and intelligence", apparso nel 1950 sulla rivista Mind. Si tratta di una situazione sperimentale consistente in un gioco a tre partecipanti: un uomo A, una donna B e una terza persona C. Quest'ultimo è tenuto separato dagli altri due e, tramite una serie di domande, deve stabilire qual è l'uomo e quale la donna. Dal canto loro, anche A e B hanno dei compiti: A deve ingannare C e portarlo a fare un'identificazione errata, mentre B deve aiutarlo. Affinché C non possa disporre di alcun indizio (come l'analisi della grafia o della voce), le risposte alle domande di C devono essere dattiloscritte o similarmente trasmesse. Il test di Turing si basa sul presupposto che una macchina si sostituisca ad A. In tal caso, se C non si accorgesse di nulla, la macchina dovrebbe essere considerata intelligente, dal momento che sarebbe indistinguibile da un essere umano. La macchina cioè dovrebbe essere considerata come dotata di una "intelligenza" pari a quella dell'uomo. Il criterio proposto da Turing è ritenuto alla base della intelligenza artificiale; tuttavia per stabilire l'intelligenza esso si basa su un presupposto meramente osservativo

[34] G.F. ITALIANO, *Intelligenza Artificiale: passato, presente, futuro*, cit., p. 210.

[35] Nel 1959 Herbert Gelertner costruì il *Geometry Theorem Prover*, capace di dimostrare teoremi matematici.

[36] Si veda il programma scritto da Arthur Samuel e dimostrato in televisione nel 1956. Inizialmente Samuel era in grado di sconfiggere il suo programma, ma dopo pochi mesi non riuscì più a vincere una partita. Fu uno dei primi esempi di *machine learning* cfr. S. FRANKLIN, *History, motivations, and core themes*, in FRANKISH K., RAMSEY W. M. (a cura di), *The Cambridge Handbook of Artificial Intelligence*, Cambridge University Press, p. 19.

e quindi necessita di ulteriori precisazioni data la complessità del concetto stesso d'intelligenza e l'impossibilità di una sua definizione universalmente accettata[37].

Dopo un approccio interessato, gli studi in materia di intelligenza artificiale si interrompono d'un tratto. Emblematica, sul punto, è la policy seguita dal Regno Unito che, nel 1973, decide di interrompere i finanziamenti finalizzati alla ricerca in materia di IA.

Una decisiva ripresa degli studi nell'ambito dell'intelligenza artificiale si ebbe con la nascita dei cd. sistemi esperti[38]. Nei primi anni di sviluppo della nuova disciplina, l'attenzione dei ricercatori si era concentrata sulla creazione di sistemi che fossero in grado di ragionare su problemi astratti. Gli studiosi si resero tuttavia conto che la capacità di ragionamento è soltanto una parte dell'universo dell'intelligenza, che ricomprende anche la capacità di fare i conti con i problemi reali[39]. Tale capacità richiede che chi si appresta a risolvere i problemi possieda alcune conoscenze necessarie. I sistemi esperti vengono incontro a questa esigenza: si tratta infatti di programmi informatici in grado di rappresentare e ragionare con la conoscenza di alcuni soggetti specializzati in un determinato settore allo scopo di risolvere problemi o dare consigli. Si può così pensare al sistema *Mycin* che fu creato per aiutare i medici a diagnosticare e trattare alcune malattie infettive del sangue e le meningiti[40]. Tali sistemi si rivelarono particolarmente utili per le applicazioni industriali.

Il vero e proprio cambio di passo si è tuttavia avuto nell'ultimo ventennio.

Una delle conquiste più importanti - anche alla luce della grande copertura mediatica - si ebbe nel 1997, quando un sistema di intelligenza artificiale costruito da IBM e capace di giocare a scacchi (conosciuto come "Deep Blue") riuscì a sconfiggere il campione mondiale di scacchi Gary Kasparov in un match composto di sei sfide[41].

Ancora più di recente, nel 2017, *Google AlphaGo* è riuscita a sconfiggere i campioni del complicato gioco *Go*[42] e un altro sistema *Libratus* a vincere contro i migliori giocatori del gioco di poker *Texas Hold'em*.

---

[37] https://www.treccani.it/enciclopedia/test-di-turing_%28Enciclopedia-della-Matematica%29

[38] Il primo esempio di questo approccio è costituito dal programma DENDRAL (1965), sviluppato da Edward Feingenbaum, Joshua Lederberg e Bruce Buchanan. DENDRAL aiutò ad identificare la struttura molecolare di molecole organiche, analizzando i dati di uno spettrometro di massa ed utilizzando le sue conoscenze in chimica. Ad esso seguì Mycin, un altro sistema esperto che aiutò i medici a diagnosticare e trattare le malattie infettive del sangue e meningiti cfr. J. RUSSEL, P. NORVIG, *Artificial Intelligence*, cit., pp. 22-23.

[39] S. FRANKLIN, *History, motivations, and core themes*, cit., p. 20.

[40] S. J. RUSSEL, P. NORVIG, *Artificial Intelligence*, cit., p. 23.

[41] S. FRANKLIN, *History, motivations, and core themes*, cit., p. 23.

[42] G.F. ITALIANO, *Intelligenza Artificiale: passato, presente, futuro*, cit., p. 214.

Al di là questi recenti traguardi, il dato più significativo è che, oggi, l'intelligenza artificiale è largamente presente nella vita quotidiana. Essa è utilizzata per tradurre automaticamente da una lingua all'altra, per riconoscere discorsi, per effettuare pianificazioni logistiche (si pensi ad un'azienda di trasporti che deve programmare destinazioni, tratte, etc.), per combattere *e-mail* spam e così via. Molte delle sue applicazioni sono già molto diffuse ed implementate in strumenti di uso comune come i telefoni cellulari.

Questa crescita e diffusione sembra destinata a subire un ulteriore sviluppo nei prossimi anni. Il futuro prospetta infatti un impiego trasversale delle applicazioni fondate sull'intelligenza artificiale, con l'introduzione di veicoli a guida autonoma che rivoluzioneranno il sistema dei trasporti, l'utilizzo di prodotti robotici per lo svolgimento di complicate operazioni mediche (*Surgical Robots*) o per l'assistenza post operatoria, l'uso dell'intelligenza artificiale per operazioni di *law enforcement* (es. droni di sorveglianza, algoritmi per la *fraud detection*) o come ausilio nell'attività educativa (es. sistemi interattivi in grado di modellarsi sull'esigenze personali del discente), fino all'impiego dei robot per lo svolgimento di faccende quotidiane (*Home and Service Robots)*. L'intelligenza artificiale pare poi suscettibile di provocare strutturali cambiamenti economici e sociali, in particolare, nell'ambito del mercato del lavoro, attraverso la sostituzione degli esseri umani nello svolgimento di determinate mansioni.

2.2. Intelligenza artificiale forte e intelligenza artificiale debole (Strong AI VS. Weak AI)

Nell'ambito dell'intelligenza artificiale si è anzitutto soliti distinguere tra intelligenza artificiale in senso forte (Strong AI) ed intelligenza artificiale in senso debole (Weak AI). L'intelligenza artificiale in senso forte (chiamata alle volte anche Artificial General Intelligence, sinteticamente AGI) mira a ricostruire nelle macchine un livello di intelligenza pari a quello dell'uomo[43] o, secondo una definizione di matrice filosofica, mira a costruire macchine che siano effettivamente pensanti[44].

Ai suoi albori, la ricerca nel campo dell'intelligenza artificiale era dominata dall'idea di potere costruire macchine "human-like".

L'intelligenza artificiale forte parte dalla concezione che anche i calcolatori abbiano una mente o finiranno per averla e che siano capaci di stati cognitivi e di pensiero nel modo

[43] S. FRANKLIN, *History, motivations, and core themes*, cit., p. 16.
[44] S. J. RUSSEL, P. NORVIG, *Artificial Intelligence*, cit. p. 1020.

in cui ne sono dotati gli esseri umani[45]. Lo studioso del linguaggio e della mente John Searle definisce intelligenza artificiale forte dicendo che: «Un computer programmato in maniera appropriata è davvero una mente, nel senso che quando a un computer vengono date le giuste istruzioni si può letteralmente dire che capisca e abbia gli altri stati cognitivi»[46]. La possibilità di sviluppare forme di intelligenza forte rimane tutt'oggi su un livello ipotetico visto che nessun sistema di IA può essere ancora comparato all'intelligenza umana.

Il punto di riferimento per attestare il raggiungimento dell'obiettivo desiderato era indicato nella capacità di passare il test di Turing, il quale, come detto, aveva proposto di sostituire la domanda «le macchine possono pensare?» con il gioco dell'imitazione[47].

Il test di Turing, basandosi sulla volontà di ricreare un'IA forte, corre il rischio di sollevare un importante problema teorico: un calcolatore che per ipotesi riuscisse a superare tale test, potrebbe essere considerato una vera intelligenza o invece sarebbe un mero 'idiota sapiente' che simula una mente senza possederla veramente? Turing, sostiene che pensare significhi produrre espressioni che abbiano un senso, non prive di significato, che siano il risultato di un concatenamento di idee che portano a un risultato logico[48]. In previsione di future critiche sulla capacità di pensare di una macchina, nel suo articolo 'Computing machinery and intelligence' sostiene che: «L'unico modo essere sicuri che una macchina pensa è quello di essere la macchina stessa e sentire se si sta pensando. (...) Allo stesso modo, la sola via per sapere che un uomo pensa è quello di essere quell'uomo in particolare. (...) Invece di discutere continuamente su questo punto, è normale attenersi alla educata convenzione che ognuno pensi»[49].

Nonostante questa premessa comunque le critiche alle sue considerazioni non mancarono. La critica più rilevante al test di Turing è sicuramente quella di John Searle che, negli anni '80, elabora il famoso "argomento della stanza cinese". Searle infatti, ipotizza di chiudere in una stanza un essere umano in grado di parlare solo l'inglese, con un manuale che specifica come una volta ricevuto un input composto da una sequenza di caratteri

[45] G. SARTOR, op. cit. , p. 273

[46] J.R. SEARLE, Minds, Brains and Programs, in The Behavioural and Brain Science, 1980 p.417

[47] G. FORNERO, *Intelligenza artificiale e filosofia*, in N. ABBAGNANO, *Storia della filosofia*, Torino, 1994 p. 523. Sulle diverse interpretazioni del Test di Turing cfr. PROUDFOOT D., COPELAND J., *Artificial Intelligence*, cit., p. 150 ss. Secondo l'interpretazione tradizionale il test di Turing fornisce una definizione operativa di intelligenza in termini di comportamento, inscrivendosi in un'impostazione di tipo *behaviourista*. L'interpretazione minoritaria, invece, afferma che il gioco dell'imitazione è un esperimento per vedere se chi si rapporta con la macchina (nel caso in esame, l'intervistatore) «immaginerà l'intelligenza», sottolineando dunque il ruolo da attribuire alla risposta dell'uomo di fronte al comportamento della macchina.

[48] G. CIACCI, G. BUONUOMO, op.cit., p. 79

[49] A.M. TURING, *Computing machinery and intelligence*. op.cit.

cinesi, debba produrre in output un'altra sequenza di caratteri cinesi. In questo modo l'individuo all'interno della stanza è in grado di rispondere correttamente a delle domande poste in cinese senza però conoscere effettivamente questa lingua[50]. L'interrogante non sarebbe quindi in grado di distinguere se sta dialogando effettivamente con un cinese o meno.

Questa ipotesi di Searle vuole rendere evidente come un calcolatore guidato da un software (manuale di istruzioni) anche se possa essere scambiato per un essere intelligente, non possieda pensieri, non ha una mente ma si limita alla cieca manipolazione di simboli. Verrebbe da chiedersi però quale sia la differenza tra le idee che passano per la testa ad un essere umano e i bytes che sfrecciano in un computer. In entrambi i casi infatti si tratta di informazioni che entrano sotto qualche forma che potremmo definire simbolica, come ad esempio i piccoli segnali nervosi che arrivano agli occhi, per poi venire processati ed uscire. Searle sostiene che deve trattarsi di cose diverse, ma che semplicemente non capiamo ancora cosa fa il cervello. Egli crede che nel nostro cervello accada qualcosa di ancora ignoto che solo una volta conosciuto ci permetterà di comprendere quei fenomeni tipicamente umani che non sono il solo 'pensare' ma anche la coscienza, la sensazione di provare delle cose, la consapevolezza di essere vivi. Searle non sostiene che un programma non potrà mai svolgere determinati compiti sia che si tratti di comporre una canzone, sia che si tratti di consolare qualcuno dopo la perdita di una persona cara, questo però sarà generato solo da un programma migliore che sarà in grado di simulare meglio il pensiero senza peraltro mai duplicare i processi che avvengono nelle menti delle persone mentre pensano. Le conclusioni raggiunte da Searle sono state però messe in discussione in molti modi. In particolare, alcuni considerano che anche se l'uomo all'interno della stanza cinese, non capisce e parla cinese, l'intero sistema composto da: stanza, uomo e manuale di cinese è invece in grado di capirlo. Se si considera la comprensione di una lingua come la capacità di rispondere a input in quella lingua producendo output appropriati nella stessa, abbiamo esattamente ciò che il sistema stanza cinese fa. L'errore di Searle sarebbe quindi quello di concentrarsi solo su un unico aspetto, cioè quello dell'uomo che risponde seguendo istruzioni simboliche, senza considerare che la mente umana nell'effettuazione di operazioni di ragionamento per comprendere una lingua è essa stessa l'insieme di complesse operazioni di collegamento tra differenti facoltà quali la memoria, la conoscenza, i sensi ecc. e prendere in

[50] J. SEARLE, *La mente è un programma?*, in Le Scienze, Marzo 1990 n. 259

considerazione solo una di queste facoltà farebbe venir meno la comprensione degli stessi processi della mente. Obiezioni sono state sollevate sul fatto che il motivo per cui l'uomo nella stanza ed anche il sistema stanza stesso non comprendano il cinese è dovuto al fatto che la comprensione di un linguaggio, richiede la capacità di connettere le parole ai loro referenti reali, e quindi è necessario aver avuto esperienza degli oggetti (o almeno parte di essi) di cui parla il linguaggio. Questo tipo di limite però riguarda un calcolatore isolato e non si applica a quei sistemi automatici che riescono ad unire capacità percettive a quelle attinenti all'elaborazione e alla registrazione delle informazioni. Quindi i limiti che sono riscontrati dalla stanza cinese, non sono limiti che riguardano in toto l'IA, ma possono essere superati facendo estendere il sistema con dispositivi che siano dotati di appositi sensori che gli permettano di interagire con il mondo circostante e che, possibilmente, gli conferiscano la capacità di movimento.

Negli anni successivi, le difficoltà legate al raggiungimento della creazione di un'intelligenza generale (come dimostrato dagli esiti del Loebner Prize, una competizione annuale avviata nel 1991 che mira a premiare i programmatori di sistemi che siano in grado di superare il Test di Turing) portarono tuttavia la maggior parte degli studiosi a concentrare gli sforzi sulla costruzione di sistemi che fossero in grado di operare in modo intelligente nell'ambito di un più ristretto dominio.

L'altra concezione di IA è quella definita 'debole'.

L'intelligenza artificiale in senso debole (chiamata anche Narrow AI) persegue allora lo scopo meno ambizioso di creare macchine capaci di concorrere con l'uomo in compiti specifici[51] o, secondo l'accennata terminologia filosofica, che si comportino come se fossero effettivamente pensanti[52]. È nell'ambito di queste ricerche che, ad esempio, sono stati costruiti sistemi di intelligenza artificiale capaci di competere con l'uomo nell'ambito dei giochi (dama, scacchi etc.) o di essere di ausilio nello svolgimento di compiti delicati (diagnosi mediche).

In accordo a questa seconda visione non sarebbe mai possibile raggiungere, a prescindere dagli sviluppi tecnologici, e dalla possibilità per il computer di superare il test di Turing, sistemi artificiali che abbiano una mente e cioè che siano in grado di pensare davvero (IA forte). L'intelligenza artificiale debole si ripropone infatti di realizzare dei sistemi capaci

[51] S. VALLOR, G. A. BEKEY, *Artificial Intelligence and the Ethics of Self-Learning Robots*, in LIN P., ABNEY K., BEKEY G. A., JENKINS R., *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*, Oxford, 2017, p. 339.

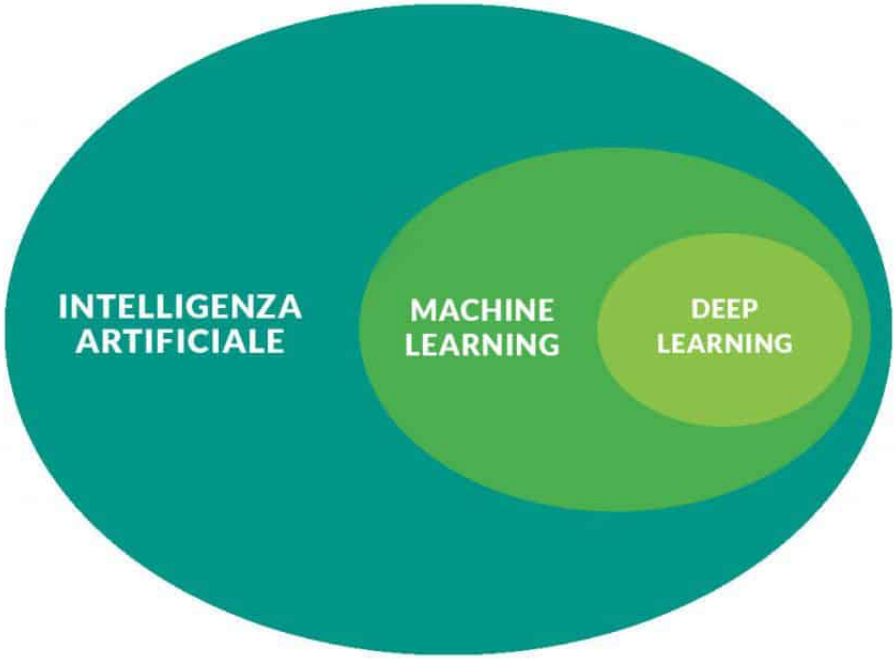
[52] S. J. RUSSEL, P. NORVIG, *Artificial Intelligence*, cit., p. 1020.

di svolgere compiti assai complessi, ma che si limitano a simulare aspetti dei processi cognitivi umani senza poterli mai davvero riprodurli[53].
La maggior parte degli esperti si confronta oggi con questa concezione più ristretta di intelligenza artificiale, mentre l'intelligenza artificiale forte è tendenzialmente concepita, almeno nelle previsioni più ottimistiche, come obiettivo di lungo periodo. Va peraltro dato atto che di recente la ricerca in questo settore sembra avere ripreso vigore anche attraverso l'opera di importanti autori[54].

2.3. Principali sistemi di apprendimento: machine-learning e deep learning

Nell'ambito dell'intelligenza artificiale esistono numerosi sotto-campi (*Machine Vision*, linguaggio naturale, sistemi esperti etc.). Tra di essi un ruolo di primo piano ai fini del presente lavoro assume il cd. *Machine Learning* (da qui in poi più sinteticamente '*ML*')[55].
Il Machine Learning fa parte del grande insieme racchiuso da quella che viene definita Cognitive Computing, ovvero tutte quelle applicazioni di intelligenza artificiale volte ad imitare con una macchina il comportamento e l'intelligenza umana e soprattutto migliorare l'interazione uomo-macchina.

[53] G. SARTOR, op. cit. , p. 273
[54] Cfr. D. PROUDFOOT, J. COPELAND, *Artificial Intelligence*, in MARGOLIS E., SAMUELS R., STICH S.P. (a cura di), *The Oxford Handbook of Philosophy of Cognitive Science*, Oxford, 2012, p. 166 ss.
[55] Termine coniato dall'informatico statunitense Athur L. Samuel, Some studies in machine learning using the game of checkers, in IBM Journal of research and development, 1959.



Fonte: 3rd place. AI models of datrix

Il Machine Learning è una tipologia di software basato su algoritmi di intelligenza artificiale che hanno la capacità di apprendere automaticamente dal set di dati che gli vengono forniti dall'uomo e anche dall'esperienza. Una peculiarità interessante di tale tecnologia è il fatto che imparando migliora anche la sua performance.

L'obiettivo di questa disciplina è di sviluppare sistemi artificiali imparando dalle capacità acquisite tramite l'esperienza. Nonostante spesso il ML sia completamente confuso con l'IA, deve essere chiaro che si tratti solamente di una sottocategoria con obiettivi specifici riguardanti l'apprendimento automatico da dati e situazioni. L'IA, al contrario, ha obiettivi molto più ampi che riguardano la costruzione di agenti intelligenti che, una volta apprese le informazioni necessarie (sia dai dati sia perché fornite esternamente in qualche altra forma), sfruttano quest'ultime per svolgere specifici compiti, i quali normalmente richiedono una qualche forma di intelligenza per essere portati a termine.

Per entrare nel concetto di *ML* conviene partire dal concetto di algoritmo.

In linea generale, con il termine algoritmo si indica una procedura utilizzata per raggiungere uno specifico risultato, caratterizzata dal fatto che: *a)* nessun passaggio della procedura richiede intuitività o creatività; *b)* al termine di ogni passaggio è chiaro quale sia quello immediatamente successivo; *c)* la procedura assicura il raggiungimento dello specifico risultato in un numero finito di passaggi. Traendo un esempio dall'esperienza della vita quotidiana costituisce un algoritmo l'espediente di chi, dovendo aprire una

porta, non si ricordi quale sia la chiave e tenti nell'ordine tutte le chiavi del mazzo fino a trovare quella corretta. Si tratta infatti di una procedura fondata su passaggi 'semplici', che non comportano né intuito, né creatività, ed il cui risultato sia certo (sempre che la chiave sia presente nel mazzo!)[56].

Mentre gli algoritmi tradizionali sono programmati manualmente per seguire passi espliciti, la caratteristica che contrassegna gli algoritmi di *ML* è quella di 'imparare' il programma automaticamente a partire dall'osservazione dei dati.

Sgombriamo subito il campo da un equivoco; l'idea che i computer possano 'imparare' va considerata per lo più come una metafora: essa non implica che i computer siano capaci di replicare in via artificiale l'avanzato sistema cognitivo di pensiero chiamato in causa nell'apprendimento degli esseri umani, quanto piuttosto che tali sistemi siano in grado di migliorare la propria performance in relazione a compiti futuri dopo avere svolto osservazioni sulla realtà[57].

Nell'ambito del *ML* esistono diversi modi attraverso cui un algoritmo può essere 'allenato':

i. apprendimento supervisionato (*supervised learning*) L'apprendimento supervisionato si caratterizza per il fatto che gli esempi forniti all'algoritmo sono composti da una serie di valori di *input* accompagnati da una etichetta' (*label*), indicante il risultato o un giudizio di valore.

Seguendo un esempio molto noto, immaginiamo che ad un algoritmo siano forniti come dati di *input* alcune informazioni relative a un gruppo di immobili (metratura, numero di stanze etc.) e delle 'etichette' indicanti il prezzo di ciascun immobile. Dall'analisi degli esempi, l'algoritmo ricava una generalizzazione che collega i dati di *input* (le caratteristiche degli immobili) ai dati di *output* (il prezzo dell'immobile). La regola generale può poi essere impiegata per predire il valore di immobili mai visti in precedenza[58].

L'algoritmo può essere utilizzato sia per predire dei valori continui (es. prezzo delle case), sia per predire/effettuare delle classificazioni. Se forniamo all'algoritmo le informazioni su vari individui e le categorie a cui appartengono, l'algoritmo produce un modello che

[56] J. COPELAND, voce *Artificial Intelligence*, in S. GUTTENPLAN, *A companion to the Philosophy of Mind*, 1996, p. 124.
[57] S. J. RUSSEL, P. NORVIG, *Artificial Intelligence*, cit., p. 693; I. H. WITTEN, E. FRANK, M.A. HALL, *Data mining. Practical Machin Learning Tools and Techniques*, Elsevier, Burligton, 2011, p. 7.
[58] L. MC ARTHUR, *Machine Learning for Philosophers*, Beneficial AI Society, Edinburgh, 2019., p. 4.

può essere utilizzato per prevedere l'appartenenza alla categoria di nuovi individui[59]. Sulla stessa scia, il *ML* è stato impiegato dagli allevatori in Nuova Zelanda come aiuto per scegliere quali mucche tenere ogni anno per la produzione del latte e quali, invece, vendere per il macello[60].

ii. *apprendimento non supervisionato* (*unsupervised learning*). In questa ipotesi, il programmatore fornisce all'algoritmo soltanto un gruppo di dati di *input*, privi di alcuna 'etichetta'. È dunque lo stesso algoritmo che si incarica di analizzare le informazioni di cui dispone, di classificarle e strutturarle, in modo da formare autonomamente conoscenze più generali. Probabilmente l'uso più comune dell'apprendimento non supervisionato è negli algoritmi di *clustering*, che sono utilizzati per separare i vari individui in gruppi "naturali" secondo l'una o l'altra metrica. Questi algoritmi a volte tracciano linee relativamente arbitrarie tra individui, ma possono essere abbastanza efficaci nello scoprire gruppi quando realmente esistono. Ad esempio, si potrebbe misurare l'atteggiamento delle persone nei confronti di varie questioni politiche e quindi determinare se vi sono gruppi naturali che possono essere definiti da tali convinzioni[61].

iii. *apprendimento rinforzato*. L'apprendimento rinforzato si basa su un sistema di ricompense e di punizioni. In questa ipotesi, non c'è un gruppo iniziale di dati, ma essi sono raccolti da alcune simulazioni. In queste ipotesi, l'algoritmo non ricerca una correlazione tra i dati, quanto piuttosto un 'modo di comportarsi' (*policy*). La linea di condotta dell'algoritmo è ciò che determina quale azione intraprendere alla luce di determinate circostanze ed è appresa attraverso il citato sistema di ricompense e punizioni. Ogni sua azione riceve un certo numero di ricompense prestabilite dal programmatore. Attraverso l'esperienza impara il modo di comportarsi che gli consente di massimizzare le ricompense[62].

Concludendo sul punto, dobbiamo tornare sul concetto di apprendimento profondo (o *deep learning*).

---

[59] D. DANKS, *Learning*, FRANKISH K., RAMSEY W. M. (a cura di), *The Cambridge Handbook of Artificial Intelligence*, Cambridge University Press, p. 154.
[60] WITTEN I. H., FRANK E., HALL M.A., *Data mining. Practical Machin Learning Tools and Techniques*, cit., p. 3
[61] D. DANKS, *Learning*, cit., p. 154.
[62] L. MC ARTHUR, *Machine Learning for Philosophers*, cit., p. 5.

La traduzione letterale è apprendimento profondo ma il Deep Learning, sotto categoria del Machine Learning e del più ampio mondo dell'Intelligenza Artificiale, sottende a qualcosa di molto più ampio del "semplice" apprendimento su più livelli delle macchine. Il termine *deep learning* indica un tipo di apprendimento che riguarda le reti neurali organizzate su più livelli. È questo il metodo alla base delle più recenti ed innovative applicazioni (*self-driving cars*, *chatbots* etc.).

Da un punto di vista scientifico, potremmo dire che il Deep Learning è l'apprendimento da parte delle "macchine" attraverso dati appresi grazie all'utilizzo di algoritmi (prevalentemente di calcolo statistico).

Il Deep Learning, infatti, fa parte di una più ampia famiglia di metodi di Machine Learning basati sull'assimilazione di rappresentazioni di dati, al contrario di algoritmi per l'esecuzione di task specifici.

Le architetture di Deep Learning sono per esempio state applicate nella computer vision, nel riconoscimento automatico della lingua parlata, nell'elaborazione del linguaggio naturale, nel riconoscimento audio e nella bioinformatica.

Raccogliendo le diverse interpretazioni di alcuni tra i più noti ricercatori e scienziati nel campo dell'apprendimento profondo, quali Andrew Yan-Tak Ng, fondatore di Google Brain, Chief Scientist di Baidu e professore e Director dell'AI Lab alla Stanford University, Ian J. Goodfellow, ricercatore riconosciuto come uno dei migliori innovatori del mondo – under 35 – dal MIT di Boston, Yoshua Bengio, uno degli scienziati più riconosciuti proprio nel campo del Deep Learning, Ilya Sutskever, Reasearch Director di OpenAI, Geoffrey Everest Hinton, una delle figure chiave del Deep Learning e dell'Intelligenza artificiale, primo ricercatore ad aver dimostrato l'uso di un algoritmo di backpropagation generalizzato per l'addestramento di reti neurali multistrato, potremmo definire il Deep Learning come un sistema che sfrutta una classe di algoritmi di apprendimento automatico che:

1) usano vari livelli di unità non lineari a cascata per svolgere compiti di estrazione di caratteristiche e di trasformazione. Ciascun livello successivo utilizza l'uscita del livello precedente come input. Gli algoritmi possono essere sia di tipo supervisionato sia non supervisionato e le applicazioni includono l'analisi di pattern (apprendimento non supervisionato) e classificazione (apprendimento supervisionato);
2) sono basati sull'apprendimento non supervisionato di livelli gerarchici multipli di caratteristiche (e di rappresentazioni) dei dati. Le caratteristiche di più alto livello

vengono derivate da quelle di livello più basso per creare una rappresentazione gerarchica;

3) fanno parte della più ampia classe di algoritmi di apprendimento della rappresentazione dei dati all'interno dell'apprendimento automatico (Machine Learning);
4) apprendono multipli livelli di rappresentazione che corrispondono a differenti livelli di astrazione; questi livelli formano una gerarchia di concetti.

Applicando il Deep Learning, avremo quindi una "macchina" che riesce autonomamente a classificare i dati ed a strutturarli gerarchicamente, trovando quelli più rilevanti e utili alla risoluzione di un problema (esattamente come fa la mente umana), migliorando le proprie prestazioni con l'apprendimento continuo.

L'apprendimento profondo (deep learning), come detto, basa il suo funzionamento sulla classificazione e "selezione" dei dati più rilevanti per giungere ad una conclusione, esattamente come fa il nostro cervello biologico che per formulare una risposta ad un quesito, dedurre un'ipotesi logica, arrivare alla risoluzione di un problema, mette in moto i propri neuroni biologici e le connessioni neurali.

Il Deep Learning si comporta allo stesso modo e sfrutta le reti neurali artificiali, modelli di calcolo matematico-informatici basati sul funzionamento delle reti neurali biologiche, ossia modelli costituiti da interconnessioni di informazioni.

Una rete neurale di fatto si presenta come un sistema "adattivo" in grado di modificare la sua struttura (i nodi e le interconnessioni) basandosi sia su dati esterni sia su informazioni interne che si connettono e passano attraverso la rete neurale durante la fase di apprendimento e ragionamento.

Con il Deep Learning vengono simulati i processi di apprendimento del cervello biologico attraverso sistemi artificiali (le reti neurali artificiali, appunto) per insegnare alle macchine non solo ad apprendere autonomamente ma a farlo in modo più "profondo" come sa fare il cervello umano dove profondo significa "su più livelli" (vale a dire sul numero di layer nascosti nella rete neurale – chiamati hydden layer: quelle "tradizionali" contengono 2-3 layer, mentre le reti neurali profonde possono contenerne oltre 150).

Le reti neurali profonde sfruttano un numero maggiore di strati intermedi (hydden layer) per costruire più livelli di astrazione, proprio come si fa nei circuiti booleani (modello matematico di computazione usato nello studio della teoria della complessità computazionale che, in informatica, afferisce alla teoria della computabilità ossia studia

le risorse minime necessarie – principalmente tempo di calcolo e memoria – per la risoluzione di un problema).
Un semplicissimo quanto efficace esempio per capire il reale funzionamento di un sistema di Machine Learning (e la differenza con un sistema di Deep Learning) ci viene fornito da Tech Target[63]:
«Mentre gli algoritmi di apprendimento automatico tradizionali sono lineari, gli algoritmi di apprendimento profondo sono impilati in una gerarchia di crescente complessità e astrazione. Per capire l'apprendimento profondo, immaginiamo un bambino la cui prima parola è "cane". Il bambino impara cos'è un cane (e cosa non lo è) indicando oggetti e dicendo la parola cane. Il genitore dice "Sì, quello è un cane" o "No, non è un cane". Mentre il bambino continua a puntare agli oggetti, diventa più consapevole delle caratteristiche che tutti i cani possiedono. Ciò che il bambino fa, senza saperlo, è chiarire un'astrazione complessa (il concetto di cane) costruendo una gerarchia in cui ogni livello di astrazione viene creato con la conoscenza che è stata acquisita dallo strato precedente della gerarchia».

A differenza del bambino, che impiegherà settimane o addirittura mesi per comprendere il concetto di "cane" e lo farà con l'aiuto del genitore (quello che viene definito apprendimento supervisionato), una applicazione che utilizza algoritmi di Deep Learning può mostrare e ordinare milioni di immagini, identificando con precisione quali immagini contengono i set di dati, in pochi minuti pur non avendo avuto alcun tipo di indirizzamento sulla correttezza o meno dell'identificazione di determinate immagini nel corso del training.
Solitamente, nei sistemi di Deep Learning, l'unica accortezza degli scienziati è "etichettare" i dati (con i meta tag), per esempio inserendo il meta tag "cane" all'interno delle immagini che contengono un cane ma senza spiegare al sistema come riconoscerlo: è il sistema stesso, attraverso livelli gerarchici multipli, che intuisce cosa caratterizza un cane (le zampe, la cosa, il pelo, ecc.) e quindi come riconoscerlo.
Questi sistemi si basano, in sostanza, su un processo di apprendimento "trial-and-error" ma perché l'output finale sia affidabile sono necessarie enormi quantità di dati. Pensare subito ai Big Data e alla facilità con cui oggi si producono e distribuiscono dati di qualsiasi forma e da qualsiasi fonte come facile risoluzione sarebbe però un errore: l'accuratezza

[63] E. BURNS, K. BRUSH, In-depth guide to machine learning in the enterprise, What is Deep Learning and How Does It Work? - TechTarget

dell'output richiede, almeno nella prima fase di addestramento, l'utilizzo di dati "etichettati" (contenenti dei meta tag) che significa che l'utilizzo di dati non strutturati potrebbero rappresentare un problema. I dati non strutturati possono essere analizzati da un modello di apprendimento profondo una volta formato e raggiunto un livello accettabile di accuratezza, ma non per la fase di training del sistema.

Non solo, i sistemi basati su Deep Learning sono difficili da addestrare a causa del numero stesso di strati nella rete neurale. Il numero di strati e collegamenti tra i neuroni nella rete è tale che può diventare difficile calcolare le "regolazioni" che devono essere apportate in ogni fase del processo di addestramento (un problema indicato come problema della scomparsa del gradiente); questo perché per il training comunemente si usano i cosiddetti algoritmi di retropropagazione dell'errore (backpropagation) attraverso il quale si rivedono i pesi della rete neurale (le connessioni tra i neuroni) in caso di errori (la rete propaga all'indietro l'errore in modo che i pesi delle connessioni vengano aggiornati in modo più appropriato). Un processo che continua in modo iterativo finché il gradiente (l'elemento che dà la direzione verso cui l'algoritmo deve muoversi) è nullo.

Nonostante le problematiche che abbiamo illustrato, i sistemi di Deep Learning hanno compiuto enormi passi evolutivi e sono migliorati moltissimo negli ultimi cinque anni, soprattutto per la grandissima quantità di dati a disposizione ma soprattutto per la disponibilità di infrastrutture ultra performanti.

Nell'ambito della ricerca sull'Artificial Intelligence, l'apprendimento automatico ha riscosso un notevole successo negli ultimi anni, consentendo ai computer di superare o avvicinarsi alle prestazioni umane corrispondenti in aree che vanno dal riconoscimento facciale al riconoscimento vocale e linguistico. L'apprendimento profondo invece consente ai computer di fare un passo in avanti, in particolare di risolvere una serie di problemi complessi.

Già oggi ci sono casi d'uso ed ambiti di applicazione che possiamo notare anche come "comuni cittadini" non esperti di tecnologia. Dalla computer vision per le auto senza conducente, fino robot droni impiegati per la consegna di pacchi o anche per l'assistenza in casi di emergenza (per esempio per la consegna di cibo o sangue per trasfusioni in zone terremotate, alluvionate o in zone che devono affrontare crisi epidemiologiche, ecc.); riconoscimento e sintesi vocale e linguistica per chatbot e robot di servizio; riconoscimento facciale per sorveglianza in paesi come la Cina; riconoscimento immagini per aiutare i radiologi a individuare i tumori nei raggi X, oppure per aiutare i ricercatori a individuare le sequenze genetiche correlate alle malattie e identificare le molecole che

potrebbero portare a farmaci più efficaci o addirittura personalizzati; sistemi di analisi per la manutenzione predittiva su una infrastruttura o un impianto analizzando i dati dei sensori dell'IoT; e ancora, la visione del computer che rende possibile il supermercato Amazon Go senza cassa.

Guardando invece ai tipi di applicazione[64] (intesi come compiti che una macchina può svolgere grazie al Deep Learning), quelli di seguito sono quelli ad oggi più maturi:

1) colorazione automatica delle immagini in bianco e nero (per la rete neurale significa riconoscere bordi, sfondi, dettagli e conoscere i colori tipici di una farfalla, per esempio, sapendo esattamente dove collocare il colore corretto);
2) aggiunta automatica di suoni a filmati silenziosi (per il sistema di Deep Learning significa sintetizzare suoni e collocarli correttamente all'interno di una situazione particolare riconoscendo immagini ed azioni, per esempio inserendo il suono del martello pneumatico, della rottura dell'asfalto e i sottofondi di una strada cittadina trafficata in un video in cui si vedono dei lavoratori che stanno rompendo l'asfalto con il martello pneumatico);
3) traduzione simultanea (per il sistema di Deep Learning significa ascoltare e riconoscere il linguaggio naturale, riconoscere la lingua parlata e tradurre il significato in un'altra lingua);

[64] Esempi di utilizzo in diversi ambiti:
- Driverless car
Traffic Sign Detection (TDS) è una funzione presente su molte auto di nuova fabbricazione che permette di riconoscere i segnali stradali. Si tratta di una applicazione di machine learning che utilizza le reti neurali convoluzionali e framework come Tensorflow.
- Cinema
All'inizio del 2021 è stata presentata una nuova metodologia di AI applicabile nella produzione cinematografica. Il nuovo approccio si avvale di diversi strumenti di deep learning in parallelo e in serie (VGG16, MLP, transfer learning) e adopera differenti tipologie di dataset (immagini con feature diverse messe in risalto) al fine di poter classificare in modo accurato le inquadrature cinematografiche e permetterne così un utilizzo professionale e operativo nel processo di film making o nell'attività di indicizzazione dei contenuti streaming. Parliamo di AI applicata all'image processing.
- Quantum Intelligence
Sul quantum computing si concentrano grandi aspettative. Un paradigma di elaborazione che richiede nuovo hardware, nuovi algoritmi e nuove soluzioni. La caratteristica principale è la capacità di semplificare molto la soluzione di alcuni problemi, riducendone la complessità esponenziale. Un esempio: la determinazione dei fattori di un numero, alla base di tanti problemi di crittografia e su cui si fondano moltissime applicazioni di sicurezza informatica.
- Ologrammi
Uno studio condotto sugli ologrammi nel 2021 dai ricercatori del Massachussets Institute of Technology (MIT) ha provato come, grazie a una nuova tecnica di deep learning denominata "olografia tensoriale", è possibile generare un video olografico in modo istantaneo, sfruttando la capacità computazionale presente su un semplice computer. La peculiarità è quella di essere composta da tensori addestrabili, che possono apprendere come elaborare le informazioni visive e di profondità in modo simile a quanto fa il cervello umano.

4) classificazione degli oggetti all'interno di una fotografia (il sistema in questo caso è in grado di riconoscere e classificare tutto ciò che vede in un'immagine, anche molto complessa dove per esempio c'è un paesaggio di sfondo, per esempio delle montagne, persone che camminano lungo un sentiero, degli animali al pascolo, ecc.);
5) generazione automatica della grafia (ci sono già oggi sistemi di Deep Learning capaci di utilizzare la grafia umana per scrivere addirittura apprendendo gli stimi della scrittura a mano degli esseri umani ed imitandola);
6) generazione automatica di testo (sono sistemi che hanno imparato a scrivere correttamente in una determinata lingua rispettando ortografia, punteggiatura, grammatica e persino imparando ad usare stili di scrittura differenti a seconda dell'output da produrre, per esempio un articolo giornalistico o una novella);
7) generazione automatica di didascalie (in questo caso il riconoscimelo delle immagini, l'analisi del contesto e la capacità di scrittura consentono ad un sistema di scrivere in automatico le didascalie di una immagine descrivendone perfettamente la scena);
8) gioco automatico (abbiamo imparato a capire le potenzialità di un sistema in grado di imparare autonomamente come giocare a un determinato gioco grazie a DeepMind – oggi parte di Google – che ha sviluppato un sistema di Deep Learning – AlphaGo – che non solo ha imparato a giocare al complessissimo gioco Go ma è riuscito anche a battere il campione mondiale, umano).

3. I nuovi orizzonti della tecnologia informatica e della robotica

Sempre più spesso si sente parlare di intelligenza artificiale, applicata con successo in vari settori. Da quelli della medicina, della difesa e dei trasporti, a quello della produzione a livello industriale di robot umanoidi e, cioè, di macchine automatiche in grado di aiutarci nei lavori più disparati[65].

Nel campo della medicina, già oggi, nel mondo, si utilizzano robot chirurgici, che altro non sono che estensioni delle mani del chirurgo controllate da remoto, i quali, fornendo ad esso un feedback sensoriale analogo a quello che avrebbe se operasse direttamente sul paziente, consentono di effettuare interventi più precisi e meno invasivi di quelli tradizionali. Si stanno inoltre sviluppando, con successo, speciali tecnologie robotiche finalizzate alla riabilitazione di pazienti affetti da gravi menomazioni, paralisi e malattie

[65] Per un approfondimento dei molteplici aspetti dell'intelligenza artificiale nelle sue svariate applicazioni, cfr. D. HEAVEN, in AA.VV., *Macchine che pensano*, Dedalo, Bari, 2018.

degenerative nonché all'impianto in soggetti che hanno subito mutilazioni, di protesi direttamente collegate tramite elettrodi a nervi, muscoli e tendini della parte amputata, capaci di consentire al paziente il recupero non solo del movimento, ma anche di gran parte delle sensazioni tattili. Di recente, per la prima volta al mondo, un gruppo di scienziati della Scuola Superiore Sant'Anna di Pisa e del Sahlgrenska University Hospital di Göteborg ha realizzato, su una persona mutilata, un impianto trans radiale stabile e permanente di una "mano robot" in grado di captare e tradurre in movimento i segnali mioelettrici provenienti dal cervello. Sempre nel campo medico-chirurgico, è inoltre in via di sviluppo un endoscopio miniaturizzato, capace di muoversi all'interno del corpo umano, dotato di sensori capaci, a fini diagnostici, di rilevare parametri biologici, ovvero, a fini terapeutici, di rilasciare in loco farmaci o effettuare interventi chirurgici.

Anche nel settore della difesa esistono già varie tipologie di robot: veicoli senza equipaggio per sorvegliare installazioni militari, droni controllati da remoto capaci di individuare obiettivi e colpire bersagli mobili senza supervisione da parte dell'uomo, robot-killer e, cioè, sistemi d'arma letali autonomi destinati a operare al posto dei soldati in missioni particolarmente rischiose.

Nel campo dei trasporti, già da tempo sono in uso in alcune città treni metropolitani a guida automatizzata, come anche le auto driverless, vale a dire veicoli a guida automatica, capaci non solo di scegliere il percorso migliore per giungere alla meta (come, del resto, già oggi è possibile facendo uso del "navigatore"), ma soprattutto di muoversi nel traffico in condizioni di assoluta sicurezza. Si tratta, cioè, di veri e propri robot, dotati di telecamere e sensori vari nonché di un'intelligenza artificiale che consente ad essi di adeguare velocità e distanza di sicurezza in relazione al tipo di strada, di traffico, di condizioni meteorologiche e di attivare la frenata d'emergenza in presenza di un ostacolo improvviso.

In breve, possiamo ben dire che quello che fino a qualche anno fa apparteneva al mondo della fantascienza, oggi è un fatto concreto.

Ai suoi albori la robotica costituiva una disciplina ricompresa nell'ambito dell'ingegneria meccanica, specificamente dedicata allo studio di macchine che fossero in grado di compiere azioni nel mondo reale, come afferrare oggetti, muoversi nello spazio, e così via[66]. Nei primi anni di sviluppo, i sistemi di controllo di tali macchine erano basati su

[66] Per una dettagliata storia dello sviluppo della robotica, cfr. P. HUSBANDS, *Robotics*, in FRANKISH K., RAMSEY W. M. (a cura di), *The Cambridge Handbook of Artificial Intelligence*, Cambridge University Press, 2014, p. 269 ss.

algoritmi e non prevedevano componenti di intelligenza artificiale. Man mano che migliorarono le capacità dei robot, si manifestò la necessità di dotarli di sistemi di controllo più intelligenti e sorse la robotica cognitiva che coinvolge componenti basate sull'intelligenza artificiale.

Oggi, pur continuando a costituire discipline distinte, i campi di ricerca della robotica e dell'intelligenza artificiale presentano notevoli sovrapposizioni[67]. Ciò appare, del resto, coerente con lo sviluppo avvenuto in seno allo stesso settore dell'intelligenza artificiale che, dagli originari sistemi incapaci di percepire e di agire sull'ambiente circostante, ha rivolto una crescente attenzione a sistemi di intelligenza artificiale capaci di percepire e di agire direttamente sull'ambiente circostante, dei quali i robot costituiscono appunto un esempio.

Come per l'intelligenza artificiale[68] non esiste una definizione condivisa di che cosa sia un 'robot'[69].

La nozione tradizionale di robot ruota intorno al requisito 'fisico' del cd. embodiment (che potremmo tradurre con il termine italiano 'incorporazione'), da intendersi come la capacità della macchina di interagire con il mondo reale[70]. Il requisito della componente fisica parrebbe ancora da ritenersi essenziale, anche se la sua centralità sembra essere messa in crisi o, quantomeno, adombrata dall'incrementale sviluppo delle sue capacità cognitive[71].

Un criterio utile per identificare il concetto di robot e per esaltarne i profili distintivi nei confronti degli altri strumenti tecnologici è offerto dal modello cd. sense- think-act[72], che definisce un sistema robotico come «a machine, situated in the world, that senses, thinks,

---

[67] FRANKLIN S., *History, motivations, and core themes*, cit., p. 27.

[68] Sul concetto di "intelligenza artificiale" e relativi ambiti applicativi, cfr. P. MELLO, *Intelligenza artificiale*, in *http://disf.org/intelligenza-artificiale (Documentazione Interdisciplinare di Scienza & Fede)*, 2002, nonché il saggio di J. BERNSTEIN, *Uomini e macchine intelligenti*, Adelphi, Milano, 2013.

[69] Cfr. C. LEROUX, R. LABRUTO et a., *A Green Paper on Legal Issues in Robotics*, euRobotics CA, public report, 2012, p. 10, che invece di dare una definizione onnicomprensiva di robot preferiscono osservare che cosa un robot è grado di fare; nonché E. PALMERINI et a., *Guidelines on regulating robotics*, The RoboLaw Project, 2014, p. 15, i quali muovono da una tassonomia dei diversi tipi di robot sulla base di alcune loro caratteristiche fondamentali.

[70] La rilevanza del requisito della 'fisicità' nella definizione di robot è ben messa in evidenza da Russel e Norvig, secondo i quali «robots are *physical agents* that perform tasks by *manipulating the physical world*», (corsivi nostri), cfr. S. J. RUSSEL, P. NORVIG, *Artificial Intelligence*, cit. p. 971.

[71] In questi termini A. SANTOSUOSSO, B. BOTTALICO, *Autonomous Systems and the Law: Why Intelligence Matters. A European Perspective*, in HILGENDORF E., SEIDEL U., *Robotics, Autonomics and the Law*, 2017, p. 27, i quali preferiscono utilizzare la più ampia categoria di «*autonomous systems*».

[72] G. A. BEKEY, *Current Trends in Robotics: Technology and Ethics*, in LIN P., ABNEY K., BEKEY G. A., *Robot Ethics: The Ethical and Social Implications of Robotics*, 2011, p. 17 ss; R. CALÒ, *Robotics and the Lessons of Cyberlaw*, in *California Law Review*, 2015, p. 529 ss; G.; E. PALMERINI, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, in *Resp. Civile e Previdenza*, 2016, p. 1815 ss.

and acts»[73]. In questa prospettiva, la nozione di robot richiede la combinazione di tutte le indicate qualità, ma ciascuna di esse può essere presente secondo un diverso grado di sviluppo.

Anzitutto, il robot deve vivere nel mondo reale. L'intelligenza artificiale può infatti occupare "diverse tipi di corpi", dando vita ad altrettante tipologie di agenti autonomi o intelligenti[74]. Il requisito citato serve pertanto a distinguere i robot da altri agenti come gli agenti software (o software robots o softbots) i quali, ancorché dotati di una certa autonomia o intelligenza, non possiedono un corpo fisico.

Il secondo elemento è la capacità di percepire l'ambiente esterno attraverso l'utilizzo di sensori. Essi consentono al robot di misurare la distanza dagli oggetti vicini (es. telecamere, radar, lidar etc.), di determinare la sua collocazione spaziale (es. GPS), di ottenere informazioni sul proprio stesso movimento.

La capacità di esercitare azioni sull'ambiente circostante è invece realizzata attraverso dei cd. effettori[75] o attuatori (come gambe, ruote, giunture meccaniche etc.) che consentono un mutamento dell'ambiente circostante.

In relazione a tale elemento si pone il problema di determinare il perimetro del concetto di azione e, in particolare, come debba interpretarsi la capacità di modificare l'ambiente esterno. In un certo senso, infatti, anche la vista di un film e altri stimoli come i social bot sono in grado di suscitare delle emozioni e di provocare delle risposte fisiologiche[76]. In un'ottica operativa, è possibile ritenere che un sistema agisce sull'ambiente quando è in grado di modificarlo direttamente e cioè di esercitare una forza fisica su di esso, dovendosi dunque distinguere il concetto di azione da quello di informazione[77].

Infine, l'elemento centrale della definizione consiste nella capacità di pensare, da intendersi come la capacità di processare le informazioni provenienti dai sensori e di assumere decisioni autonome. In questa particolare accezione, la nozione di robot viene associata al carattere di autonomia, dovendosi pertanto escludere quelle macchine del tutto tele-operate da un attore umano[78].

---

[73] G. A. BEKEY, *Current Trends in Robotics: Technology and Ethics*, cit., p. 18.

[74] W. BARFIELD, *Towards a law of artificial intelligence*, in BARFIELD W., PAGALLO U., *Research Handbook on the Law of Artificial Intelligence*, Ed. Edward Elgar Publishing, 2018, p. 15. p. 12.

[75] Questa la terminologia utilizzata da S.J. Russel e P. Norvig. *Ivi*, p. 971.

[76] R. CALÒ, *Robotics and the Lessons of Cyberlaw*, cit., pp. 530-531.

[77] Per una critica a questa impostazione cfr. A. SANTOSUOSSO, B. BOTTALICO, *Autonomous Systems and the Law: Why Intelligence Matters. A European Perspective*, cit., p. 32 ss.

[78] G. A. BEKEY, *Current Trends in Robotics: Technology and Ethics*, cit., p. 18.

4. Il contesto normativo in materia di intelligenza artificiale e robotica

4.1. Il quadro normativo europeo in materia di robotica

L'attività di ricognizione delle norme applicabili a questo settore non è semplice, posto che si tratta di una materia evidentemente trasversale. Esiste, infatti, una pletora di norme applicabili al tema dell'intelligenza artificiale e dalla robotica che spaziano dalle leggi applicabili ai droni, alle leggi del diritto internazionale di guerra per quanto concerne le armi automatiche, alle leggi di regolamentazione del mercato finanziario etc.[79]

Il quadro normativo europeo che assume rilievo nell'ambito della robotica può essere idealmente rappresentato attraverso l'immagine di tre cerchi concentrici: il cerchio più interno è costituito dalla direttiva 2006/42/CE (cd. 'direttiva macchine') che, pur riguardando nello specifico il settore delle macchine, può ricomprendere i robot nella loro veste di artefatti meccanici; il cerchio intermedio comprende le misure relative alla tutela della salute, della sicurezza e degli interessi dei consumatori, tra cui rientrano la direttiva 2001/95/CE, la decisione 768/2008/CE e la decisione 765/2008/CE; il cerchio più esterno comprende infine la direttiva 1999/44/CE, che riguarda la vendita dei beni di consumo[80].

A questa normativa di carattere generale si affianca una possibile regolazione settoriale dovuta alla appartenenza del prodotto a classi più specifiche come accade, ad esempio, per i prodotti robotici che si inseriscono nell'ambito medico.

Esistono infine categorie di prodotti che, al momento attuale, sono sprovvisti di una apposita regolazione e che necessitano di uno specifico intervento da parte del legislatore.

4.1.1. Il cerchio interno: la 'direttiva macchine'

La direttiva macchine si inserisce nel più ampio contesto di armonizzazione dei requisiti di salute e sicurezza dei prodotti, al fine di assicurarne la libera circolazione all'interno del mercato europeo[81].

---

[79] A. SANTOSUOSSO, C. BOSCARATO, F. CAROLEO, *Robot e diritto: una prima ricognizione*, in *Nuova Giurisprudenza Civile Commentata*, 2012, p. 4; E. PALMERINI, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, cit., p. 1815 ss.

[80] Cfr. C. LEROUX, R. LABRUTO et a., *A Green Paper on Legal Issues in Robotics*, cit., p. 21 ss.; nonché, più sinteticamente, A. SANTOSUOSSO, C. BOSCARATO, F. CAROLEO, *Robot e diritto: una prima ricognizione*, cit., p. 8; A. SANTOSUOSSO, B. BOTTALICO, *Autonomous Systems and the Law: Why Intelligence Matters. A European Perspective*, cit., p. 35 ss.

[81] A. SANTOSUOSSO, C. BOSCARATO, F. CAROLEO, *Robot e diritto: una prima ricognizione*, cit., p. 8.

In base all'art. 1, le disposizioni della direttiva trovano applicazione per le seguenti categorie di prodotti: macchine; attrezzature intercambiabili; componenti di sicurezza; accessori di sollevamento; catene, funi e cinghie; dispositivi amovibili di trasmissione meccanica; quasi- macchine.

I sistemi robotici ricadono nel concetto di «*macchine*», che l'art. 2 del testo normativo definisce come:

«*a) -insieme equipaggiato o destinato ad essere equipaggiato di un sistema di azionamento diverso dalla forza umana o animale diretta, composto di parti o di componenti, di cui almeno uno mobile, collegati tra loro solidamente per un'applicazione ben determinata,*

- insieme di cui al primo trattino, al quale mancano solamente elementi di collegamento al sito di impiego o di allacciamento alle fonti di energia e di movimento,

- insieme di cui al primo e al secondo trattino, pronto per essere installato e che può funzionare solo dopo essere stato montato su un mezzo di trasporto o installato in un edificio o in una costruzione,

- insiemi di macchine, di cui al primo, al secondo e al terzo trattino, o di quasi-macchine, di cui alla lettera g)[82]*, che per raggiungere uno stesso risultato sono disposti e comandati in modo da avere un funzionamento solidale,*

- insieme di parti o di componenti, di cui almeno uno mobile, collegati tra loro solidalmente e destinati al sollevamento di pesi e la cui unica fonte di energia è la forza umana diretta».

La direttiva pone una serie di obblighi per i produttori o i loro rappresentanti autorizzati da osservare prima di poter mettere sul mercato una macchina.

Tra di essi rientrano: a) l'obbligo di soddisfare gli essenziali requisiti di sicurezza e di tutela della salute; b) l'obbligo di fornire le informazioni necessarie; c) l'obbligo di sottoporre le macchine alle procedure di valutazione della conformità disciplinate dall'art. 12 della direttiva.

[82] Ai sensi dell'art. 2 *lett*. g) per per «*quasi macchine*» si devono intendere quegli «*insiemi che costituiscono quasi una macchina, ma che, da soli, non sono in grado di garantire un'applicazione ben determinata. Un sistema di aziona- mento è una quasi-macchina. Le quasi-macchine sono unicamente destinate ad essere incorporate o assemblate ad altre macchine o ad altre quasi-macchine o apparecchi per costituire una macchina disciplinata dalla presente direttiva*».

All'esito delle procedure di valutazione, il produttore può redigere la dichiarazione di conformità e apporre la marcatura «CE»[83] ex art. 16. Attraverso l'apposizione della marcatura il produttore si assume la responsabilità circa la conformità del prodotto[84].

4.1.2. Il cerchio maggiore ed il cerchio esterno: le disposizioni sulla sicurezza generale dei prodotti e sulla vendita dei beni di consumo.

Alle disposizioni più specifiche contenute nella direttiva macchine si aggiungono le più generali previsioni normative di cui alla direttiva 2001/95/CE sulla sicurezza generale dei prodotti e della direttiva 1999/44/CE sulla vendita dei beni di consumo. Poiché il tema sarà ripreso più dettagliatamente nel prosieguo, ci limitiamo qui ai tratti essenziali.

La direttiva 2001/95/CE impone un requisito generale di sicurezza per ogni prodotto messo sul mercato e destinato al consumo. La disciplina ivi stabilita si applica a tutte le tipologie di prodotto (cd. legislazione orizzontale), ma si interseca con le previsioni normative specificamente rivolte a determinate categorie di beni (macchine, cosmetici, autoveicoli etc., cd. legislazione verticale)[85]. L'aspetto più saliente della direttiva - e, più in generale, delle direttive relative alla sicurezza dei prodotti - consiste nella particolare rilevanza attribuita alla normazione di carattere tecnico, vale a dire a quelle regole di carattere tecnico-scientifico elaborate da organismi privati denominati «enti di normalizzazione» (UNI e CEI in Italia, DIN in Germania etc.)[86].

La direttiva 1999/44/CE, infine, riguarda l'armonizzazione delle leggi, dei regolamenti e dei provvedimenti amministrativi degli Stati membri dell'Unione Europea in merito ad alcuni aspetti relativi alla vendita dei beni di consumo e alle relative garanzie per raggiungere un livello minimo uniforme di tutela del consumatore nell'ambito del mercato comunitario[87]. Tra i tratti più significativi della direttiva vi è la previsione dell'obbligo per il venditore di consegnare un bene «conforme» a quanto stabilito nel contratto di vendita, con l'introduzione della correlata nozione di «difetto di conformità»

[83] Per la marcatura «CE» cfr. reg. n. 765/2008/CE.

[84] A. SANTOSUOSSO, C. BOSCARATO, F. CAROLEO, *Robot e diritto: una prima ricognizione*, cit., p. 9.

[85] E. AL MUREDEN, *La responsabilità del fabbricante nella prospettiva della standardizzazione delle regole sulla sicurezza dei prodotti*, in AL MUREDEN E., *La sicurezza dei prodotti e la responsabilità del produttore: casi e materiali*, Giappichelli, Torino, 2017, p. 10.

[86] U. CARNEVALI, *Prevenzione e risarcimento nelle direttive comunitarie sulla sicurezza dei prodotti*, or in AL MUREDEN E., cit., 10.

[87] A. SANTOSUOSSO, C. BOSCARATO, F. CAROLEO, *Robot e diritto: una prima ricognizione*, cit., p. 12.

e la previsione di alcuni specifici rimedi per l'acquirente, tra cui quello della richiesta di riparazione o sostituzione del bene[88].

4.1.3. Settori specifici VS. assenza di regolazione

I sistemi robotici sono poi tendenzialmente sottoposti ad una *regolazione settoriale* che li riguarda come classe di prodotti specifici.

Come accennato in premessa, è questo il caso dei robot che costituiscono dispositivi medici (protesi, nanocapsule, robot chirurgici), la cui regolamentazione è contenuta nella direttiva 1993/42/CEE sui dispositivi medici e nella direttiva 1990/385/CEE sui dispositivi medici impiantabili attivi[89].

Esistono infine dei sistemi robotici emergenti la cui disciplina è, al momento attuale, *assente o in fase di gestazione*. Come buon esempio della categoria, soffermiamoci brevemente sui recenti sviluppi della regolamentazione delle auto a guida autonoma. Ciò ci consentirà anche di evidenziare quale sia il livello di sviluppo tecnologico attualmente preso in considerazione a livello legislativo.

La circolazione stradale è oggetto della Convenzione sulla circolazione stradale, conclusa a Vienna l'8 novembre del 1968. Fino a poco tempo fa, alcune previsioni dell'articolato normativo (in particolare, l'art. 8[90]) escludevano la circolazione di auto a guida autonoma su strade pubbliche, imponendo che ogni veicolo dovesse essere costantemente sottoposto al controllo del guidatore[91]. Il testo è tuttavia stato recentemente emendato nel senso di consentire l'introduzione dei nuovi mezzi tecnologici, pur continuandosi a richiedere la possibilità per il guidatore di neutralizzare il sistema automatizzato[92].

---

[88] In generale cfr. M. PALADINI, *I rimedi al difetto di conformità nella vendita di beni di consumo*, in E. TOSI, *La tutela dei consumatori in internet e nel commercio elettronico*, 2012, p. 351 ss.

[89] E. PALMERINI, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, cit., p. 1815 ss.

[90] Art. 8 comma 1 «*Ogni veicolo in movimento o ogni complesso di veicoli in movimento deve avere un conducente*»; comma 5 «*Ogni conducente deve avere costantemente il controllo del proprio veicolo o deve poter guidare i propri animal*i».

[91] E. HILGENDORF, *Automated Driving and the Law*, cit., p. 179; N. VELLINGA, *Automated Driving and the Future of Traffic Law*, cit., p. 77;

[92] Art. 8 comma 5 *bis* «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», introdotto dagli emendamenti del 26 marzo 2014, in vigore dal 23 marzo 2016. Per una discussione sul punto cfr. S. LUTZ, *Autonome Fahrzeuge als rechtliche Herausforderung*, in *NJW*, 2015, p. 119 ss.

In termini generali, l'implementazione legale delle auto a guida autonoma segue due passaggi legislativi: l'autorizzazione a svolgere test su strada per lo sviluppo dei suddetti veicoli e la regolazione della loro effettiva immissione in mercato.

Il primo Paese europeo a introdurre una regolamentazione è stata la *Germania*[93]. La legge tedesca, che ha introdotto modifiche al codice della strada, si riferisce solo a situazioni in cui il conducente è in grado di riprendere il controllo sul sistema tecnico di guida (cioè ai veicoli fino a livello 4) e non, invece, ai veicoli nei quali il conducente non ha possibilità di esercitare alcun controllo (livello 5). In altri termini, la legge disciplina la guida *automatizzata*, non la guida *autonoma*. Ciò tuttavia non esclude che vi possano essere delle fasi in cui l'attività di guida è svolta interamente dal computer di bordo. In questa prospettiva, la legge stabilisce che il conducente - la cui nozione è estesa fino a ricomprendere colui che sia in grado di attivare una funzione di guida molto o del tutto automatizzata - che attiva il sistema automatico, rispettando la funzione per cui esso è destinato, possa distaccarsi dall'attività di guida, pur dovendo rimanere disponibile a percepire eventuali segnali da parte del sistema che gli richiedano di riprendere il controllo del mezzo ovvero circostanze anomale chiaramente riconoscibili[94]. Quanto invece ai costruttori, essi sono obbligati a dotare i veicoli di una strumentazione idonea a richiamare il conducente, oltre che, più in generale, a realizzare mezzi capaci di seguire le norme sulla circolazione stradale[95]. Inoltre, i veicoli devono essere dotati di sistemi

---

[93] E. HILGENDORF, *Auf dem Weg zu einer Regulierung des automatisierten Fahrens: Anmerkungen zur jüngsten Reform des StVG*, in *KriPoZ*, 2017, p. 225 ss; *ID. Automatisiertes Fahren und Recht – ein Überblick*, in *JA*, 2018, p. 801 ss.; V. LÜDEMANN, C. SUTTER, K. VOGELPOHL, *Neue Pflichten für Fahrzeugführer beim automatisierten Fahren – eine Analyse aus rechtlicher und verkehrspsychologischer Sicht*, in *NZV*, 2018, p. 411 ss.; C. KÖNIG, *Gesetzgeber ebnet Weg für automatisiertes Fahren – weitgehend gelungen*, in *NZV*, 2017, p. 249 ss. Nella letteratura italiana cfr. M. G. LOSANO, *Il progetto di legge tedesco sull'auto a guida automatizzata*, in *Diritto dell'Informazione e dell'Informatica*, 2017, p. 1 ss.

[94] § 1b (StVG) «*Rechte und Pflichten des Fahrzeugführers bei Nutzung hoch-oder vollautomatisierter Fahrfunktionen 1) Der Fahrzeugführer darf sich während der Fahrzeugführung mittels hoch-oder vollautomatisierter Fahrfunktionen gemäß § 1a vom Verkehrsgeschehen und der Fahrzeugsteuerung abwenden; dabei muss er derart wahrnehmungsbereit bleiben, dass er seiner Pflicht nach Absatz 2 jederzeit nachkommen kann. (2) Der Fahrzeugführer ist verpflichtet, die Fahrzeugsteuerung unverzüglich wieder zu übernehmen, 1. wenn das hoch- oder vollautomatisierte System ihn dazu auffordert oder 2. wenn er erkennt oder auf Grund offensichtlicher Umstände erkennen muss, dass die Voraussetzungen für eine bestimmungsgemäße Verwendung der hoch- oder vollautomatisierten Fahrfunktionen nicht mehr vorliegen*». Qualche perplessità ha suscitato in dottrina il fatto che il legislatore non abbia tipizzato le situazioni in cui il conducente è obbligato a riprendere il controllo del mezzo in caso di anomalia. Sul punto cfr. P. STEINERT, *Automatisiertes Fahren (Strafrechtliche Fragen*), in *SVR*, 2019, p. 6. Un giudizio favorevole è peraltro espresso da E. HILGENDORF. *Automatisiertes Fahren und Recht – ein Überblick, cit., p. 802, secondo cui il legislatore ha stabilito una regolamentazione sufficientemente chiara, mentre sarà compito della giurisprudenza precisare ulteriormente quali siano le situazioni che richiedono l'intervento del conducente.*

[95] M. G. LOSANO, *Il progetto di legge tedesco sull'auto a guida automatizzata*, cit., pp. 12-13.

§ 1a (StVG) «*Kraftfahrzeuge mit hoch- oder vollautomatisierter Fahrfunktion (1) Der Betrieb eines Kraftfahrzeugs mittels hoch- oder vollautomatisierter Fahrfunktion ist zulässig, wenn die Funktion bestimmungsgemäß verwendet wird. (2) Kraftfahrzeuge mit hoch- oder vollautomatisierter Fahrfunktion im Sinne*

capaci di memorizzare i dati necessari ad indicare quando il controllo del mezzo sia affidato al conducente e quando invece il controllo sia affidato al sistema automatico[96].
Considerato il ruolo attivo ricoperto dal conducente del veicolo, il legislatore tedesco non ha per il momento inteso introdurre modifiche al suo regime di responsabilità[97]. La legge prevede infine un meccanismo di revisione per cui, in base ai dati relativi ai primi anni di implementazione e ai più recenti sviluppi tecnologici, sarà possibile effettuare degli opportuni emendamenti.
Anche il *Regno Unito* con l'*Automated and Electric Vehicles Act 2018* ha approvato una specifica legislazione concernente le auto a guida autonoma.
La Sec. 1 dà incarico al Segretario di Stato di redigere una lista dei mezzi di trasporto a guida autonoma, definiti come quei veicoli «*capable, in at least some circumstances or situations, of safely driving themselves*» e suscettibili di essere utilizzati su strade pubbliche.
Sul versante della responsabilità, è anzitutto l'assicuratore che la legge chiama a rispondere degli eventuali danni provocati dalla circolazione del veicolo, ivi comprese la morte o il ferimento di terzi (Sec. 2, Par. 1). Ove tuttavia ricorrano determinate circostanze e, in particolare, il veicolo non sia assicurato, la responsabilità ricade sul proprietario del medesimo (Sec. 2, Par. 2). Alcune previsioni si occupano delle ipotesi di *contributory negligence*, tra di esse si segnala la Sec. 3, Par. 2, che esclude la responsabilità dell'assicuratore o del proprietario del veicolo «*where the accident that it*

---

dieses Gesetzes sind solche, die über eine technische Ausrüstung verfügen, 1. die zur Bewältigung der Fahraufgabe – einschließlich Längs- und Querführung – das jeweilige Kraftfahrzeug nach Aktivierung steuern (Fahrzeugsteuerung) kann, 2. die in der Lage ist, während der hoch- oder vollautomatisierten Fahrzeugsteuerung den an die Fahrzeugführung gerichteten Verkehrsvorschriften zu entsprechen, 3. die jederzeit durch den Fahrzeugführer manuell übersteuerbar oder deaktivierbar ist, 4. die die Erforderlichkeit der eigenhändigen Fahrzeugsteuerung durch den Fahrzeugführer erkennen kann, 5. die dem Fahrzeugführer das Erfordernis der eigenhändigen Fahrzeugsteuerung mit ausreichender Zeitreserve vor der Abgabe der Fahrzeugsteuerung an den Fahrzeugführer optisch, akustisch, taktil oder sonst wahrnehmbar anzeigen kann und 6. die auf eine der Systembeschreibung zuwiderlaufende Verwendung hinweist. Der Hersteller eines solchen Kraftfahrzeugs hat in der Systembeschreibung verbindlich zu erklären, dass das Fahrzeug den Voraussetzungen des Satzes 1 entspricht. (3) Die vorstehenden Absätze sind nur auf solche Fahrzeuge anzuwenden, die nach § 1 Absatz 1 zugelassen sind, den in Absatz 2 Satz 1 enthaltenen Vorgaben entsprechen und deren hoch- oder vollautomatisierte Fahrfunktionen 1.in internationalen, im Geltungsbereich dieses Gesetzes anzuwendenden Vorschriften beschrieben sind und diesen entsprechen oder 2. eine Typgenehmigung gemäß Artikel 20 der Richtlinie 2007/46/EG des Europäischen Parlaments und des Rates vom 5. September 2007 zur Schaffung eines Rahmens für die Genehmigung von Kraftfahrzeugen und Kraftfahrzeuganhängern sowie von Systemen, Bauteilen und selbstständigen technischen Einheiten für diese Fahrzeuge (Rahmenrichtlinie) (ABl. L 263 vom 9.10.2007, S. 1) erteilt bekommen haben. (4) Fahrzeugführer ist auch derjenige, der eine hoch- oder vollautomatisierte Fahrfunktion im Sinne des Absatzes 2 aktiviert und zur Fahrzeugsteuerung verwendet, auch wenn er im Rahmen der bestimmungsgemäßen Verwendung dieser Funktion das Fahrzeug nicht eigenhändig steuert.

[96] Cfr. § 63.

[97] L. WÖRNER, *Der Weichensteller 4.0 Zur strafrechtlichen Verantwortlichkeit des Programmierers im Notstand für Vorgaben an autonome Fahrzeuge*, in *ZIS,* 2019, p. 42.

caused was wholly due to the person's negligence in allowing the vehicle to begin driving itself when it was not appropriate to do so».

Con riferimento all'*Italia*, la materia è attualmente regolata dal Decreto 28 febbraio 2018 del Ministero delle Infrastrutture e dei Trasporti. Il Decreto è stato emanato in attuazione dell'art. 1 comma 72 della l. 27 dicembre 2017, n. 205[98], che ha autorizzato la sperimentazione su strada delle soluzioni di *Smart Road* e di guida connessa e automatica. Il testo normativo consente, previa autorizzazione del Ministero dei Trasporti, la sperimentazione su strada delle auto a guida automatizzata (art. 9). Durante la fase di sperimentazione, il veicolo deve essere costantemente sottoposto al controllo di un supervisore. Quest'ultimo deve essere in grado di poter commutare tempestivamente l'operatività del veicolo in modo automatico in operatività dello stesso in modo manuale e conserva la responsabilità del veicolo in entrambe le modalità operative (art. 10).

Sul versante definitorio, l'art. 1 definisce il «*veicolo a guida automatica*» 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*», precisando che «*non è considerato veicolo a guida automatica un veicolo omologato per la circolazione sulle strade pubbliche italiane secondo le regole vigenti e dotato di uno o più` sistemi di assistenza alla guida, che vengono attivati da un guidatore al solo scopo di attuare comportamenti di guida da egli stesso decisi e che comunque necessitano di una continua partecipazione attiva da parte del conducente alla attività` di guida*».

4.1.4. I progetti specifici relativi alla rilevanza dei robot per il diritto

In ambito europeo è attualmente in corso un vivace dibattito relativo alla regolamentazione della robotica. In questa sede ci limitiamo a riportare sinteticamente alcuni dei progetti più recenti e rilevanti per la materia, mentre i loro contenuti verranno

[98] Art. 1 comma, 72 «*Al fine di sostenere la diffusione delle buone pratiche tecnologiche nel processo di trasformazione digitale della rete stradale nazionale (Smart Road) nonché allo scopo di promuovere lo sviluppo, la realizzazione in via prototipale, la sperimentazione e la validazione di soluzioni applicative dinamicamente aggiornate alle specifiche funzionali, di valutare e aggiornare dinamicamente le specifiche funzionali per le Smart Road e di facilitare un'equa possibilità di accesso del mondo produttivo ed economico alla sperimentazione, è autorizzata la sperimentazione su strada delle soluzioni di Smart Road e di guida connessa e automatica. A tale fine, entro trenta giorni dalla data di entrata in vigore della presente legge, con decreto del Ministro delle infrastrutture e dei trasporti, sentito il Ministro dell'interno, sono definiti le modalità attuative e gli strumenti operativi della sperimentazione. Per le finalità di cui al presente comma è autorizzata la spesa di un milione di euro per ciascuno degli anni 2018 e 2019». Cfr. S. SCAGLIARINI, Smart Roads e driverless cars nella legge di bilancio: opportunità e rischi di un'attività economica «indirizzata e coordinata a fini sociali», in Quaderni Costituzionali, pp. 497-500.*

(e in parte già lo sono stati) ripresi nel corso della trattazione. Un'attenzione maggiore sarà invece dedicata alla recente Risoluzione del Parlamento Europeo in materia di robotica.

Nel 2012 i partecipanti al progetto *euRobotics* hanno pubblicato il documento *Suggestion for a green paper on legal issues in robotics*[99], che contiene un'embrionale descrizione delle principali questioni giuridiche che ostacolano lo sviluppo della robotica in Europa, insieme ad alcune raccomandazioni e all'indicazione di possibili strade per affrontarle. Come si legge nel testo, lo scopo del report non è quello di fornire risposte conclusive, quanto piuttosto di stimolare un dibattito sull'argomento. Il documento offre un'analisi interdisciplinare delle caratteristiche essenziali dei nuovi sistemi robotici ed esamina la loro possibile qualificazione giuridica, sia nella prospettiva tradizionale dei robot come oggetto-strumento-prodotto, sia in quella innovativa della possibile istituzione della nuova categoria giuridica della personalità elettronica dei robot.

Sulla stessa scia, nel 2014, è stato rilasciato il documento conclusivo del progetto europeo denominato *RoboLaw*. L'intento del progetto, non dissimile dal precedente, è quello di offrire una approfondita analisi dei problemi etici e legali sollevati dalla robotica e di fornire alcune linee guida per il legislatore europeo e nazionale[100]. La novità più rilevante consiste nel fatto che il tema è sviluppato attraverso un approccio casistico:

dopo avere premesso alcuni chiarimenti di ordine generale (tra cui il concetto di robot e problemi più pressanti per il mondo giuridico), il documento si concentra infatti sulle applicazioni robotiche più significative e, segnatamente, su auto a guida autonoma, robot chirurgici, protesi robotiche, robot impiegati nell'*health care*.

Lo stesso metodo di indagine è stato poi ripreso nel documento *ELS issues in robotics and steps to consider them: Robotics and Regulations* che nel 2016 ha concluso il progetto *RockEu*[101].

---

[99] Il documento costituisce precisamente «*a proposal for a "Green Paper"*», ove quest'ultimo nella terminologia della Commissione Europea rappresenta «*a discussion document intended to stimulate debate and launch a process of consultation, at European level, on a particular topic*». Esso può servire da lavoro preparatorio per un «*White Paper*» e cioè un documento contenente delle proposte per un'azione dell'Unione Europea in un settore specifico. Cfr. C. LEROUX, R. LABRUTO et a., *A Green Paper on Legal Issues in Robotics*, euRobotics CA, cit., p. 7.

[100] E. PALMERINI et a., *Guidelines on regulating robotics*, The RoboLaw Project, cit., p. 8. Per una sintesi cfr. E. PALMERINI, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, cit., p. 1815 ss.; HOLDER C., V. KHURANA, F. HARRISON, L. JACOBS, *Robotics and law: Key legal and regulatory implications of the robotics age*, in *Computer Law & Security Review*, p. 2016, p. 383 ss.

[101] Cfr. A. SANTOSUOSSO, B. BOTTALICO, *Autonomous Systems and the Law: Why Intelligence Matters. A European Perspective*, cit.

Un approfondimento maggiore merita, come accennato, la *Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*. Il documento si confronta con diversi temi concernenti lo sviluppo della robotica e l'intelligenza artificiale come, ad esempio, l'esigenza che i soggetti coinvolti nello sviluppo e nella commercializzazione di tali prodotti integrino gli aspetti relativi alla sicurezza e all'etica fin dal principio (cioè già a livello di *design*) oppure le conseguenze della diffusione dei robot per il mercato del lavoro e le potenziali ricadute sul tasso di occupazione[102]. Qui ci interessa tuttavia mettere a fuoco la parte specificamente dedicata ai profili di responsabilità connessi alla produzione e all'impiego dei sistemi robotici.

Leggiamo dunque insieme alcuni 'stralci' del documento.

La Risoluzione prende le mosse dall'idea che «*l'umanità si trova ora sulla soglia di un'era nella quale robot, bot, androidi e altre manifestazioni dell'intelligenza artificiale sembrano sul punto di avviare una nuova rivoluzione industriale, suscettibile di toccare tutti gli strati sociali*»[103] *e dall'attestazione secondo cui «l'andamento attuale, che tende a sviluppare macchine autonome e intelligenti, in grado di apprendere e prendere decisioni in modo indipendente, genera nel lungo periodo non solo vantaggi economici ma anche una serie di preoccupazioni circa gli effetti diretti e indiretti sulla società nel suo complesso»*[104].

In secondo luogo, essa si sofferma sui tratti peculiari degli attuali sistemi robotici. Essi sono individuati, oltre che nella capacità di *«svolgere attività che tradizionalmente erano tipicamente ed esclusivamente umane»,* nella possibilità di *«apprendere dall'esperienza e di prendere decisioni quasi indipendenti»*[105]. Ciò contribuisce a considerare i robot *«sempre più simili ad agenti che interagiscono con l'ambiente circostante e sono in grado di alterarlo in modo significativo»* e a gettare dubbi sulla loro qualificazione alla stregua di *«meri strumenti nelle mani di altri attori (quali il fabbricante, l'operatore, il proprietario, l'utilizzatore, ecc.)"*[106].

[102] Per un primo commento cfr. SCHAFER B., *'Closing Pandora's box? The EU proposal on the regulation of robots'*, in *The journal of the Justice and the Law Society of the University of Queeensland*, 2016, vol. 19, p. 55 ss; E. STRADELLA, *La regolazione della Robotica e dell'Intelligenza artificiale: il dibattito, le proposte, le prospettive. Alcuni spunti di riflessione,* in *MediaLaws – Rivista dir. media,* 2019, in corso di pubblicazione.

[103] Cfr. Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica. Introduzione, Lett. B).

[104] *Ivi*, Introduzione, Lett. G).

[105] *Ivi*, Responsabilità, Lett. Z).

[106] *Ivi,* Responsabilità, Lett AB).

Le conseguenze più rilevanti di questa ricostruzione si avvertono in tema di responsabilità (nel caso di specie, vale la pena ripeterlo, è esaminata soltanto la responsabilità civile, contrattuale ed extracontrattuale). La *«capacità di adattamento e di apprendimento» dei robot, da cui discende un certo margine di «imprevedibilità nel loro comportamento»*[107], inducono infatti a domandarsi se *«le regole ordinarie in materia di responsabilità siano sufficienti o se ciò renda necessari nuovi principi e regole volte a chiarire la responsabilità legale dei vari attori per azioni e omissioni imputabili ai robot, qualora le cause non possano essere ricondotte a un soggetto umano specifico, e se le azioni o le omissioni legate ai robot che hanno causato danni avrebbero potuto essere evitate»*.

Nel prosieguo la Risoluzione esprime dubbi circa l'adeguatezza della normativa in tema di responsabilità (civile) da prodotto (direttiva n. 85/374/CEE) a fare fronte ai danni provocati dai sistemi robotici. Se non interpretiamo male, il documento osserva che la responsabilità da prodotto, nella sua versione colposa e/o oggettiva, mal si concilia con l'ipotesi di danni che, non solo possono essere imprevedibili e quindi difficilmente evitabili (rispetto a cui mancherebbe cioè l'elemento colposo), ma nemmeno possono essere ricondotti a un difetto del prodotto, posto che la capacità di apprendimento è da intendersi come voluta (e dunque non sarebbe ravvisabile quell'elemento, che costituisce il necessario presupposto dello schema oggettivo della responsabilità).

Essa invita pertanto a riflettere sulla «*natura* [dei sistemi robotici] *alla luce delle categorie giuridiche esistenti*» e sull' «*eventuale necessità di creare una nuova categoria con caratteristiche specifiche e implicazioni proprie*»[108].

Sul versante propositivo, la Risoluzione raccomanda anzitutto alla Commissione di elaborare una definizione di «*sistemi ciberfisici, di robot autonomi intelligenti e delle loro sottocategorie*»[109] imperniata sui seguenti requisiti: autonomia, autoapprendimento (indicato come criterio facoltativo), esistenza di un supporto fisico, capacità di adattamento del proprio comportamento all'ambiente, assenza di vita in termini biologici.

È bene osservare che si tratterebbe di una delle prime definizioni legislative del concetto di robot di carattere generale, vale a dire trasversale rispetto a una pluralità di domini applicativi[110].

---

[107] *Ivi*, Responsabilità, Lett. AI)

[108] *Ivi*, Responsabilità, Lett. AC)

[109] *Ivi*, Principi generali riguardanti lo sviluppo della robotica e dell'intelligenza artificiale per uso civile, pt. 1.

[110] B. SCHAFER, *'Closing Pandora's box? The EU proposal on the regulation of robots'*, cit., p. 55 ss., il quale peraltro ritiene che alcuni importanti tematiche non siano affrontate affatto (come, ad esempio, in materia di *privacy*).

Con riferimento invece alla tematica della responsabilità, essa suggerisce di esplorare la strada della previsione di un regime di assicurazione obbligatoria sul modello di quello in vigore per il settore automobilistico, nonché l'istituzione di un fondo per garantire la possibilità di risarcire i danni in caso di assenza di copertura assicurativa. Infine, senza escludere alcuna soluzione, riconosce anche la possibilità di creare «*uno status giuridico specifico per i robot nel lungo termine, di modo che almeno i robot autonomi più sofisticati possano essere considerati come persone elettroniche responsabili di risarcire qualsiasi danno da loro causato, nonché eventualmente il riconoscimento della personalità elettronica dei robot che prendono decisioni autonome o che interagiscono in modo indipendente con terzi*»[111].

Per concludere sul punto, fermiamo l'attenzione su due osservazioni critiche[112] che sono state sollevate dalla dottrina nell'immediatezza dell'emanazione della Risoluzione.

La prima riguarda il contenuto della definizione di robot proposto dal documento e, più in generale, la scelta di fornire una definizione onnicomprensiva di robot 'intelligente'. Parte della dottrina osserva, in effetti, che la definizione proposta, ancorché conforme alle indicazioni provenienti dal mondo tecnico, non per è per ciò solo da ritenersi adeguata anche sul versante giuridico; in secondo luogo, dubita dell'utilità stessa di una definizione generale di robot, ritenendo preferibile, ai fini della regolamentazione positiva, elaborare definizioni settoriali del concetto di robot, che tengano conto dello specifico settore di appartenenza (medico, militare, settore dei trasporti).

La seconda annotazione è che la Risoluzione si sofferma esclusivamente sui sistemi robotici dotati di componenti di intelligenza artificiale, tralasciando altre applicazioni fondate sull'intelligenza artificiale (es. software agents) che, oltre a presentare gli stessi problemi dei robot 'intelligenti', sono già state in passato oggetto di analisi da parte degli studiosi e potrebbero offrire utili spunti quantomeno in riferimento alla responsabilità contrattuale degli agenti autonomi.

5. L'informatica del diritto

[111] *Ivi*, Responsabilità, pt. 59, Lett F).

[112] Un'opinione particolarmente negativa rispetto al testo della Risoluzione è espressa dalla *Open Letter to the European Commission Artificial Intelligence and Robotics*, firmata da un gruppo di esperti nel campo dell'intelligenza artificiale e della robotica. In essa si legge in particolare che: «*The creation of a Legal Status of an "electronic person" for "autonomous", "unpredictable" and "self-learning" robots is justified by the* incorrect affirmation that damage liability would be impossible to prove. *From a technical perspective, this statement offers many bias based on an overvaluation of the actual capabilities of even the most advanced robots, a superficial understanding of unpredictability and self- learning capacities and, a robot perception distorted by Science-Fiction and a few recent sensational press announcements*».

Nel 1963 l'American Bar Association[113] pubblica un articolo[114] in cui, per la prima volta, si pone il quesito: *«Will computers revolutionize the practice of law and the administration of justice, as they will in almost everything else?»*. Nell'articolo veniva esaminata l'incidenza delle nuove tecnologie sulla società concludendo che queste avrebbero inevitabilmente rivoluzionato tutti i settori compreso ovviamente quello giurisprudenziale. La giurisprudenza in un primo momento viene chiamata a dare ordine a tutte le nuove problematiche scaturite dal rapido sviluppo tecnologico che si intersecano sempre di più frequentemente con la vita di ogni individuo. Gli operatori del diritto si trovarono a dover esaminare e disciplinare questioni di una natura totalmente nuova.

Le problematiche nascono e si trasformano di pari passo con l'inarrestabile progresso scientifico.

Il primo ostacolo che si oppone sulla strada di chi tenta di dettare norme sul nuovo fenomeno è quello di dover acquisire competenze informatiche, seppur basilari, indispensabili per poter capire il funzionamento e le problematiche relative a questi nuovi sistemi, per poi poterli disciplinare in maniera ottimale. Solo chi sappia cosa è un codice oggetto e codice sorgente può capire il funzionamento del software, solo chi conosce il funzionamento degli indirizzi di Internet può comprendere cosa siano i nomi di dominio, non si può pensare di disciplinare il tema degli accessi abusivi a sistemi informatici senza che si conoscano i modi di attacco e le relative difese degli stessi.

Per quanto complesso, però, questo è solamente un aspetto della questione, col tempo la giurisprudenza stessa ha cominciato ad utilizzare in prima persona questi sistemi facendo acquisire al giurista non solo nuovi strumenti di lavoro, ma un nuovo modo per eseguirlo. Affrontando questi temi solitamente si prende in considerazione quello che viene definito 'diritto dell'informatica costituito da tutte le leggi che sono state emanate per disciplinare l'uso dei computer (reati informatici, tutela della privacy, proprietà intellettuale ecc.)[115], si trascura invece l''informatica del diritto'[116] dove è il diritto stesso che si fa oggetto dell'elaborazione elettronica[117].

[113] L'American Bar Association, fondata il 21 agosto 1878, è un'associazione volontaria di avvocati e studenti di giurisprudenza, che non è specifica per nessuna giurisdizione negli Stati Uniti.

[114] R.C. LAWLOR, What Computers Can Do: Analysis and Prediction of Judicial Decisions, in 'American Bar Association Journal', 49, 1963, p.337

[115] Come ad esempio le leggi che riguardano la proprietà intellettuale o i reati informatici

[116] Definita anche 'informatica giuridica in senso stretto'

[117] R. BORRUSO, "Riflessioni sull'informatica giuridica", in: L'informatica del diritto, Milano 2004, p.297

L'informatica del diritto può avere diversi ruoli nell'attività e nella cultura del giurista; in primo luogo può accrescere l'efficienza del suo lavoro attraverso l'utilizzo di strumenti che gli permettano di svolgerne prima e meglio molti aspetti: la redazione di atti, l'accesso a documenti, la tenuta della contabilità ecc. L'utilità che può apportare l'informatica rimane molto limitata se le attività giuridiche continuano a svolgersi secondo forme pre-informatiche, una piena potenzialità dei sistemi informatici richiede una riorganizzazione delle varie attività giuridiche e del loro modo di esecuzione.

Un esempio emblematico lo possiamo vedere in relazione al processo civile telematico[118]: affinché l'attività giudiziale possa sfruttare in maniera piena le nuove tecnologie non è sufficiente fornire agli uffici giudiziari dei calcolatori più potenti e software migliori, ma bisogna determinare dei protocolli sulla comunicazione processuale telematica, sulle forme che devono avere gli atti e realizzare un'infrastruttura tecnologica che possa garantire efficienza e sicurezza[119].

L'utilizzo di sistemi informatici è ormai largamente diffuso nelle professioni legali, basti pensare all'uso delle banche dati, di cui ormai nessuno pensa di poter fare a meno[120].

Lo studioso Roland Volg, identifica cinque aree di maggiore crescita di tecnologia nella pratica forense:

1) Ricerca legale: in particolare per quanto riguarda le tecniche di recupero delle informazioni legali e giuridiche per fornire elementi utili ai fini della decisione.

Chiunque voglia studiare diritto o applicarlo deve necessariamente conoscere le leggi e per fare ciò è necessario saperle ricercare.

In un ordinamento che è sempre meno a dimensione d'uomo ricercare le leggi da applicare al caso concreto diventa sempre più difficoltoso visto l'enorme numero di leggi e di atti aventi forza legge.

---

[118] Il processo civile telematico (PCT) è la disciplina del processo civile svolto attraverso strumenti digitale e, in particolare, tramite il deposito telematico degli atti di causa. Il PCT è attualmente vigente nella Repubblica Italiana. Ad oggi il PCT coinvolge il contenzioso civile, la volontaria giurisdizione, il processo del lavoro, le esecuzioni mobiliari, le esecuzioni immobiliari e le procedure concorsuali (sebbene - in via obbligatoria - solo per la fase successiva alla dichiarazione di apertura della procedura). In parallelo il Ministero della Giustizia ha abilitato un servizio per la consultazione pubblica dei dati non giudiziali attinenti al Giudice di pace. Le basi normative del processo civile telematico sono contenute nel Decreto del presidente della Repubblica 13 febbraio 2001, n. 123, Regolamento recante disciplina sull'uso degli strumenti informatici e telematici nel processo civile, nel processo amministrativo e nel processo dinanzi alle sezioni giurisdizionali della Corte dei Conti, che introdusse il processo civile telematico nell'ordinamento

[119] P. LICCARDO, Introduzione al processo civile telematico, in: Rivista trimestrale di diritto e procedura civile (2000), PP.1165 ss.

[120] M. MAUGERI, I robot e la possibile «prognosi» delle decisioni giudiziali, Contenuto in A. CARLEO, Decisione robotica, Bologna 2019 p. 159

Per fronteggiare quindi questa problematica sono state progettate banche dati dove poter ricercare qualsiasi legge con la relativa giurisprudenza.
Di queste banche dati in Italia quella più famosa è curata dalla Corte di Cassazione accompagnata da banche dati realizzati da editori giuridici privati su CD-ROM consultabili mediante PC[121].
2) Big data law: con ciò si intendono le varie tecniche di elaborazione di NLP e machine learning che permettono di prevedere risultati e creare modelli di casi, contratti e documenti legali.
3) Contract Automation: letteralmente si può tradurre con 'automazione del contratto ' e con ciò si intende l'intervento dell'IA in materia contrattuale che comprende i vari ambiti di:
- Drafting ('redazione'): cioè la redazione automatizzata di contratti standard.
- Contract Management ('gestione del contratto'): che permette di avere sotto controllo tutti i contratti attivi con scadenze, adempimenti e pagamenti.
- Contract Review ('revisione del contratto): la revisione automatizzata dei contratti.
- Extraction ('estrazione'): la possibile estrazione di informazioni mirate nei documenti contrattuali.
- Due diligence: comprensiva di software in grado di rilevare anomalie o previsioni specifiche in una determinata mole di contratti.
4) Expertise Automation ('competenza automatizzata'): quei sistemi che aiutano l'esperto di diritto (ad esempio l'avvocato) nello svolgimento della propria professione.
5) ODR (Online Dispute Resolution): sistemi che permettono di risolvere le controversie online. La Gran Bretagna insieme ai Paesi Bassi e la Lettonia sono i paesi europei più all'avanguardia sotto questo punto di vista, a livello europeo, per la risoluzione delle controversie transfrontaliere grazie al regolamento n. 524/2013, è stato istituito un quadro comune disponibile su internet (European small claims).
Questi servizi online di risoluzione delle controversie si stanno gradualmente estendendo e vengono, sempre più integrati nel processo giudiziario. L'ambito di applicazione non è più solo la controversia di modesto valore, ma anche il contenzioso in materia tributaria, quello relativo ai servizi della previdenza sociale e le cause di divorzio. Il ruolo principale di questi sistemi di ODR è chiaramente quello di contribuire all'applicazione di servizi stragiudiziali di conciliazione, mediazione e arbitrato, al punto che tali servizi vengono

[121] Banca dati Italjure.

anche utilizzati nel corso del procedimento giudiziario sotto la supervisione del giudice che dovrà decidere l'esito della controversia.

Nonostante la grande diffusione di queste ODR non bisogna mai dimenticare che in Europa è stato recentemente adottato un quadro normativo protettivo che all'art 22 del GDPR (Regolamento Generale sulla Protezione dei dati, Reg. 2016/679) vincola gli Stati membri a tutelare i cittadini dando la possibilità alle persone di rifiutarsi di essere sottoposte a decisioni basate esclusivamente su un trattamento automatizzato dei dati.

6) Giustizia "predittiva": cioè la capacità di elaborare previsioni attraverso sistemi probabilistici che utilizzano algoritmi operanti su base logicostatistica, in grado di determinare l'esito di un processo. Questa capacità "predittiva" può essere utilizzata in tre modi diversi tra loro a seconda dei dati che vengono inseriti nell'elaboratore:

- Prevenzione della criminalità: attraverso l'analisi di dati estrapolati da denunce presentate alla Polizia, relativi per esempio a rapine o furti che si sono verificati ripetutamente in zone specifiche con modalità analoghe, determinati sistemi di intelligenza artificiale sono in grado di prevedere luoghi e orari in cui, verosimilmente, potranno verificarsi altri reati dello stesso tipo. Sembra di trovarsi nel mondo fantascientifico ideato da Phil Dick di Minority Report[122] dove esiste un dipartimento 'Pre-crimine' in cui la polizia grazie alle premonizioni di tre individui dotati di poteri extrasensoriali e precognitivi, riesce ad impedire gli omicidi prima che essi si verifichino e ad arrestare i potenziali colpevoli.

Tale programma permette di elaborare i dati di tutti gli scippi, rapine e borseggi e furti che vengono desunti dalle denunce effettuate dai cittadini e dalle informazioni raccolte dalla polizia di prossimità insieme alle notizie diffuse dagli organi di stampa e sui social network.

I crimini predatori urbani si compiono in luoghi precisi, sono realizzati infatti da soggetti che hanno bisogno di un profitto che deve essere realizzato in un arco temporale di breve durata, sono crimini che si ripetono nel tempo e che seguono strategie uguali in tutti i centri urbani.

Il luogo in cui si verifica questa tipologia di crimini ha sempre due caratteristiche una di tipo oggettivo e l'altra di tipo soggettivo:

1) presenza di prede e target appetibili;

2) presenza di via di fuga, rifugi e sistemi di copertura criminale del luogo.

[122] "Minority report" di Phil K. Dick, Orion Publishing and Co, 2009

In questo modo si vengono a creare delle vere e proprio 'riserve di caccia' all'interno delle quali vengono inglobate fasi e operazioni regolari come l'entrata e uscita da abitazioni, uffici scuole, arrivo di treni autobus o altri mezzi.
Se queste dinamiche, che si verificano in maniera sequenziale, vengono sovrapposte ai vari crimini compiuti su base regolare, si è in grado di decodificare le sequenze e quindi di prevedere i singoli delitti prima che si verifichino.
Attraverso l'utilizzo di un algoritmo euristico e non statistico infatti, tali dati vengono poi elaborati su una carta topografica digitale non solo le così definite zone di caccia ma anche gli alert temporali geo-referenziati che si aggiornano ogni mezz'ora su dove e quando potrebbe accadere un reato, andando a descrivere il tipo di crimine, il modus operandi dell'autore e il tipo di preda e target.
Una volta terminata la sua analisi, il software si esprime attraverso un interfaccia WebApp che permette alle forze dell'ordine quindi di usufruirne in tempo reale, sui loro dispositivi elettronici cellulari, tablet ecc.. In questo modo viene favorita, non solo la predisposizione dei presidi nel posto e nel momento giusto, ma anche l'attuazione di strategie operative che rendono gli aggressori meno efficaci e più vulnerabili.
Usando questi strumenti è stato possibile attuare la migliore azione d'intervento in ogni occasione cosa che ha ridotto considerevolmente le probabilità e il rischio di operazioni pericolose. Ottimizzare le risorse operative impiegandole al meglio ed in maniera organizzata ha reso il sistema più efficace, ha consentito di ridurre il numero di crimini, e permesso ai cittadini di sentirsi più sicuri, ma ha anche ridotto le risorse economiche impiegate in questo tipo di attività.

6. Modelli di responsabilità per self-learning robot: campo di indagine

6.1. I modelli elaborati in ambito penale: il pensiero di G. Hallevy

La diffusione dei sistemi di intelligenza artificiale pone anzitutto questioni di ordine costituzionale. Al costituzionalista, infatti, compete individuare le forme idonee per assicurare che i valori e i diritti tutelati dalla carta fondamentale e della Carta di Nizza possano trovare adeguata protezione anche al cospetto dei sistemi di intelligenza artificiale. Occorre inoltre porsi alla ricerca di norme che agiscano come argine rispetto all'esercizio di un potere che non promana più esclusivamente dalle autorità pubbliche, ma si ritrova sempre più concentrato nelle mani di operatori privati (Teubner, 2004). Il rapporto tra Stato e cittadini si fonda, come noto, sul riconoscimento in capo a questi

ultimi di un sistema di garanzie. Queste ultime intendono porre al riparo gli individui da possibili abusi e arbìtri da parte dei temporanei detentori del potere pubblico. Questa dinamica impone oggi un ripensamento alla luce delle caratteristiche diffuse del potere "privato". Si tratta di un potere detenuto da chi utilizza, spesso alimentandosi di ingenti quantità di dati (big data), tecnologie algoritmiche e sistemi di intelligenza artificiale per realizzare attività che presentano implicazioni rilevanti per i diritti e le libertà delle persone (De Gregorio, 2019).

Questa necessaria estensione delle garanzie non comporta per l'esclusione dei poteri pubblici, che a loro volta possono essere fruitori e utilizzatori di sistemi di intelligenza artificiale per finalità disparate, senz'altro consentite dall'ordinamento, tra cui l'efficientamento dell'attività amministrativa. Infatti, l'intelligenza artificiale non coinvolge soltanto la sfera dei dritti dell'individuo ma anche altri principi costituzionali quali quelli formanti l'organizzazione della giustizia e l'attività della pubblica amministrazione.

La circostanza che criticità emergano anche nel rapporto tra cittadini e poteri pubblici, e non soltanto nell'ambito dei rapporti orizzontali inter-privati, illustra efficacemente come sia necessario un ripensamento di alcuni aspetti del rapporto di cittadinanza (sempre più digitale). Le norme costituzionali, pensate esclusivamente per attagliarsi al rapporto tra potere pubblico e individuo, paiono vacillare al cospetto di questo nuovo scenario.

La tutela dei diritti appare, invero, soltanto uno dei versanti rispetto ai quali è auspicabile che si registri un'evoluzione in grado di restituire un elevato livello di controllo da parte degli individui. Nella prospettiva del diritto pubblico si palesano altresì esigenze di governance della tecnologia, che implicano la necessità per i regolatori di realizzare opportune forme di concertazione a livello nazionale e sovranazionale.

L'esigenza di rafforzare la tutela dei diritti fondamentali si coglie guardando alle embrionali forme di codificazione di nuovi diritti, di cui si ha traccia, per esempio, nel già ricordato GDPR, al Considerando 71 e all'art. 22. Tali referenti paiono dare fondamento a un diritto "alla spiegazione" nell'ambito dei processi decisionali automatizzati, così da costituire al contempo una barriera a tutela degli individui rispetto a decisioni del tutto estranee a un intervento umano che incidano significativamente sui loro diritti. La norma appare veicolare un messaggio evidente: restituire centralità al fattore umano, pur senza che la necessaria precedenza logica e giuridica dell'agente umano possa ergersi a ostacolo all'innovazione tecnologica e al miglioramento dei processi che parti pubbliche e private pongono in essere. Il diritto alla spiegazione cela

evidentemente un'esigenza primitiva di trasparenza, che soprattutto (ma non esclusivamente) nel campo dei provvedimenti amministrativi riflette un principio cruciale per conferire al cittadino la possibilità di verificare le modalità di esercizio del potere ed esercitare su di esso un controllo, compreso il sindacato giurisdizionale. Questa esigenza non manca di essere avvertita anche nell'ambito delle applicazioni nel settore privato, dove sono invece carenti specifici meccanismi che consentano analogo controllo da parte di chi subisce gli effetti di una decisione automatizzata (in dottrina, Wachter-Mittelstadt,-Floridi, 2017; Kaminski, 2019). Accanto al diritto alla spiegazione si potrebbe valutare la codificazione di un ulteriore principio, ossia il diritto a conoscere il proprio interlocutore e la sua natura, onde assicurarsi non soltanto della trasparenza del processo decisionale automatizzato ma anche della sua effettiva corrispondenza a una logica umana ovvero algoritmica. Il diritto a non essere sottoposti a un trattamento interamente automatizzato si iscrive nella disciplina specifica sulla tutela dei dati personali, ma ambisce ovviamente a un riconoscimento più ampio, che lo emancipi dalla dimensione della data protection. Sotto questo profilo, è importante osservare il trend che ha condotto negli ultimi anni la Corte di giustizia a interpretare il diritto alla privacy e alla protezione dei dati con effetti orizzontali, in particolare nella decisione Google Spain del 2014 (causa C-131/12), che a sua volta concerneva l'attività di un motore di ricerca consistente nella indicizzazione automatica di dati personali contenuti in pagine web di siti sorgente. In questa e altre pronunce la Corte di giustizia ha offerto importanti indicazioni a sostegno di un'applicazione dei diritti fondamentali anche nei rapporti orizzontali inter-soggettivi, quale per esempio quello che si instaura tra motore di ricerca e individuo, i cui dati sono oggetto di memorizzazione tra i risultati delle ricerche (Pollicino, 2018). Proprio questo trend parrebbe sostenere l'opportunità di codificare o comunque riconoscere nuove situazioni giuridiche che non trovano immediata tutela nelle carte dei diritti in quanto si mostrano sguarnite di protezione e dunque di giustiziabilità soprattutto nei rapporti inter-privati. Infatti, di fronte all'attività dei poteri pubblici, princìpi come la trasparenza impongono una possibilità di accesso e di controllo sull'operato dell'amministrazione. Invece, nei rapporti privati tali tutele non sarebbero direttamente applicabili in assenza di una codificazione normativa. Il GDPR ha tentato di realizzare una tale codificazione sotto il profilo della tutela dei dati personali, ma essa inevitabilmente aspira a un confezionamento entro sedi normative più generali, financo, per alcuni tratti, di rango costituzionale.

Le implicazioni di carattere costituzionale sul fronte dei diritti e dell'eguaglianza non sono circoscritte soltanto a questo ambito, ma conoscono manifestazioni in forme variegate (Casonato, 2019; Simoncini, 2019).

Tensioni si pongono anzitutto rispetto all'eguaglianza dei cittadini, da intendersi in particolare come parità di chances e sotto il profilo "sostanziale", non meramente ancorata a un vincolo formalistico. La diffusione di sistemi di intelligenza artificiale è destinata, per esempio, a determinare conseguenze di grande momento in ambito occupazionale: per un verso, sostituendo gli individui nell'esecuzione di mansioni che possono essere svolte analogamente da una macchina, per altro verso sottraendo però opportunità lavorative. L'impatto di questi sistemi riguarderà tanto le professioni intellettuali quanto le attività diverse da queste ultime; è ragionevole pronosticare, tuttavia, che saranno soprattutto quelle artigiane e operaie (normalmente esercitate da personale con livello meno elevato di istruzione), a risentire del rapporto di succedaneità tra uomo e macchina, con ricadute sociali che potrebbero attingere l'eguaglianza sostanziale dei cittadini. Un altro versante esemplificativo riguarda l'ambito della salute, ove è ragionevole pronosticare un più diffuso ricorso a tecniche e dispositivi medicali di nuova generazione, che però potrebbero, per gli elevati costi di sviluppo e di messa a regime, causare una disparità nell'accesso alle cure somministrate tra abbienti e meno abbienti.

Un secondo profilo meritevole di attenzione si ricollega alla sfera dei diritti politici. L'impiego di tecniche algoritmiche favorisce dinamiche di profilazione degli utenti basate sulla raccolta delle loro preferenze, in grado di rifletterne convinzioni e orientamenti politici, morali, filosofici, religiosi, etc. Come dimostra anche la vicenda Cambridge Analytica, la raccolta di queste informazioni a carattere personale può alimentare una bubble democracy nella quale gli utenti sono destinatari di messaggi personalizzati, che in quanto selezionati in base all'analisi algoritmica delle loro preferenze, si presume riscuoteranno maggiore successo (Sunstein, 2009). Questa logica di "customizzazione" rischia tuttavia di deformare la reale estensione del quadro delle opinioni e visioni politiche. Si condiziona il rafforzamento delle posizioni di partenza (confirmation bias) anziché favorire una loro ponderazione ed eventuale rivisitazione tramite il confronto critico con altre di segno opposto, che esistono nella realtà ma divengono "invisibili" sul web e sui social network in particolare (Pariser, 2011).

Sul versante della governance è avvertita l'esigenza di individuare o costituire un'autorità indipendente che possa assurgere a garante di uno sviluppo dei sistemi di intelligenza artificiale in conformità a linee e orientamenti condivisi a livello sovranazionale, in modo

da preservare la centralità della persona umana, della sua dignità e dei diritti fondamentali. L'amministrazione, del resto, non pu astenersi dall'intervenire a fronte di fatti economici rilevanti in cui viene sviluppata la capacità delle macchine di riprodurre o attuare operazioni tipiche delle funzioni cognitive umane, compiendo azioni con un certo grado di autonomia. La public regulation potrebbe essere affidata ad un apparato estraneo al circuito politico, indipendente rispetto al governo e a singoli ministri, e caratterizzato dalla tecnicità dei componenti. La specializzazione del personale è necessaria dal momento che l'attività richiede conoscenze specifiche che non si ritrovano ordinariamente nell'ambito della funzione pubblica, ed è utile a prevenire conflitti di interessi. In via alternativa, potrebbero essere istituiti uffici di regolazione sul modello statunitense, vigilati o incardinati presso i dipartimenti dell'esecutivo e con competenze settoriali. La collocazione ideale di un'autorità di questo tipo è sovranazionale, perché di rilevanza internazionale sono i temi, le tecniche, i caratteri delle questioni affrontate. Ma la stessa è suscettibile di essere collocata anche a livello domestico, pur sempre entro un network europeo sul modello di altre regolazioni economiche.

A una simile Autorità potrebbero essere attribuiti poteri sia amministrativi che "quasi legislativi" e "quasi giudiziali", ricorrendo a tecniche di hard e soft law, e cumulando diverse funzioni che allo stato non appaiono univocamente ascrivibili agli attori pubblici esistenti. Potrebbe svolgere funzioni di carattere regolatorio dell'iniziativa privata, provvedendo anche alla formulazione di indirizzi generali nel settore. Ovvero funzioni di certificazione pubblica a garanzia dell'affidabilità e della trasparenza per i produttori e per gli utenti. Ovvero ancora funzioni di vigilanza o di accertamento rispetto a possibili rischi di manipolazione e profilazione. A questa Autorità a potrebbe essere rimesso il compito di disciplinare i requisiti di trasparenza, sicurezza, affidabilità, sostenibilità, accountability di chi opera nel settore – anche mediante sistemi di accreditamento –, nonché le modalità di tutela dei soggetti vulnerabili. L'Autorità dovrebbe regolarmente riferire in Parlamento sull'attività svolta. La scelta delle politiche pubbliche necessarie per l'attuazione dell'intelligenza artificiale, dovrebbe, invece, rimanere affidata agli organismi politico-rappresentativi.

A questo sistema potrebbe validamente aggiungersi la costituzione di un organo consultivo permanente che replichi il modello multi-stakeholder – attuato, per esempio, con l'Internet Governance Forum – aprendosi ai vari attori del settore pubblico e privato per promuovere il dibattito sulle implicazioni etiche, giuridiche ed economiche dei sistemi di intelligenza artificiale. Nel complesso, la prospettiva proposta sembra

confermare la necessità di un sistema di governo pubblico in grado di incidere dall'alto sulle posizioni economiche degli operatori privati, secondo una filosofia differente da quella che ispira la Internet Corporation for Assigned Names and Numbers (ICANN). In qualità di ente privato che opera senza fini di lucro sul sistema dei nomi a dominio, degli indirizzi e dei protocolli internet, esso ha dato vita ad un modello di governance contrattuale, basato sulla necessità della stipulazione di accordi privatistici tra l'ente e gli operatori. Eppure, in tale articolato sistema di governo, le istituzioni pubbliche e gli elementi di autoritatività non scompaiono affatto, poiché la rete necessita di un sistema unitario in cui singole parti possono differenziarsi, ma a condizione che ciò avvenga secondo regole tali da assicurare l'uniformazione e l'integrazione funzionale.

Sempre sul versante pubblicistico, come si è evidenziato, l'intelligenza artificiale promette un impatto di grande momento rispetto all'attività della pubblica amministrazione, e dunque nell'ambito precipuo del diritto e del processo amministrativo. La già ricordata vicenda giurisprudenziale che ha trovato seguito nella pronuncia del Consiglio di Stato ha descritto importanti indicazioni sul rapporto tra regola algoritmica e regola giuridica (Luciani, 2018). L'utilizzo da parte della pubblica amministrazione di sistemi algoritmici impone pertanto di costituire una serie di vincoli che possano conformare la tecnologia ai principi sottesi all'attività amministrativa. Naturalmente, al fine di assicurare un utilizzo virtuoso e rispondente all'interesse pubblico dell'intelligenza artificiale potrebbe essere utile stabilire alcuni presìdi mediante legge ordinaria, per esempio nella seguente formula: «Le pubbliche amministrazioni istituiscono una rete di comitati tecnici per il controllo e il monitoraggio costante dei processi di intelligenza artificiale composti da scienziati, economisti e giuristi, con il compito di fornire al Governo gli elementi da illustrare al Parlamento sulla verifica annuale dei processi di sviluppo e applicazione dell'intelligenza artificiale. I comitati tecnici valutano anche gli impatti sociali e giuridici dell'uso dell'intelligenza artificiale da parte del settore pubblico al fine di modulare opportunamente gli obiettivi settoriali».

Questi interventi potrebbero permettere di arginare le insidie insite nell'esistenza di un ineliminabile bias dovuto alla programmazione degli algoritmi da parte di agenti umani, sensibili, a una serie di fattori "esterni". Si mira così ad assicurare che in questo ambito i parametri di funzionamento dell'algoritmo costituiscano il risultato di un processo decisionale pubblico di cui sia garantita la massima trasparenza. Come contraltare, ricorre per l'esigenza di un controllo democratico sui pubblici poteri da parte dei cittadini, che presuppone un sufficiente grado di alfabetizzazione corrispondente all'elevato contenuto

tecnico degli algoritmi, che trova eco anche normativamente nel principio di totale accessibilità degli atti e dei documenti amministrativi (cosiddetto "FOIA", d.lgs. n. 97/2016). Tali interventi potrebbero trovare attuazione, per esempio, a partire da disposizioni come l'art. 8 del Codice dell'amministrazione digitale. Analoghe misure dovrebbero mirare a consentire alla stessa pubblica amministrazione e ai funzionari che vi operano di mantenere un effettivo potere di dominio sul prodotto di una decisione automatizzata, loro imputabile, di cui agli effetti giuridici essi continuano a rispondere.
Questi interventi paiono rappresentare le condizioni ottimali affinché i processi automatizzati, già diffusi in misura apprezzabile nell'ambito di atti amministrativi vincolati o a bassa discrezionalità amministrativa, possano conoscere un'estensione della loro applicazione anche nell'ambito della formazione delle decisioni amministrative contraddistinte da un più elevato livello di discrezionalità. Sotto questo profilo, particolari criticità andranno affrontate rispetto alla motivazione dell'atto amministrativo, adempimento la cui complessità può variare significativamente a seconda del grado di discrezionalità richiesto. Sul versante dell'intensità del controllo giurisdizionale, le proprietà dell'algoritmo non andranno considerate come un ambito riservato alla pubblica amministrazione, non attingibile dal sindacato giurisdizionale, se non in termini di ragionevolezza e proporzionalità, bensì dovranno essere oggetto di una piena e diretta verifica istruttoria.

7. Intelligenza artificiale e diritto civile

L'intelligenza artificiale e le nuove tecnologie avranno un significativo impatto anche sulle regole del diritto privato. In primo luogo, nell'area concernente "le persone", studiosi quali Teubner[123], hanno discusso in merito alla esigenza di creare una nuova figura giuridica dotata della capacità di agire e della capacità di essere titolare di situazioni giuridiche soggettive: il robot. Il nuovo soggetto dovrebbe affiancarsi alle persone naturali e agli enti nel quadro dei rapporti privatistici (Teubner, 2018). I vantaggi risiederebbero nella possibilità di limitare la responsabilità dei programmatori e degli utilizzatori dei servizi del nuovo potenziale soggetto al fine di promuovere lo sviluppo tecnologico. Nelle trattazioni più approfondite si osserva che il futuro soggetto, dotato di intelligenza artificiale, potrebbe essere sottoposto a regole che allo stato vigono per gli

[123] Si veda *"Verso la tutela giuridica dei sistemi intelligenti. Prospettive critiche della soggettività robotica"* di G. DE BONA, in Journal of Ethics and Legal Technologies – Volume 4(2) – November 2022

enti. In questo modo, ad esempio, sarebbe possibile assicurare l'esistenza di una garanzia patrimoniale utile a far fronte a eventuali pretese risarcitorie di terzi.

In proposito, è stato opportunamente messo in luce che, diversamente rispetto ai soggetti che attualmente intrattengono rapporti nell'ambito del diritto privato, il robot non persegue un proprio interesse. L'interesse è pur sempre riconducibile a un diverso soggetto, inteso in senso tradizionale, che si avvale degli strumenti messi a disposizione dell'intelligenza artificiale. Questo fondamentale rilievo è sufficiente per affermare che non sussiste la necessità di creare nuovi soggetti giuridici. La tecnologia e l'intelligenza artificiale sono (e devono permanere) al servizio dei soggetti giuridici intesi in senso tradizionale. Peraltro, la netta presa di posizione non deve indurre il giurista a distogliere l'attenzione dai procedimenti decisionali, in parte imprevedibili, che caratterizzano le nuove tecnologie. Il profilo è di primaria importanza per comprendere come, nei diversi settori, il diritto privato debba disciplinare atti posti in essere con l'ausilio della tecnologia, che in alcuni settori si sostituirà alle condotte umane.

Alcune innovazioni sono già attuali e coinvolgono la materia cardine dei rapporti privatistici, ossia il contratto. Il pensiero è subito rivolto allo smart contract che, nel contesto del c.d. "decreto Semplificazioni" all'art. 8-ter (aggiunto dalla l. 12/2019 in sede di conversione del d.l. 135/2018)[124], è stato di recente definito come "un programma per elaboratore che opera su tecnologie basate su registri distribuiti e la cui esecuzione vincola automaticamente due o più parti sulla base di effetti predefiniti dalle stesse". In questo ambito, di indubbia rilevanza da un punto di vista economico, l'ambizione dei programmatori è quella di escludere completamente l'operatività del diritto. Secondo la visione più estrema, il codice dovrebbe prendere il posto del diritto, poiché è proprio il codice computerizzato – in queste ricostruzioni – a fungere da "diritto". In modo provocatorio Savelyev ha osservato inoltre che gli smart contracts rappresentano l'inizio della fine del diritto contrattuale[125].

Di fronte ad affermazioni di questo genere, i giuristi devono rispondere con cautela, venire a conoscenza degli aspetti tecnici e affrontarne i profili problematici. Si tratta di una tematica spesso intrisa di luoghi comuni, da superare facendo uso delle tradizionali categorie civilistiche. In questa prospettiva, è necessario mettere in luce che, nonostante il seducente appellativo, coniato negli anni novanta da Szabo, gli smart contracts non sono

[124] https://www.altalex.com/documents/news/2021/06/29/smart-contract-profili-di-qualificazione-giuridica

[125] Contract Law 2.0: «Smart» Contracts As the Beginning of the End of Classic Contract Law Higher School of Economics Research Paper No. WP BRP 71/LAW/2016, 24 Pages Posted: 14 Dec 2016

affatto intelligenti[126]. Essi eseguono semplicemente una prestazione al ricorrere di un determinato evento (trigger event), che opera come una sorta di condizione: una volta verificatosi l'evento, l'esecuzione del contratto diviene inesorabile, alla stregua delle operazioni di un distributore automatico che, in seguito all'immissione della moneta, dispensa il prodotto e l'eventuale resto. Lo smart contract in quanto tale non è dunque dotato di intelligenza artificiale e il procedimento automatizzato, allo stato, riguarda soltanto l'esecuzione e non la conclusione del contratto. Ne deriva che, a rigore, lo smart contract non potrebbe neppure considerarsi un contratto[127].

Le questioni più significative si pongono per gli smart contracts connessi alla tecnologia blockchain. Le transazioni avvengono su piattaforme informatiche che prendono il posto degli intermediari e assicurano che la corretta esecuzione del contratto, concluso attraverso l'incontro tra la proposta e l'accettazione, non possa più essere messa in discussione. Pertanto, in linea teorica, gli smart contracts non conoscono inadempimenti. Inoltre, sono in grado di generare vantaggi in termini economici, in virtù della eliminazione dei costi di transazione, principalmente connessi alle negoziazioni e alle attività di intermediazione. In questo quadro, il grande cambiamento rispetto al passato riguarda l'affidamento nelle capacità di adempiere dell'altro contraente. Chi compie una transazione su una piattaforma blockchain generalmente non conosce l'altro contraente, il suo affidamento concerne soltanto il funzionamento della piattaforma. In questo senso, si afferma che gli smart contracts sono self-executing e self-enforcing. In realtà, il procedimento non sempre è pienamente automatizzato poiché spesso le informazioni rilevanti per l'esecuzione del contratto vengono fornite da un terzo che, nel gergo tecnico, viene denominato oracle. L'affidamento dei contraenti deve pertanto estendersi anche a quest'ultima figura.

Nonostante le incomprensioni e le incongruenze, gli smart contracts avranno un impatto significativo nel mercato. Tuttavia, non corrisponde al vero che il codice si sostituirà al diritto; il diritto continuerà certamente ad essere applicabile. Eventuali falle nei sistemi operativi potrebbero agevolare fraudolente manomissioni da parte di soggetti terzi, che richiederebbero l'applicazione delle norme sulla responsabilità. Inoltre, non può escludersi che uno smart contract violi norme imperative o che venga concluso da un soggetto incapace di agire. Molto rilevante, anche in virtù dei risvolti per il mercato concorrenziale, si profila l'applicazione del diritto dei consumatori, che contiene norme

[126] https://www.4clegal.com/hot-topic/smart-contracts-breve

[127] The Formation of Blockchain-based Smart Contracts in the Light of Contract Law, di M. DJUROVIC, A. JANSSEN, in European Review of Private Law, Volume 26, Issue 6 (2018) pp. 753 – 771

idonee a tutelare la parte debole del rapporto contrattuale, limitando l'autonomia dei contraenti. Le riflessioni non dovrebbero riguardare soltanto l'astratta questione dell'applicabilità delle norme imperative, bensì essere volte ad individuare in che modo tali norme possano adattarsi a protocolli informatici. In questo senso, sarà necessario valutare l'applicabilità della direttiva 93/13/CEE, concernente le clausole abusive nei contratti stipulati con i consumatori. Non è escluso che la normativa europea possa applicarsi a codici informatici e che la tecnologia possa rendere più efficace il controllo sostanziale dei contratti nei rapporti B2C.

Le procedure di esecuzione automatizzata possono comportare un aumento dell'effettività del diritto. Oltre a quanto indicato con riguardo alle clausole abusive, un esempio sovente menzionato in dottrina riguarda la compensazione pecuniaria alla quale è tenuto il vettore aereo, in caso di ritardo o cancellazione del volo, ai sensi dell'art. 5, par. 3, del Regolamento CE n. 261/2004. Rendendo la tecnologia degli smart contracts obbligatoria per i vettori aerei sarebbe possibile assicurare ai consumatori gli indennizzi in via automatica. Attualmente, in media solo il 5 per cento dei passeggeri esercita i propri diritti in caso di ritardo o cancellazione del volo. Gli smart contracts potrebbero avere un impatto significativo sui rapporti tra vettori aerei e passeggeri e incidere sulle modalità di erogazione dei servizi. Da questo punto di vista, si teme che l'aumento di effettività del diritto possa rendere quest'ultimo insostenibile da un punto di vista economico. Se i vettori aerei fossero sistematicamente obbligati al pagamento dell'indennizzo in caso di ritardo o cancellazione del volo, si assisterebbe con tutta probabilità a un incremento significativo del prezzo dei biglietti aerei. In questo senso, in maniera paradossale, Basedow ha affermato che il diritto non sempre si presta ad essere pienamente effettivo.

Al di là di questa problematica, occorre rilevare che le nuove tecnologie potrebbero apportare un significativo miglioramento delle norme giuridiche, rendendole più adeguate alla soluzione del caso concreto. In base alla normativa europea relativa alla responsabilità del vettore aereo pochi minuti di differenza nella durata del ritardo possono assumere un'importanza decisiva ai fini del riconoscimento dell'indennizzo. Con la tecnologia degli smart contracts sarebbe possibile quantificare un indennizzo in proporzione al ritardo accumulato. Il problema specifico si collega a quello di carattere più generale concernente la personalizzazione delle norme giuridiche. Con l'avvento dei big data, i software potrebbero modellare le norme giuridiche sulla base delle peculiarità di ogni singolo consociato (c.d. "granular norms"). Secondo Casey e Niblett, ci potrebbe

determinare il superamento dell'utilizzazione di standard e clausole generali nel diritto privato[128].

I processi decisionali algoritmici assumono rilievo nel contesto dei mercati attraverso la pratica del c.d. 'High-frequency trading' (HFT). Si tratta di una modalità di intervento sui mercati compiuta con sofisticati strumenti software, e talvolta anche hardware, che permettono di porre in essere negoziazioni ad alta frequenza, guidate da algoritmi matematici, che agiscono su mercati di azioni, opzioni, obbligazioni, strumenti derivati, commodities. La durata di queste transazioni può essere brevissima, con posizioni di investimento fatte proprie per periodi di tempo variabili, da poche ore fino a frazioni di secondo. Lo scopo di questo approccio è quello di lucrare su margini estremamente esigui, anche pochi centesimi. Per ottenere significativi ricavi mediante margini minimi, la strategia HFT deve necessariamente operare su grandi quantità di transazioni giornaliere. L'HFT pu causare distorsioni nel mercato, in quanto gli scambi ad alta velocità conferiscono un vantaggio a chi li mette in atto rispetto a chi si serve di strumenti tradizionali. Inoltre, la continua azione degli operatori HFT tra mercati diversi sullo stesso titolo conferisce ai mercati un'estrema volatilità, senza tuttavia che i movimenti speculativi si consolidino su una posizione di investimento sul capitale azionario di specifiche aziende. Ciò dimostra che le operazioni prescindono da valutazioni sui fondamenti economici esibiti dalle aziende. La potenziale pericolosità dell'HFT per la stabilità dei mercati è confermata dal caso Goldman Sachs, uno dei più importanti operatori a far ampio uso della suddetta tecnologia. Un suo dipendente, Sergey Aleynikov, si è impossessato dei codici sorgente con cui la compagnia accedeva a simili operazioni di trading: la vicenda di Aleynikov si concluse con l'arresto da parte dell'FBI e l'applicazione di una pena detentiva.

La pericolosità dell'HFT è ulteriormente testimoniata da interventi normativi volti a porre un freno alle relative attività speculative e a garantire una certa trasparenza con riferimento alle modalità operative. Sotto questi profili, deve essere valutata con attenzione la revisione della direttiva europea, c.d. MiFID II (Markets in Financial Instruments Directive) e ulteriori regolamenti attuativi, che obbligano gli high frequency traders a registrarsi come imprese di investimento e a rendere pubblici i loro algoritmi, fornendo garanzie sull'attendibilità dei loro software. In ambito nazionale, la CONSOB segue con particolare concentrazione gli sviluppi tecnologici e ha pubblicato discussion

[128] A.J. CASEY, A. NIBLETT, Self-Driving Contracts, in 43 The Journal of Corporation Law, 2017, p. 101 ss.

papers e altri contributi nel tentativo di pianificare interventi regolatori idonei a fronteggiare le principali problematiche a livello interno e sovranazionale.

Da un altro angolo di visuale, i problemi relativi alla responsabilità civile investono soprattutto l'applicabilità delle norme relative all'illecito in ipotesi in cui una scelta, che si rivela dannosa, sia stata compiuta in base alle elaborazioni di un algoritmo. Si pongono problemi soprattutto in ordine alla prova della sussistenza degli elementi della responsabilità civile. Più in generale, sono in discussione le funzioni della responsabilità civile, che dovrebbe continuare a garantire una certa carica deterrente nei confronti dei possibili tortfeasors e, al contempo, rispondere all'esigenza di indennizzare adeguatamente i soggetti danneggiati. Le nuove tecnologie rendono necessaria l'elaborazione di nuovi parametri per attribuire la responsabilità e, posta la diversa gestione del rischio, impongono ai giuristi di ripensare alcune norme concernenti la responsabilità oggettiva.

Per quanto attiene ai parametri di riferimento per attribuire la responsabilità, si tratta di valutare la condotta di sistemi in grado di apprendere autonomamente e di decidere sulla base dei dati raccolti. Un primo parametro utilizzabile è quello relativo alle capacità umane. In presenza di un danno causato da un sistema operativo autonomo, ci si dovrebbe chiedere se un uomo nella stessa posizione sarebbe stato in grado di evitare il danno. La soluzione è criticata da parte della dottrina, poiché il software dovrebbe avere capacità superiori all'uomo e, in molti casi, le due condotte non sarebbero paragonabili, posto che le nuove tecnologie possono generare rischi diversi rispetto a quelli connessi a una condotta umana. In ogni caso (come osservato da Teubner, 2018), le capacità dell'uomo potrebbero costituire – almeno nel primo periodo – una soglia minima di "diligenza" richiesta al sistema operativo. In prospettiva, Chagal-Feferkorn propone di sviluppare un parametro autonomo per i sistemi intelligenti, denominato "algoritmo ragionevole", diverso a seconda del settore preso in esame. Altri autori, come Wagner, concentrano invece l'attenzione sui dati statistici concernenti le attività esercitate da software dotati di intelligenza artificiale.

Non mancano voci critiche in merito alla possibilità di sviluppare parametri adeguati a valutare la condotta dei sistemi operativi autonomi. Il diritto potrebbe assicurare la tutela dei danneggiati mediante normative speciali che prevedono ipotesi di responsabilità oggettiva. In questo contesto, emerge il problema della responsabilità del produttore con riguardo a sistemi operativi che apprendono autonomamente. Rispetto alle nuove tecnologie, le norme di conio europeo, contenute nella direttiva 85/374/CEE sulla

responsabilità del produttore, appaiono del tutto inadeguate poiché si riferiscono a tecnologie obsolete e non chiariscono con precisione in quali casi un sistema operativo in grado di apprendere autonomamente è da considerare difettoso. Allo stato, un gruppo di lavoro istituito dalla Commissione si sta occupando dello sviluppo di guidelines, che dovrebbero permettere di orientare l'interprete chiamato ad applicare le norme della direttiva.

Un ambito peculiare in cui si presentano le descritte problematiche è quello della responsabilità da circolazione di veicoli a guida autonoma, in cui è sotto osservazione l'apparato di regole relative alla responsabilità del proprietario del veicolo, alla responsabilità del produttore e alla assicurazione obbligatoria (v. 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", del 17 maggio 2018). Da più parti si discute di questioni che attengono a diversi soggetti interessati: i produttori, gli utenti, le compagnie assicurative. Pur afferendo all'area del diritto privato, il settore dei veicoli a guida autonoma incide certamente su interessi generali poiché, allo stato, riguarda le norme sulla responsabilità civile che trovano più sovente applicazione e concerne interessi economici di notevole rilevanza.

In un primo periodo, come è suggerito dalla suddetta Comunicazione, le attuali norme relative alla responsabilità da circolazione di veicoli, assistite dalla previsione dell'assicurazione obbligatoria, potrebbero continuare a fornire una regolamentazione soddisfacente. In prospettiva futura, occorrerà riflettere su nuove forme di responsabilità che coinvolgano maggiormente i produttori delle vetture. Questi ultimi sono infatti i soggetti che immettono i sistemi operativi nel mercato e gestiscono i rischi connessi al funzionamento del software. Il tema è affrontato a livello globale e, tenuto conto delle principali tesi, si tratterà di decidere se modificare o interpretare diversamente le norme in materia di responsabilità del produttore (come suggerito da Geistfeld, 2017), oppure se creare un nuovo sistema di responsabilità no-fault, in cui i danneggiati dovrebbero essere risarciti da un apposito fondo istituito con contributi erogati dai produttori di veicoli a guida autonoma (come proposto da Abraham-Rabin, 2019).

8. Intelligenza artificiale e diritto penale

L'impatto dell'intelligenza artificiale. e delle questioni ad essa connesse, appare rilevante anche con riferimento al diritto penale.

Sul piano della responsabilità per condotte penalmente rilevanti realizzate da soggetti non umani, sembra opportuno avviare al più presto una discussione obiettiva, priva di preconcetti, nella consapevolezza del rischio rappresentato da una deriva verso forme di responsabilità penale oggettiva in capo a coloro che hanno progettato e attivato il soggetto agente non umano.

Si tratta di una discussione che sembra spingersi al di là della realtà contingente. Tuttavia occorre liberarsi di ogni condizionamento distopico per accettare che numerosissime sono già oggi le fonti decisionali automatizzate le cui conseguenze possono violare beni giuridici tutelati dalla legge penale. Ferma la granitica previsione dell'art. 27, comma 1, Cost. – che ha di fatto impedito, ad oggi, il riconoscimento di responsabilità penale in capo alle persone giuridiche (con il conseguente ricorso al concetto di responsabilità amministrativa dell'ente per il reato commesso nel suo interesse) – il problema dell'imputazione di responsabilità per scelte riferibili a decisioni automatizzate non fa parte, appunto, di un futuro distopico, ma della realtà. Indubbiamente la questione pone in discussione l'intero concetto di pena e coinvolge nella discussione prospettive filosofico-culturali che trascendono la dogmatica penalistica, innervandosi ancora una volta nel dibattito sul riconoscimento di una soggettività giuridica. Il rischio da evitare, in primis, è quello di rivolgersi a paradigmi di tipo oggettivo che, formalmente estranei al diritto penale, talvolta si affermano attraverso l'intervento giurisprudenziale, per mezzo dell'amplificazione di posizioni di garanzia pur previste dalla legge.

10. Intelligenza artificiale e processo

Non va sottaciuto il rilevante contributo che l'intelligenza artificiale può fornire nel campo dell'organizzazione del servizio giustizia, del funzionamento dei sistemi giudiziari nazionali nonché del diritto processuale. Ivi l'applicazione delle tecnologie algoritmiche può indubbiamente rappresentare un fattore di miglioramento dell'efficienza e della qualità dei processi, sebbene l'implementazione della tecnologia debba avvenire in modo responsabile e nel rispetto dei diritti fondamentali garantiti.

L'intelligenza artificiale applicata al settore dell'azione pubblica nel contesto della produzione, diffusione, valutazione e miglioramento dei servizi pubblici e dei beni collettivi, è intesa come un insieme di dispositivi di carattere tecnologico, tecnico e matematico la cui regolazione va distinta per le fasi di sviluppo, utilizzo e monitoraggio e per i livelli di conoscenza fattuale, di architettura e di agency che esso implica. Nel

settore dell'azione giudiziaria, in particolare, le promesse della giustizia predittiva si sono riverberate nella espansione delle cosiddette legal tech, nell'aumento della attenzione degli attori internazionali per la elaborazione di standard di accountability e responsiveness (di sviluppo ed uso), nella sperimentazioni in taluni ordinamenti di strumenti di aiuto alla decisione del giudice, di elaborazione in via preliminare del potenziale di costo/beneficio di un contenzioso e della sua soluzione giurisdizionale.

Un utilizzo avanzato di questo insieme di strumenti ad alta intensità tecnologica, articolati sulla applicazione della scienza matematica e della scienza informatica, è in essere solo in alcuni Paesi europei. Eppure già nella distribuzione automatizzata dei casi, nella analisi di banche dati giurisprudenziali (anche se non realizzata attraverso la traduzione del linguaggio naturale in una rappresentazione digitale) e nella gestione ponderata dei carichi di lavoro e dei flussi, lo strumento dell'automazione ha fatto la sua entrata nell'organizzazione del servizio giustizia ormai da tempo.

Sul piano della qualità della giustizia le promesse e le aspettative legate all'applicazione di strumenti di intelligenza artificiale vanno nella direzione di: (i) riduzione del contenzioso attraverso rimedi alternativi automatizzati; (ii) riduzione dei margini di errore nella valutazione preventiva del rischio di soccombenza (tale profilo riguarda in particolare le professioni ordinistiche); (iii) riduzione dei tempi attraverso la trattazione in via automatizzata delle controversie seriali e standardizzabili; (iv) riduzione dei margini di differenziazione distrettuale e circondariale per tipologie di risposta a simili tipologie di contenzioso.

Muovendo dai risultati conseguiti da società già operanti nel settore è possibile affermare che, attualmente, esistono software in grado di anticipare correttamente l'esito di un giudizio. Il dato assume rilevanza nell'ambito delle professioni forensi e della gestione del contenzioso, posto che i software potrebbero essere in grado – quantomeno – di influire sulla scelta del privato di agire in giudizio. Una prospettiva maggiormente di frontiera auspica l'utilizzazione di software predittivi per la soluzione di questioni di natura bagatellare.

La massima aspirazione di alcuni fautori dell'intelligenza artificiale. è quella di eliminare del tutto il ruolo del diritto nell'espletamento di significative attività dell'uomo. Ci potrebbe implicare l'adozione di meccanismi alternativi di risoluzione delle controversie completamente automatizzati, eventualmente collegati a piattaforme blockchain (Ortolani, 2019). L'obiettivo, da molti considerato utopistico, è quello di creare un "giudice robot" (il c.d. "robo-judge"), in grado di decidere controversie sulla base

dell'elaborazione statistica di dati. Sebbene siano state messe in evidenza le potenzialità della tecnologia negli accertamenti del fatto compiuti in ambito civile, la ricerca scientifica si presenta ancora limitata con riferimento al possibile ruolo dell'intelligenza artificiale.

Con la transizione da applicazione algoritmiche deterministiche ad applicazioni probabilistiche, la decisione non deriva semplicemente dalla pedissequa applicazione del ragionamento deduttivo, ma è basata sul contemporaneo apprezzamento di una pluralità di interessi ed istanze che vengono in rilievo nel caso concreto. Si realizza in tal senso un passaggio sintetizzabile nella formula dalla delega di processo alla delega di decisione, che implica necessariamente l'accettazione dell'esistenza del margine di errore. La tecnologia potrebbe assicurare maggiore velocità e un significativo risparmio di risorse. Come in altri campi, le innovative soluzioni che la tecnologia è in grado di offrire per la soluzione delle controversie civili dovranno necessariamente conseguire un elevato livello di accettazione sociale prima di essere implementate. Merita attenzione, altresì, l'impatto che i sistemi di open access, associati a strumenti algoritmici predittivi, possono produrre sul piano del valore del precedente e dell'indipendenza degli organi giudicanti.

La regolazione etica e giuridica di tali profili deve tenere conto degli snodi organizzativi entro cui domanda ed offerta di giustizia si incontrano. Quattro appaiono le dimensioni funzionali che necessitano di essere poste alla attenzione della norma etica e giuridica in questo contesto: conoscenza, rito del processo, status dell'organo terzo dirimente le controversie, dati personali. I principi che vanno richiamati in tale senso sono: (i) responsabilità professionale e trasparenza della produzione della conoscenza; (ii) applicazione delle garanzie processuali di difesa; (iii) autonomia della giurisdizione e indipendenza del giudice; (iv) privacy e sicurezza dei dati. Essi possono essere declinati in norme etiche e in norme giuridiche.

Sul piano etico un codice deontologico che assicuri la responsabilità professionale degli sviluppatori e di coloro che intervengono nel processo di integrazione e applicazione dei dispositivi di automazione appare necessario e in linea con le forme di regolazione delle expertise peritali e di consulenza di cui ci si avvale in via endoprocedimentale. Ancora la estensione delle garanzie processuali massime – in tal senso il caso ordinamentale italiano appare in una prospettiva comparata di sicura ispirazione – per i riti nei quali le parti possono avvalersi (con riconoscimento nel processo) di strumenti di giustizia predittiva deve essere assicurata con il più alto livello di tutela costituzionale.

L'autonomia del giudice a fronte della disponibilità di conoscenza induttiva derivata dalla analisi di banche dati giurisprudenziali va assicurata sia con norme che obblighino le giurisdizioni supreme a un set di standard comuni per il trattamento di tale conoscenza da parte delle giurisdizioni di primo e secondo grado, sia con norme che integrino nella argomentazione del giudice la esplicitazione della fonte da cui tale conoscenza viene desunta. Sul piano giuridico ordinamentale dovrebbero essere previste delle unità specializzate a livello di giurisdizioni di secondo grado – dove il giudizio di fatto trova la sua sintesi – per la valutazione dei dispositivi di analisi delle banche dati. Infine, la qualità dei dati e la loro tutela va prevista sul piano giuridico. Sub condicione di anonimato del giudice, le banche dati devono rispondere a sistemi di regolazione pubblica responsabili della sicurezza. In tal senso, andrebbe previsto in via legislativa la disposizione nei ministeri di uffici preposti allo sviluppo reti e politiche di mantenimento della sicurezza. Le garanzie di privacy dovranno essere conformi alle norme sovranazionali in materia (GDPR).

Anche la sfera del processo – civile, penale, amministrativo, contabile – è chiamata a confrontarsi con le implicazioni del digital turn, non soltanto per via del senso, assai diffuso, di ineluttabile transizione verso un quadro in cui tutte le attività umane debbono necessariamente offrirsi alla rivoluzione basata sull'inedito connubio tra incommensurabile potenza computazionale e big data. É, infatti, incontestabile la condizione di grave inefficienza e inefficacia della giustizia, che pone tutti gli operatori nella posizione di dover studiare come le potenzialità dell'era digitale e, in particolare, l'intelligenza artificiale, possano eventualmente migliorare l'attuale situazione.

In particolare, con riferimento al problema della prova, con considerazioni mirate soprattutto sul processo penale ma estensibili, con opportune distinzioni, al processo in generale, può essere osservato quanto segue.

L'impiego di strumenti algoritmici nella ricerca e nella formazione della prova implica una inevitabile opacità del processo di creazione del dato probatorio.

Occorre, pertanto, per bilanciare l'asimmetria conoscitiva tra le parti del processo, valorizzare in pieno il principio di parità delle armi.

Questa asimmetria conoscitiva, fortemente radicata nella tradizione, post-inquisitoria, italiana e di molti altri Stati europei continentali, è inevitabilmente legata alla presenza di una parte pubblica nell'indagine e nel processo penale. Essa rischia di essere oltremodo amplificata dal ricorso a mezzi di ricerca della prova e mezzi di prova la cui opacità, accentuata dall'impiego di soluzioni di machine learning o deep learning, può sottrare

completamente la prova stessa al confronto dialettico delle parti sulla sua attendibilità. Ci trasformerebbe la "prova algoritmica" in un mezzo apoditticamente attendibile e, quindi, il giudice, in mero annotatore di un processo di valutazione della prova interamente assorbito dalla natura algoritmica della medesima.

Allo scopo di attuare più efficacemente le suddette considerazioni, occorre inoltre identificare e distinguere, all'interno della "prova algoritmica", i profili di rischio di inattendibilità legati al piano della teoria scientifica sintetizzata nell'algoritmo e quelli legati, invece, alla costruzione e al funzionamento dell'algoritmo stesso. Risulta evidente dalla esperienza di alcune giurisdizioni locali statunitensi che l'impiego processuale di strumenti di natura predittiva nasconde plurimi livelli di opacità e di rischio, dipendenti non soltanto dalla natura digitale e automatizzata dello strumento probatorio. Infatti, il software che confeziona il dato utilizzato con valore probatorio all'interno del processo, innanzitutto, si basa su una teoria scientifico-probabilistica che potrebbe non rispondere ai parametri "Daubert" i quali, trascendendo la sfera della Corte Suprema statunitense, sono divenuti, nel tempo, una sorta di decalogo per l'ammissibilità della prova scientifica nella maggior parte degli ordinamenti occidentali. Il primo profilo di contraddittorio processuale deve riguardare, allora, la validità e la validazione della teoria scientifica sottesa all'algoritmo da parte della comunità scientifica di riferimento. Il profilo dell'opacità dell'algoritmo in cui essa viene sintetizzata rappresenta un secondo, ineluttabile, tema di confronto tra le parti, in contraddittorio, davanti al giudice imparziale.

Un altro passaggio consiste nel rivalutare la tradizionale distinzione tra fase cognitiva ed esecutiva, in relazione all'impiego di valutazioni predittive del comportamento violento e del rischio di recidivanza. In diretta connessione con quanto sottolineato nel punto precedente, si osserva un crescente ricorso a strumenti algoritmici di predizione del rischio di recidivanza e/o di comportamenti violenti (come anche nel caso State v. Loomis già ricordato), basati su teorie bayesiane che sfruttano correlazioni ottenute dall'analisi automatizzata di dati di varia provenienza. Come accennato, l'area geografica di riferimento, al momento, è per lo più quella nordamericana e australiana, ma con progressive incursioni anche in Europa (in particolare, Inghilterra e Galles). Poiché tali strumenti predittivi si basano su teorie psico-criminologiche più o meno accettate dalla comunità scientifica, il vaglio di cui si accennava anche supra rimane essenziale. Altrettanto essenziale è la rivalutazione del ruolo che la prova dell'attitudine caratteriale e psicologica può avere nel processo penale. Si tratta di un elemento certamente rilevante

nella quantificazione della pena e nella modulazione del trattamento sanzionatorio in concreto, ma non certo nell'accertamento dei fatti oggetto del singolo procedimento penale. In questa prospettiva, la tradizionale distinzione che vede l'Italia e altri Paesi continentali opporsi all'idea che la valutazione caratteriale possa incidere sull'accertamento della responsabilità per i fatti accaduti torna ad assumere grande significato. La prova (predittiva, algoritmica, automatizzata) caratteriale può acquisire elementi conoscitivi su come l'imputato potrebbe essere portato a comportarsi in futuro, ma nulla dice rispetto alla responsabilità di questi rispetto ai fatti di cui è accusato.

Sul diverso piano dell'indipendenza del giudice merita di essere preso in considerazione attentamente l'impatto che i sistemi di open access, associati a strumenti algoritmici predittivi, possono produrre sul piano del valore del precedente e dell'indipendenza degli organi giudicanti. L'accesso completo e automatizzato alle decisioni giudiziarie rappresenta un tassello del più ampio movimento per l'openness delle pubbliche amministrazioni che in alcuni ordinamenti, come ad esempio, la Francia, è già stato assicurato da tempo. Certamente tale strumento agevola molto i professionisti legali, favorendo una migliore conoscenza e circolazione anche delle decisioni di merito. Pertanto, tale esteso bacino di sentenze presenta fortissima appetibilità per chi sia interessato a costruire, a beneficio – principale ma non esclusivo – delle law firms, strumenti predittivi della decisione giurisdizionale. Alimentando una rete neurale con l'insieme di tutti i precedenti, pur non vincolanti, si possono stabilire, correlazioni utili per immaginare come il giudice si pronuncerà rispetto a una determinata questione e all'interno del collegio che concorre a comporre.

Ciò può rappresentare, come accennato, un grande vantaggio per la strategia difensiva e per l'allocazione delle risorse all'interno dello studio legale. Esso ha diverse ricadute assai perniciose: 1) le correlazioni stabilite possono non essere corrette; 2) l'evoluzione dell'interpretazione giurisprudenziale sarebbe fortemente compromessa. Il ricorso a simili strumenti scoraggerebbe, infatti, il patrocinio di posizioni potenzialmente 'non vincenti'. Rischierebbe così di essere depotenziato l'essenziale fattore di evoluzione che deriva dal tradizionale impegno dell'avvocatura nel promuovere nuovi paradigmi interpretativi che, scalando la struttura delle corti, arrivano ad affermarsi come giurisprudenza dominante, di legittimità, e talvolta, a trasformarsi in modifiche normative; 3) anche nei sistemi continentali che non sono basati sul valore del precedente, il diffondersi di strumenti di quantitative legal prediction può provocare un effetto di centralizzazione del valore del precedente. Ciò può indurre il singolo giudice persona

fisica a temere conseguenze di vario genere (disciplinare, civile) per essersi discostato, con la sua decisione, dall'esito preconizzato dallo strumento predittivo. L'effetto potrebbe inizialmente indurre un apparentemente neutro onere suppletivo di motivazione, per distaccarsi dal precedente, come accade negli ordinamenti di common law, che potrebbe tuttavia progressivamente incidere sul senso di indipendenza del giudice. Questi, infatti, potrebbe sentirsi, più o meno consciamente, indotto ad aderire sistematicamente al risultato della predizione.

Alla luce delle suddette osservazioni, bisogna distinguere attentamente il concetto di conoscibilità dei precetti e di prevedibilità della sanzione, essenziali a soddisfare il moderno concetto di legalità, da quello di "prevedibilità" della decisione del giudice. Infatti, se è vero che la più moderna concezione del principio di legalità passa per il riconoscimento dell'essenzialità della garanzia di conoscibilità del precetto normativo da seguire e di prevedibilità della sanzione che sarà applicata in caso di accertata trasgressione, si deve riconoscere che lo scenario evocato dalla quantitative legal prediction non è ispirato all'attuazione della predetta garanzia o che, quantomeno, pu sovrapporvi effetti contrari e perniciosi.

Sempre con riferimento alla giurisdizione, centrale appare il ruolo del processo amministrativo, destinato a divenire, ad un tempo, il luogo del sindacato sul potere dell'amministrazione che si serve di algoritmi e il luogo dell'esame del corretto esercizio dell'eventuale potere del sistema pubblico di regolare e conformare l'uso, da parte dei soggetti privati, dell'intelligenza artificiale.

Sotto il primo profilo, assume particolare rilievo il sindacato sulle c.d decisioni algoritmiche dell'amministrazione. Pur nelle comprensibili oscillazioni, la giurisprudenza amministrativa, come sopra visto, ha già avuto modo di precisare le condizioni fondamentali per assicurare la qualità di tali decisioni. Quel che val la pena di sottolineare è che, ai fini della legittimità del provvedimento basato sull'utilizzazione degli algoritmi, deve essere assicurata una piena accessibilità di questi ultimi – assimilabili ad un concetto tecnico e non ad un concetto giuridico indeterminato – alla loro formazione e alla loro evoluzione, anche con riferimento alla loro provenienza.

Sotto il secondo profilo, deve essere osservato che esso riguarda uno scenario ancora lontano dall'essere realizzato. Quel che forse, in questa sede, può essere anticipato è che potrà costituire un utile ausilio la giurisprudenza del giudice amministrativo sull'attività di regolazione, che richiede un supplemento di partecipazione, anche in funzione di controllo dei soggetti interessati (c.d legalità procedurale) e sul sindacato delle valutazioni

tecniche. Mediante tale sindacato, difatti, potrà essere valutata con attenzione la maggiore attendibilità delle scelte operate dall'autorità amministrativa.

9. Proseguimento dell'indagine

Se fino a pochi anni fa il principale problema di tutti gli scienziati coinvolti nella ricerca relativa all'Intelligenza Artificiale era quello di poter dimostrare la realistica possibilità di utilizzare sistemi intelligenti per usi comuni, oggi che questo obiettivo è ampiamente raggiunto ci si chiede spesso quale possa essere il futuro dell'Intelligenza Artificiale. Sicuramente molta strada deve essere ancora fatta, soprattutto in determinati settori, ma la consapevolezza che l'Intelligenza Artificiale oggi rappresenta una realtà, e non più un'ipotesi, e i dubbi sono soprattutto relativi alle diverse possibilità di utilizzo dei sistemi intelligenti e al loro impatto sul tessuto sociale ed economico.
Secondo una definizione accettata a livello internazionale, l'Intelligenza Artificiale è quella disciplina, appartenente all'informatica, che studia i fondamenti teorici, le metodologie e le tecniche che consentono di progettare sistemi hardware e sistemi di programmi software capaci di fornire all'elaboratore elettronico delle prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana.
La rapida evoluzione dell'Intelligenza Artificiale e le sempre più numerose applicazioni che essa trova nei diversi settori della vita quotidiana impongono una profonda riflessione sulle sue implicazioni in ambito giuridico. Qui interessano, in particolare, quelle riguardanti il diritto penale (sostanziale), dove lo straordinario sviluppo dell'Intelligenza Artificiale dell'ultimo decennio ha sollevato questioni assai delicate.

Prima di concludere, è utile fornire alcune indicazioni su come abbiamo strutturato il presente lavoro ed i relativi profili di interesse.
I tre principali profili critici del connubio tra Diritto Penale e "artificial intelligence" sono:
- Il quesito della legittimità dell'utilizzo di algoritmi predittivi come mezzo di prevenzione della criminalità, tanto nelle strategie di compliance aziendale quanto nell'attività di polizia;
- La delicata tematica concernente il ricorso ad algoritmi nei processi decisionali e, in particolare, nelle diverse fasi di risk assessment del processo penale per calcolare il tasso di pericolosità sociale e lo specifico rischio di recidiva;

- Il problema relativo alla imputazione della responsabilità, allorquando un reato sia commesso dal sistema intelligente che abbia agito al di fuori della sfera della signoria dell'uomo.

L'importanza del tema è stata percepita anche dalla comunità internazionale e, segnatamente, dal Consiglio d'Europa, impegnato a promuovere iniziative volte a uniformare la legislazione in materia di diritto ed IA. La base normativa su cui fonda l'iniziativa seminariale sono senz'altro la "Carta etica per l'uso dell'intelligenza artificiale nel sistema di giustizia penale e nei relativi ambienti" del 2018, e l'istituzione del CAHAI (Ad Hoc Committee on Artificial Intelligence) nel 2019.

A tal fine, risulterà decisivo adottare un approccio multilevel al tema, e si andranno necessariamente ad analizzare ed approfondire contributi e studi provenienti da altri ordinamenti senza dubbio precursori nel presente ambito.

CAPITOLO II
IA E ATTIVITÀ DI LAW ENFORCEMENT

1. Smart city e predictive policing; 2. AI e Law Enforcement; 3. La polizia predittiva: ratio ed impiego; 3.1. Sistemi di individuazione degli hotspots; 3.2. Sistemi di crime linking; 4. Il valore delle informazioni raccolte tramite i tools; 5. Le criticità; 6. Un modello di polizia predittiva: XLAW

1. Smart city e predictive policing

"Signor Marks in nome della sezione precrimine di Washington D.C. la dichiaro in arresto per il futuro omicidio di Sarah Marks e Donald Dubin che avrebbe dovuto avere luogo oggi 22 aprile alle ore 8 e 04 minuti"[129].

Un complesso sistema chiamato Precrimine riesce a prevedere i crimini ancor prima del loro accadimento, aiutando così la polizia ad arrestare i presunti colpevoli, punendo di fatto la sola intenzione di consumare il fatto, ma non che lo stesso sia mai consumato.

Siamo nella Washington del 2054, e questa è la trama di un film[130].

Nel mondo interconnesso e intelligente che ha sviluppato gli attuali modelli di progettazione delle smart city - aree urbane ben definite capaci di implementare le soluzioni tecnologiche ai bisogni della cittadinanza -, una condizione fondamentale per l'elaborazione di strategie efficienti con previsioni attendibili, da parte dei software predittivi per il contrasto della criminalità, è rappresentata dalla disponibilità di una enorme quantità di dati, informazioni raccolte e analizzate poi da sistemi predittivi; di questi dati, una parte interessante viene già catturata tramite l'utilizzo della sensoristica in campo (videosorveglianza, microfoni ambientali, sensori di misura, etc), raffigurata fisicamente dai device IoT (Internet of Things).

Ebbene proprio in questo contesto i sociologi hanno osservato come la crescita del fenomeno delle città intelligenti si stia riflettendo sulla sperimentazione delle tecniche di polizia predittiva.

[129] Dal film: "Minority Report", di Steven Spielberg

[130] Si veda "Come in Minority Report con Tom Cruise: l'IA prevede i crimini negli Usa", di C. Guzzonato, il 14 luglio 2022

Se, da un lato, una smart city permette di incrementare i Big Data a disposizione dei software predittivi destinati alle operazioni di polizia, dall'altro, questi ultimi, ne condividono la strategica capacità, l'indubbia efficienza nella prevenzione e nel contrasto della criminalità.

Ora, tornando al tema centrale, possono le forze di polizia utilizzare i "Big Data" come strumento di analisi per migliorare l'efficienza dei propri compiti istituzionali e dei risultati attesi?
Lo strumento tecnologico della predictive policing è rappresentato dall'analisi dei dati acquisiti dalle fonti più disparate, quindi utilizzandone i risultati si può tentare, con le dovute cautele, di prevenire e rispondere nella maniera più favorevole possibile al contenimento della consumazione di probabili attività illecite.

L'idea di anticipare propositi criminosi piuttosto che rispondere semplicemente ad essi, risale agli anni 2000, quando il "the Office of Justice Assistance" in collaborazione al "Los Angeles Police Department" tennero il "Predictive Policing Symposium" per approfondire questa avveniristica idea, e il suo impatto sul futuro professionale degli stessi operatori.
Un simposio che vide il fattivo coinvolgimento di ricercatori, criminologi, sociologi, specialisti del diritto, funzionari di diversi dipartimenti di polizia, coralmente impegnati nell'esplorare le possibili implicazioni politiche, la violazione dei diritti civili, l'impatto sulla sfera privacy, sul trattamento dei dati personali, che la tecnologia dell'investigazione predittiva potesse generare.

Inoltre, un ulteriore pericolo di queste metodologie da non sottovalutare fu correlato direttamente all'uso di processi matematici, appunto predittivi ed analitici, utilizzati nelle attività di polizia per mettere in luce qualsiasi potenziale attività criminale; metodi, peraltro, che ricadono all'interno di svariate categorie di servizio: previsione dei reati, previsione dei colpevoli e/o la loro identità, classificazione delle vittime di un reato.
Del resto, se riflettiamo attentamente, per prevenire un qualcosa bisogna anzitutto prevederlo, dunque, quale migliore aiuto di quello dato dai modelli matematici, dagli algoritmi specializzati, dall'intelligenza artificiale, tutti ingredienti contenuti, appunto, in specializzati software di analisi predittiva.

Tuttavia, questa "anticipazione analitica" rischia però di spostare troppo le forze dell'ordine dal concentrarsi su ciò che è accaduto verso cosa potrebbe accadere, rischiando di distribuire, e con scarsa efficacia, le risorse umane sul fronte del contrasto al crimine, trasformandone negativamente i risultati.

Charlie Beck, Chief of Los Angeles Police Department, sottolineava positivamente da sempre l'uso delle tecniche predittive nella scoperta dei reati, infatti l'esame analitico di determinati dati hanno ben aiutato i funzionari del suo dipartimento nell'anticipare con precisione taluni eventi criminosi[131].

Nello stesso tempo, però, metteva in guardia tutti dall'uso distorto di queste metodologie, che non vanno mai intese, o ancor peggio confuse, come la sostituzione futura di tutte le tecniche investigative di polizia collaudate e certificate ex lege, ovverosia, quelle procedure codificate basate sull'evidenza dei fatti, il riscontro probatorio, sulle analisi di intelligence.

Tuttavia, quanto detto fin qui ci spinge ad una ulteriore considerazione, vale a dire, come nel dominio della pubblica sicurezza l'uso dell'intelligenza artificiale, per la previsione in realtime di plausibili attività illecite, rappresenti oggi una priorità molto sentita anche all'interno della comunità scientifica, impegnata da sempre nello sviluppare modelli statistici in un crescendo di efficacia e affidabilità.

Modelli di algoritmi sviluppati con una capacità di analisi concentrata essenzialmente su due principali obiettivi: l'individuazione in primis del luogo e del momento di una probabile azione criminale, indirizzandosi successivamente poi sulle cause e le possibili vittime.

Per far ciò, la maggior parte dei progetti di predictive policing utilizzano informazioni storicizzate, ma riferite a dati certi e disponibili, calibrate su accadimenti non troppo lontani, che opportunamente analizzate con specifiche tecnologie, possono delineare, prevenire, individuare, il trend di particolari comportamenti criminogeni.

Ad esempio, una tecnica di machine learning, utilizzando dati e algoritmi statistici, ci permetterà di creare modelli heat maps necessari per poter comprendere il possibile momento temporale di un accadimento violento in luoghi che, secondo calcoli matematici, saranno gli scenari ideali (crime mapping) degli eventi; peraltro, proprio su tali previsioni si potranno modellare i mirati servizi di controllo, su territori classificati ad "alto rischio".

[131] "Comprendere la polizia predittiva: cos'è e come funziona", a cura di Giovanni Villarosa, il 14 Dicembre 2022

Altro esempio di modello computazionale è rappresentato dal risk terrain modeling (RTM), software specializzato nella previsione di possibili attività di spaccio di sostanze stupefacenti circoscritte a determinati territori urbani.

Anche in Italia, da molto tempo, diversi operatori della pubblica sicurezza si stanno impegnando nella progettazione diretta di sistemi intelligenti a supporto dei vari uffici investigativi, nell'attenta opera di prevenzione dei crimini e del controllo del territorio, per contrastare tutti i possibili fenomeni di assembramenti violenti, con un focus particolare su quella predatoria, molto più diffusa, e per l'allarme sociale che ingenera.

Dal 2007 ad oggi sono stati sviluppati e impiegati dai nostri funzionari di polizia due software in particolare, operativi in diverse città italiane: il Key Crime, utilizzato dalla Questura di Milano, e l'altro, XLaw, un programma sviluppato specificatamente per la classe dei reati predatori, sperimentato dapprima dalla Questura di Napoli su alcuni quartieri, ma successivamente esteso ai territori urbani più ampi, come le città di Prato, Salerno, Venezia, Modena, Parma.

2. AI e Law Enforcement

Nel documento di presentazione del Convegno annuale di esperti di Polizia, organizzato dall'OSCE, dedicato quest'anno (2019) proprio al tema "Artificial Intelligence and Law Enforcement", può leggersi quanto segue:

«nei loro sforzi per aumentare l'efficienza e l'efficacia e per stare al passo con le innovazioni tecnologiche, le autorità e le agenzie di law enforcement di tutto il mondo stanno esplorando sempre più i potenziali dell'IA per il loro lavoro. La crescente quantità di dati ottenuti e archiviati dalla polizia ha anche richiesto metodi e strumenti più sofisticati per la loro gestione e analisi, per l'identificazione di modelli (pattern), la previsione dei rischi e lo sviluppo di strategie per allocare le risorse umane e finanziarie dove sono maggiormente necessarie. Anche se l'uso dell'IA nel lavoro delle forze dell'ordine è un argomento relativamente nuovo, alcuni strumenti basati sull'intelligenza artificiale sono già stati testati e sono persino attivamente utilizzati dai servizi di polizia di diversi Paesi del mondo. Questi includono software di analisi di video e immagini, sistemi di riconoscimento facciale, di identificazione biometrica, droni autonomi e altri robot e strumenti di analisi predittiva per prevedere le "zone calde" del crimine o anche

per identificare potenziali criminali futuri, in particolare i criminali ad elevata pericolosità»[132].

L'impiego di sistemi di IA nelle attività di law enforcement è, quindi, già una realtà, e anzi se ne prevede una crescita ed intensificazione nei prossimi anni a vari livelli[133]. Del resto, l'importanza strategica dell'impiego di sistemi di IA nelle attività di law enforcement e i preziosi risultati grazie ad essi raggiungibili, sono ben messi in evidenza da un episodio riferito da Giuseppe Italiano in un suo recente saggio:

«già nel recente passato si sono verificati casi in cui l'utilizzo di opportune (anche semplici) analisi algoritmiche avrebbe potuto prevenire il verificarsi di pericolosi eventi terroristici. Ad esempio, è famoso il caso di Umar Farouk Abdulmutallab, noto anche come il "terrorista delle mutande" (Underwear Bomber), che è riuscito a imbarcarsi sul volo Amsterdam-Detroit nel giorno di Natale 2009, con dell'esplosivo cucito all'interno della biancheria intima che indossava, e che ha cercato di farsi esplodere durante il volo. Per una fortunata coincidenza [l'intervento di alcuni passeggeri insospettitisi], l'attacco terroristico non è andato a buon fine [...]. L'intelligence aveva dati e informazioni a sufficienza per valutare il grado di pericolosità del terrorista e aveva anche elementi sufficienti per inserirlo nella black list, così da negargli la possibilità di imbarco su voli diretti negli Stati Uniti. Ma in quel caso l'intelligence non è semplicemente riuscita a connettere le molteplici informazioni, provenienti da varie fonti, che erano a sua disposizione. Come dire, nel patrimonio informativo c'erano tutti i dati necessari, ma è semplicemente mancata l'utilizzazione di un buon algoritmo per mettere in correlazione tutti questi dati»[134].

Riflettendo su tale episodio, Giuseppe Italiano rileva, quindi, che:

«sicuramente, e non soltanto in questa circostanza, tecniche di IA sono e possono essere impiegate con successo nell'analisi delle informazioni disponibili, delle transazioni, dei file di log, del traffico sulla rete, e di tutte le "impronte" che ogni individuo lascia in rete e nei sistemi digitali, allo scopo di identificare possibili anomalie e attività sospette, o semplicemente per comporre in una visione coerente le informazioni provenienti da sorgenti multiple ed eterogenee, ed estrarne conoscenza, in modo tale da prendere in

[132] Il documento completo di presentazione del 2019 OSCE Annual Police Experts Meeting: Artificial Intelligence and Law Enforcement: An Ally or an Adversary?, 23-24 September, Wien, può essere letto su questa rivista (Redazione, Artificial Intelligence and Law Enforcement: an Ally or an Adversary?, 23 settembre 2019).

[133] In argomento, v. anche il documentato studio di A. G. Ferguson, The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement, New York University Press, 2017, pp. 3 ss.

[134] G. F. Italiano, Intelligenza artificiale: passato, presente, futuro, cit., p. 222.

maniera automatica decisioni oppure fornire il supporto a decisori umani, che devono essere in grado di reagire sempre più velocemente agli stimoli esterni».

Di recente l'INTERPOL e l'UNICRI (United Nations Interregional Crime and Research Institute) hanno pubblicato un Report in materia di AI, Robots e Law Enforcement, avente ad oggetto l'analisi dei possibili rischi e delle opportunità legati all'utilizzo dell'Intelligenza Artificiale nella lotta alla criminalità. Il Report raccoglie le conclusioni di una conferenza tenutasi a Singapore nel Luglio 2018, in cui Interpol e Unicri hanno riunito rappresentanti delle law enforcement agencies dei vari Paesi partecipanti, rappresentanti del mondo industriale e del mondo accademico al fine di analizzare lo stato dell'arte dell'utilizzo delle nuove tecnologie nei diversi ordinamenti e di delinearne i possibili sviluppi.

Partendo dalla considerazione per cui "crime has gone high tech", il Report sottolinea come nuove forme di criminalità connesse all'utilizzo distorto delle AI sono già da tempo particolarmente diffuse: basti pensare alle varie tipologie di cyber-attack, alla diffusione di fake news con l'obiettivo di creare o inasprire l'instabilità politica, agli strumenti di modifica dei documenti (face swapping e spoofing tools) volti a comprometterne la validità e via dicendo.
Se, dunque, le attività criminose sono sempre più technology-oriented, ne consegue che le autorità di controllo devono essere in grado di affrontare queste nuove sfide e, soprattutto, di utilizzare a proprio vantaggio gli innovativi strumenti tecnologici per prevenire e reprimere la criminalità. Due sono, difatti, gli aspetti evidenziati nel Report che determinano la necessità di introdurre strumenti di intelligenza artificiale nelle attività di law enforcement.
Da un lato, viene in rilievo una sorta di "asimmetria tecnologica" in cui le autorità di polizia e di controllo verrebbero a trovarsi rispetto ai gruppi criminali. Il Report evidenzia, in particolare, un diffuso "expertise gap" che deve necessariamente essere colmato per evitare che le autorità "restino indietro".
Quanto all'utilizzo delle AI per l'attività di law enforcement, il Report osserva come tale attività sia "information-based", fondandosi in gran parte sulla raccolta e l'analisi di numerose informazioni. Il ricorso alle nuove tecnologie può divenire allora essenziale per garantire una sistematizzazione e un'analisi di tali dati più efficiente. Tale analisi potrebbe essere anche di tipo predittivo, ove si faccia ricorso a tecniche di machine

learning, individuando, ad esempio, sulla base dei dati raccolti, i trend illeciti più frequenti e ottimizzando le risorse in base ai relativi outcome (si parla, a tal proposito, di "predictive policing").

Il Report contiene anche l'esame di alcune delle possibili applicazioni pratiche delle nuove tecnologie per l'attività di law enforcement. Tra i vari ambiti individuati vale la pena ricordare: strumenti di autopsia avanzati che identificano le cause di decesso; strumenti predittivi di individuazione dei trend illeciti più frequenti; software di computer vision per i furti d'auto e altre forme di riconoscimento facciale; strumenti automatizzati che identifichino e sanzionino in via autonoma i truffatori online; strumenti di machine learning per l'analisi delle conversazioni telefoniche; sistemi di sorveglianza basati sull'utilizzo dei droni.

Il Report, infine, si sofferma su alcune delle questioni problematiche di carattere etico legate all'utilizzo delle nuove tecnologie nella law enforcement, in particolar modo per quel che attiene alla tutela della privacy dei cittadini. In tale prospettiva, si sostiene che l'applicazione di strumenti di intelligenza artificiale alla law enforcement dovrebbe avvenire in maniera tale da garantire fairness, accountability, transparency ed explainability delle relative attività[135].

Le AI, dunque, dovrebbero essere applicate in questo settore tenendo sempre in considerazione alcuni principi di diritto fondamentali, quali quelli del giusto procedimento, della presunzione di innocenza, della libertà di espressione e della non discriminazione. L'attività di enforcement deve, inoltre, essere accountable nei confronti della comunità e in tale direzione si pongono anche le caratteristiche della trasparenza e dell'explainability, tutte volte a far sì che i processi decisionali che sono alla base dello svolgimento delle varie attività, nonché delle misure eventualmente adottate, siano conoscibili e comprensibili all'esterno[136].

3. La polizia predittiva: ratio ed impiego

La polizia predittiva è un sistema di analisi che tende alla creazione di modelli predittivi sulla verificazione di crimini. Al di là del profilo tecnologico, il punto di partenza della polizia predittiva è data da una serie di teorie criminologiche.

[135] https://www.irpa.eu/ai-law-enforcement/
[136] http://fra.europa.eu/it/news/2023/ethical-and-legal-aspects-ai-law-enforcement

In breve, si muove dall'idea che criminali e vittime seguono modelli di vita comuni, influenzati, a loro volta, da fattori geografici e temporali. Si assume, inoltre, che il criminale che si muove all'interno di tale contesto, sia un agente "razionale"; precisamente, la decisione di commettere un reato presuppone la valutazione di una serie di elementi, quali l'area, l'idoneità dell'obiettivo e il rischio di essere scoperti.
Il software predittivo, dunque, non fa che replicare il lavoro che, su queste basi, veniva precedentemente svolto da operatori umani, con la differenza che riesce ad analizzare velocemente enormi quantità di dati, con risultati sempre più accurati. Non sorprende, pertanto, che l'impiego di tali software in ambito investigativo stia riscuotendo crescente seguito in tutto il mondo.
I software predittivi seguono differenti metodologie. Il metodo più comune è quello "place based". Come suggerisce la locuzione, tale metodologia si incentra sull'attenzione a particolari luoghi ad alto rischio o tasso di criminalità. Si pensi a PredPol, portato della collaborazione tra il dipartimento di polizia di Los Angeles e l'esperto Jeff Brantigham dell'Ucla, che elabora le previsioni sulla base di dati storici della criminalità. Altri software ancora prendono in considerazione fattori come la luminosità, la vicinanza ai negozi di liquori, le condizioni metereologiche.
Diverso è l'approccio su cui si regge Keycrime, in dotazione della questura di Milano. La caratteristica principale del software è il crime linking che consente di segnalare le serie criminali effettuate dalle stesse persone, preconizzando il verificarsi di ulteriori reati. Più precisamente, partendo dai dati investigativi raccolti sulla scena del crimine dai testimoni e dalle registrazioni delle telecamere di videosorveglianza, Keycrime offre un'indicazione probabilistica sull'evoluzione di una serie criminale.

I software di polizia predittiva – siano essi assistiti o meno da sistemi di IA[137] – possono dividersi fondamentalmente in due categorie:
– quelli che, ispirandosi alle acquisizioni della criminologia ambientale, individuano le c.d. "zone calde" (hotspots), vale a dire i luoghi che costituiscono il possibile scenario dell'eventuale futura commissione di determinati reati;
– quelli che, ispirandosi invece all'idea del crime linking, seguono le serialità criminali di determinati soggetti (individuati o ancora da individuare), per prevedere dove e quando costoro commetteranno il prossimo reato.

[137] Non sempre risulta chiaro se, e in quale misura, i software di cui parleremo nelle pagine seguenti si basino su sistemi di IA. Ciò dipende anche dal fatto che alcuni di questi software sono di proprietà privata e sono coperti da segreto industriale, sicché i dettagli sul loro funzionamento non sono resi pubblici.

Va subito detto che, almeno per ora, sia gli uni che gli altri sistemi possono fornire adeguate previsioni solo in relazione a limitate, determinate categorie di reati (ad esempio, reati attinenti alla criminalità da strada, come rapine e spaccio di stupefacenti), e non in via generalizzata per tutti i reati.

3.1. Sistemi di individuazione degli hotspots.

Rientra nel primo tipo di sistemi il Risk Terrain Modeling (RTM): un algoritmo che, rielaborando quantità enormi di dati inerenti i fattori ambientali e spaziali favorenti la criminalità, sembrerebbe consentire la predizione della commissione di reati di spaccio di sostanze stupefacenti in determinate aree urbane[138]. I ricercatori hanno elaborato questo sistema sottoponendo all'algoritmo RTM dati inerenti i fattori ambientali e spaziali più frequentemente connessi alla commissione dei reati suddetti: presenza di luminarie stradali scarse o non funzionanti, vicinanza di locali notturni, di fermate di mezzi pubblici, di stazioni ferroviarie, di snodi di strade ad alta percorribilità, di bancomat, di compro-oro, di parcheggi scambiatori, infine, di scuole. Ciò ha consentito di elaborare una vera e propria "mappatura" di alcune grandi aree metropolitane al fine di individuare le "zone calde" dove più elevato risulta il rischio di spaccio di sostanze stupefacenti, con conseguenti benefici in termini di programmazione e attuazione di interventi di prevenzione della delinquenza connessa allo spaccio.

Parimenti finalizzato all'individuazione degli hotspots ma in relazione ad un numero più elevato di reati (non solo quelli di spaccio) è anche un software, già in uso da alcuni anni negli Stati Uniti e nel Regno Unito, originariamente elaborato da alcuni ricercatori dell'UCLA (Università della California di Los Angeles) in collaborazione con la locale polizia, e oggi venduto, parrebbe con grande successo commerciale, da un'azienda privata americana col marchio PredPol, il cui sito pubblicizza tale dispositivo con le seguenti parole:

«utilizzando solo tre tipologie di dati – tipo di reato, data/ora del reato e luogo del reato – per fare previsioni, la tecnologia PredPol ha aiutato le forze dell'ordine a ridurre drasticamente il tasso di criminalità in giurisdizioni di tutti i tipi e di tutte le dimensioni, negli Stati Uniti e all'estero. PredPol può vantare una comprovata sperimentazione: il

[138] In argomento, v. J.M.Caplan, L.W. Kennedy, J.D. Barnum, E.L Piza, Crime in Context: Utilizing Risk Terrain Modeling and Conjunctive Analysis to Explore the Dynamics of Criminogenic Behavior Setting, in Journal of Contemporary Criminal Justice, 33(2), 2017, pp. 133 ss.; J.M.Caplan, L.W. Kennedy, Risk Terrain Modeling: Crime Prediction and Risk Reduction, Univ. of California Press, 2016; L.W. Kennedy, J.M.Caplan, E.L Piza, Risk Clusters, Hotspots and Spatial Intelligence: Risk Terrain Modeling as an Algorithm for Police Resource Allocation Strategies, in Journal of Quantitative Criminology, 2010, pp. 339 ss.

Dipartimento della polizia di Los Angeles ha registrato un calo del 20% dei reati previsti di anno in anno e una divisione della locale polizia ha potuto sperimentare per la prima volta, un'intera giornata senza ricevere denunce di reati. Il Dipartimento dello Sceriffo della contea di Jefferson ha registrato una riduzione del 24% nelle rapine e una riduzione del 13% nei furti con scasso. A Plainfield, New Jersey, da quando si usa PredPol si è avuta una riduzione del 54% delle rapine e una riduzione del 69% dei furti di auto»[139].
Sembrerebbe ispirarsi ad una analoga logica predittiva anche un dispositivo in uso presso la polizia italiana, di cui si parlerà nello specifico in seguito: si tratta del sistema informatico X-LAW, originariamente predisposto dalla Questura di Napoli, che parrebbe aver già ottenuto ottimi risultati sul territorio italiano nel campo della prevenzione di talune tipologie di reati. Stando alle notizie riferite[140], il software X-LAW si basa su un algoritmo capace di rielaborare una mole enorme di dati estrapolati dalle denunce inoltrate alla Polizia di Stato. Tale rielaborazione consente di far emergere fattori ricorrenti o fattori coincidenti, come ad esempio la ripetuta commissione di rapine negli stessi luoghi, da parte di persone con lo stesso tipo di casco o di moto, e con analoghe modalità. Ciò consente di tracciare una mappa del territorio dove vengono evidenziate le zone a più alto rischio fino a raggiungere il livello massimo in determinati orari, così consentendo – nelle zone e negli orari 'caldi' – la predisposizione delle forze dell'ordine per impedire la commissione di tali reati e per cogliere in flagranza i potenziali autori degli stessi.

3.2. Sistemi di crime linking.

Si rifà, invece, all'idea del crime linking, seguendo le serialità criminali di determinati soggetti (individuati o ancora da individuare), per prevedere dove e quando essi commetteranno il prossimo reato, il software Keycrime, originariamente elaborato presso la Questura di Milano, e poi divenuto di proprietà di un'azienda privata[141]. Altri software parimenti ispirati all'idea del crime linking, e quindi all'individuazione delle persone, più

[139] https://www.predpol.com/, visitato il 9 agosto 2019.
[140] Notizie riferite da M. Iaselli, X-LAW: la polizia predittiva è realtà, in Altalex.com, 28 novembre 2018.
[141] In argomento, v. A.D. Signorelli, Il software italiano che ha cambiato il mondo della polizia predittiva, in Wired.it, 18 maggio 2019; C. Parodi, V. Sellaroli, Sistema penale e intelligenza artificiale: molte speranze e qualche equivoco, in Diritto penale contemporaneo, 2019, fasc. 6, p. 56. Per una descrizione di Keycrime, fornita dal suo stesso ideatore, Mario Venturi, v.
Id., La chiave del crimine, in Profiling, 5, 4, 2014.

che delle zone calde, sono stati elaborati, e sono in uso, in Germania (Precobs)[142], in Inghilterra (Hart - Harm Assessment Risk Tool)[143], e negli Stati Uniti[144].

Questi software si basano sull'idea di fondo che alcune forme di criminalità si manifesterebbero in un arco temporale e in una zona geografica molto circoscritti (c.d. near repeat crimes, o reati a ripetizione ravvicinata): ad esempio, la commissione di una rapina sembrerebbe essere associata ad un elevato rischio di commissione di una nuova rapina, da parte degli stessi autori e in una zona geografica assai prossima al luogo del primo delitto, entro le successive 48 ore e, sia pur con un tasso di rischio decrescente, fino a tutto il mese successivo. Attraverso la raccolta e l'incrocio di una gran mole di dati, provenienti da varie fonti (ad esempio, immagini riprese da una telecamera o informazioni relative a precedenti analoghi reati), questi software cercano, infatti, di "profilare" il possibile autore della serie criminale e prevederne la prossima mossa.

Peraltro, i risultati forniti da questi software in alcuni casi potrebbero essere usati non solo a fini predittivi, ma anche per ricostruire la carriera criminale del soggetto profilato, vale a dire per avere una traccia di indagine da seguire per imputargli non solo l'ultimo reato commesso (in occasione del quale egli è stato individuato), ma anche i precedenti reati costituenti la serie criminale ricostruita grazie all'archiviazione e all'elaborazione dei dati.

4. Il valore delle informazioni raccolte tramite i tools

Abbiamo visto fin qui tutte le principali caratteristiche di questa metodologia di predire possibili atti violenti, o nel dare delle valide informazioni desunte dalle analisi video della videosorveglianza urbana, informazioni queste piuttosto utili alla polizia locale per poter prevenire, ad esempio, possibili raduni rave, attività queste tipiche della sicurezza urbana, anche se con possibili risvolti di ordine e sicurezza pubblica, ma sempre del tutto illegali. Pur tuttavia, questa fattiva attività analitica costituita da metodi, tecnologie e algoritmi, anche quando potenzialmente utili alle forze dell'ordine, genera un ulteriore

[142] Su Precobs, v. l'accurata voce di Wikipedia https://en.wikipedia.org/wiki/Precobs.

[143] Sul software HART, v. M. Oswald, J. Grace, S. Urwin, G. Barnes, Algorithmic risk assessment policing models: lessons from the Durham HART model and "Experimental" proportionality, in Information & Communications Technology Law, 2018, pp. 223 ss.; nella dottrina italiana, v. M. Gialuz, Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa, in Diritto penale contemporaneo, 29 maggio 2019, pp. 10 ss. Il software HART è stato sottoposto a studi di validazione da parte di alcuni ricercatori della Cambridge University: cfr. il presente indirizzo web.

[144] Per un progetto pilota avviato nella città di Chicago, v. J. Predictions put into practice: a quasi-experimental evaluation of Chicago's predictive policing pilot, in .

interrogativo: quale valore avranno in sede giudiziaria le informazioni raccolte tramite i "tools" di predictive policing?
Peraltro, questi strumenti di polizia predittiva, ma in linea di principio tutti i sistemi di intelligenza artificiale legati al settore, ad oggi non sono stati ancora oggetto di una specifica regolamentazione, sia nell'uso che nella validazione, delegando piuttosto condizioni e modalità d'impiego, nonché la valutazione e il probabile valore giuridico dei risultati assunti, alle abituali procedure operative quando non addirittura alle analisi del singolo operatore di polizia.

Invero, un primo approccio regolamentorio c'è stato il 3 dicembre 2018 mediante l'adozione della Carta Etica Europea "sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi adottata dalla CEPEJ" (commissione europea per l'efficacia della giustizia).

Una Carta, tipico documento di soft law, che mette in luce, fissandoli, cinque principi cardinali che dovranno rappresentare una cornice imprescindibile sulla futura disciplina della materia: il rispetto dei diritti fondamentali, il principio di non discriminazione, il principio di qualità e sicurezza dei dati e delle decisioni giudiziarie, il principio di trasparenza e imparzialità e la possibilità di controllo da parte dell'utente.
Successivamente, poi, un ulteriore passaggio sul tema è datato 21 aprile 2021, quando la Commissione Europea presentò 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".
La proposta non vuole ordinare in se l'intelligenza artificiale, in quanto tale, quanto piuttosto l'utilizzo nell'ambito UE di sistemi logici che contengono tecnologie fortemente impattanti la sfera privacy e la protezione dei dati personali per mezzo di software e di sistemi informatici intelligenti che dovranno, invece, far riferimento al data protection reform package, costituito nel Regolamento UE 2016/679 (GDPR) e nella Direttiva UE 2016/680 (direttiva considerata lex specialis).

Appare evidente come l'impiego di software predittivi nelle attività di polizia di prevenzione, ma particolarmente in quella giudiziaria laddove utilizzati per individuare una responsabilità penale, non potrà certo costituire l'unico elemento basilare su cui poggerà poi il dispositivo decisionale finale in sede di giudizio.

E questo è ancora più vero quando in gioco c'è la libertà personale dell'imputato, diritto imprescindibile protetto e garantito dalle Carte europee sui diritti fondamentali dell'uomo (artt. 5, 6, 8 CEDU – art. 6 CDFUE).
Del resto, già all'interno dell'art. 192, comma 2 CCP è stabilito che "l'esistenza di un fatto non può essere desunta da indizi a meno che questi siano gravi, precisi e concordanti"; dunque, è del tutto pacifico come qualsiasi output del software dovrà essere considerato come uno dei tanti elementi che potranno formare un giudizio di responsabilità penale.
Non va dimenticato, inoltre, che qualsiasi risultato scaturito dall'analisi investigativa predittiva dovrà essere sempre assunto nel pieno rispetto dei vincoli imposti dal codice di procedura penale (artt. 191 e 220), con prove ben delineate all'interno della cornice imposta dalla locuzione "al di là di ogni ragionevole dubbio" (art.533), una prerogativa di ragionevolezza che difficilmente sarà rinvenibile all'interno delle pieghe di un qualsiasi algoritmo, ancor meno computabile con metodologia statistica.

Si badi bene, però, che la stretta sinergia che si vorrebbe creare tra smart city e predictive policing, la cui genesi parte da una premessa comune, ovverosia, dall'esperimento di rendere una città lo strumento a disposizione di complesse "esperienze" informatico/matematiche legate al calcolo statistico, argomenti dogmatici molto spinosi.
Infine, le verosimili conseguenze fin qui analizzate come possibile incidenza negativa sui principi costituzionali del nostro diritto penale, meritano certamente una profonda e ragionata analisi del binomio applicativo tra città intelligente e modelli di polizia predittiva, perché anche a fronte di possibili e innegabili vantaggi, in termini di efficacia sociale del progetto e di prevenzione/repressione dei reati, ebbene, ci si dovrà misurare con un tema molto spigoloso: quello del pregiudizio algoritmico, pregiudiziale già centrale nella critica dei sistemi di sicurezza della video facial recognition.

Con l'ultima risoluzione che verrà analizzata, ossia quella del 6 ottobre 2021 - l'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale – si conclude, temporaneamente, il percorso che il Parlamento europeo sta svolgendo per creare un "ecosistema" trasparente nel campo dell'IA.
Il Parlamento europeo affronta il tema dell'utilizzo dell'intelligenza artificiale nel diritto penale. Il Parlamento dichiara positivo il contributo di determinati tipi di applicazioni di

IA al lavoro delle autorità di contrasto e giudiziarie in tutta l'Unione; ciò nondimeno, constata pure che lo sviluppo dell'IA può implicare rischi enormi per i diritti fondamentali e le democrazie basate sullo Stato di diritto. Elenca, quindi, una serie di requisiti che questi sistemi debbono possedere affinché possano essere lecitamente utilizzati.

In generale, il Parlamento afferma che l'utilizzo dell'IA, che non può in nessun caso "*nuocere alla integrità fisica degli esseri umani, né attribuire diritti o imporre obblighi giuridici agli individui*",[145] deve essere vietato se incompatibile con i diritti fondamentali; più precisamente, ritiene fondamentale che si appresti uno specifico quadro giuridico chiaro e preciso, che sia in grado di disciplinare le condizioni, le modalità e le conseguenze dell'utilizzo degli strumenti di IA nei settori dell'attività di contrasto e giudiziario,[146] nonché i diritti delle persone interessate e le procedure - efficaci e facilmente accessibili - di reclamo e di ricorso "*anche per via giudiziaria*"[147].

Le soluzioni di IA, si legge, devono rispettare appieno i princìpi di "*dignità umana, 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*."[148]

In questa prospettiva, la Commissione ritiene che la diffusione, lo sviluppo e l'utilizzo di questi strumenti, debbano essere assoggettati ad una previa "*valutazione dei rischi e a una rigorosa verifica* ex ante *della necessità e della proporzionalità*".[149] Da ciò, deriva anche la fiducia che i cittadini riporranno nell'utilizzo dell'IA in questi settori, fiducia consolidata dal fatto che "*che qualsiasi strumento di IA sviluppato o utilizzato dalle autorità di contrasto o giudiziarie dovrebbe come minimo essere sicuro, solido e adatto allo scopo previsto, rispettare i principi di equità, minimizzazione dei dati, responsabilità, trasparenza, non discriminazione e spiegabilità, e il suo sviluppo, la sua diffusione e il suo utilizzo dovrebbero essere soggetti a una valutazione dei rischi e a una*

[145] Risoluzione del Parlamento europeo del 6 ottobre 2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale, punto12. (2020/2016(INI))

[146] POSTERARO N., *Una risoluzione del Parlamento europeo sull'uso dell'intelligenza artificiale nel diritto penale*, in Istituto di Ricerche sulla pubblica amministrazione, 23 novembre 2021, https://www.irpa.eu/una-risoluzione-del-parlamento-europeo-sulluso-dellintelligenza-artificiale-nel-diritto-penale/.

[147] Risoluzione del Parlamento europeo del 6 ottobre 2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale, punto14. (2020/2016(INI))

[148] Risoluzione del Parlamento europeo del 6 ottobre 2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale, punto2. (2020/2016(INI))

[149] Risoluzione del Parlamento europeo del 6 ottobre 2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale, punto20. (2020/2016(INI))

rigorosa verifica della necessità e della proporzionalità, e le relative salvaguardie dovrebbero essere proporzionate ai rischi individuati; sottolinea che la fiducia tra i cittadini nell'utilizzo dell'IA sviluppata, diffusa e utilizzata nell'Unione dipende dal pieno rispetto di tali criteri".[150]

Sottolinea, poi, che gli aspetti legati alla sicurezza e alla protezione dei sistemi di IA utilizzati devono essere valutati con attenzione per *"prevenire le conseguenze potenzialmente catastrofiche di attacchi dolosi".*[151]

In quest'ottica, sono invitate le autorità di contrasto e giudiziarie, ad adoperare esclusivamente applicazioni di Intelligenza Artificiale che rispettino il principio della tutela della vita privata, e della protezione dei dati sin dalla progettazione e, a questo ultimo proposito, nel ricordare a quali condizioni il trattamento dei dati possa dirsi lecito e corretto, afferma che le norme sulla protezione dei dati definite dall'UE per l'attività di contrasto costituiscono la base per qualunque futura regolamentazione dell'IA per l'impiego di attività di contrasto e nel settore giudiziario.[152]

In più parti, viene evidenziata l'importanza di evitare che l'uso di applicazioni basate sull'IA, comportando distorsioni "intrinseche", amplifichi le discriminazioni esistenti - in particolare, nei confronti delle persone che appartengono a determinate minoranze tecniche o comunità razziali e diventi fattore di disuguaglianza, frattura sociale o esclusione - .[153]

Pertanto, sottolineato che molte tecnologie di identificazione, fondate su algoritmi, compiono un numero smisurato di errori di identificazione e di classificazione, il Parlamento europeo afferma come non sia opportuno affidare troppa fiducia nella natura apparentemente oggettiva e scientifica degli strumenti di IA, ed evidenzia quanto sia importante:

- che i *team* che progettano, sviluppano, testano, eseguono la manutenzione, diffondono e forniscono tali sistemi di IA per le attività di contrasto e giudiziarie siano interdisciplinari e rispecchino, ove possibile, la diversità della società in generale;

[150]Risoluzione del Parlamento europeo del 6 ottobre 2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale, punto4. (2020/2016(INI))

[151] Risoluzione del Parlamento europeo del 6 ottobre 2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale, punto11. (2020/2016(INI))

[152] Risoluzione del Parlamento europeo del 6 ottobre 2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale, punto1. (2020/2016(INI))

[153]POSTERARO N., *Una risoluzione del Parlamento europeo sull'uso dell'intelligenza artificiale nel diritto penale*, in Istituto di Ricerche sulla pubblica amministrazione, 23 novembre 2021, https://www.irpa.eu/una-risoluzione-del-parlamento-europeo-sulluso-dellintelligenza-artificiale-nel-diritto-penale/.

- che sia assicurata una adeguata formazione dei decisori, affinché gli stessi siano consci delle (e sappiano gestire le) distorsioni che gli strumenti utilizzati possono potenzialmente generare, quando utilizzati;
- che gli Stati membri promuovano politiche contro la discriminazione ed elaborino piani d'azione nazionali contro il razzismo nel settore delle attività di polizia e nel sistema giudiziario, posto che "*l'inclusione nelle serie di dati in materia di addestramento dei sistemi di IA di casi di razzismo da parte delle forze di polizia nell'adempimento dei propri doveri comporterà inevitabilmente distorsioni di natura razzista nei risultati, nei punteggi e nelle raccomandazioni basati sull'IA.*"[154]

Sempre con lo scopo di eludere le discriminazioni, il Parlamento si oppone, in seguito, all'utilizzo dell'IA da parte delle autorità di contrasto, per fare previsioni sui comportamenti degli individui o di gruppi, sulla base di dati storici e condotte precedenti, dell'appartenenza a un gruppo, l'ubicazione o qualunque altra caratteristica al fine di identificare le persone che potrebbero commettere un reato.[155]

In vari punti, il Parlamento, sul modello già evidenziato in altre occasioni, sottolinea il potere della supervisione umana, sostenendo che *"sia necessario un controllo umano specifico a monte dell'utilizzo di determinate applicazioni critiche"*[156]*;* afferma, poi, che il provvedimento, nei contesti giudiziari o di contrasto, qualora produca effetti giuridici o analoghi, debba essere sempre assunto da un essere umano. Chiede, in particolare, che siano conservate la competenza esclusiva dei giudici, e l'adozione di decisioni caso per caso, ed inoltre, che sia inserito un divieto sull'uso dell'IA e delle relative tecnologie per l'emanazione delle decisioni giudiziarie.

Viene espressamente inserita la necessità di un regime chiaro ed equo per attribuire, ad una persona fisica o giuridica, la responsabilità giuridica delle potenziali conseguenze negative prodotte da tali tecnologie digitali avanzate.[157]

---

[154]Risoluzione del Parlamento europeo del 6 ottobre 2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale, punto23. (2020/2016(INI))

[155]POSTERARO N., *Una risoluzione del Parlamento europeo sull'uso dell'intelligenza artificiale nel diritto penale*, in Istituto di Ricerche sulla pubblica amministrazione, 23 novembre 2021, https://www.irpa.eu/una-risoluzione-del-parlamento-europeo-sulluso-dellintelligenza-artificiale-nel-diritto-penale/.

[156]Risoluzione del Parlamento europeo del 6 ottobre 2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale, punto11. (2020/2016(INI))

[157] Risoluzione del Parlamento europeo del 6 ottobre 2021 sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale, punto13. (2020/2016(INI))

La risoluzione precisa inoltre che, per garantire il rispetto dei diritti fondamentali dell'uomo, gli algoritmi debbano essere tracciabili, trasparenti e sufficientemente documentati; il rischio, altrimenti, è che, tra le altre cose, essi abbiano un impatto avverso sul diritto alla difesa degli indagati, posto che li metterebbero in difficoltà nell'ottenere informazioni significative sul loro funzionamento e nel confutare i risultati in tribunale.[158]

Il Parlamento chiede un *audit* indispensabile periodico, di tutti i sistemi di IA utilizzati dalle autorità di contrasto e dal potere giudiziario, nel caso in cui essi possano influire in maniera rilevante sulla vita delle persone, da parte di un'autorità indipendente, per testare e valutare i sistemi di algoritmi, il contesto, le finalità, la precisione, le prestazioni e la portata una volta che questi siano operativi, al fine di individuare, indagare, diagnosticare e rettificare eventuali effetti indesiderati e negativi e assicurare che i sistemi di IA funzionino come previsto.[159]

Gli Stati membri sono tenuti a fornire informazioni complete sugli strumenti utilizzati dalle proprie autorità di contrasto e giudiziarie, sulle tipologie di strumenti utilizzati, sulle finalità per cui sono utilizzati, sui tipi di reati cui si applicano e i nomi delle società o organizzazioni che hanno sviluppato tali strumenti*;* la Commissione deve quindi aggiornare periodicamente le informazioni in un'unica sede, pubblicando le informazioni sull'utilizzo delle intelligenze artificiali da parte delle agenzie dell'Unione incaricate delle attività di contrasto e giudiziarie.

Ancora, sottolinea che l'utilizzo e la raccolta di dati biometrici per finalità di identificazione a distanza, ad esempio attraverso il riconoscimento facciale in luoghi pubblici, possono presentare rischi specifici per i diritti fondamentali. *Su* queste basi, invita la Commissione, tramite strumenti legislativi e non legislativi, a introdurre il divieto di trattamento dei dati biometrici, comprese le immagini facciali, per finalità di applicazione della legge, tale da determinare sorveglianza di massa negli spazi accessibili al pubblico, e a interrompere il finanziamento della ricerca o diffusione della biometrica o di programmi che potrebbero portare alla sorveglianza di massa indiscriminata nei luoghi pubblici; sottolinea, in questo contesto, che andrebbe prestata particolare attenzione e dovrebbe essere applicato un quadro rigoroso all'utilizzo dei droni nelle operazioni di polizia.

[158]POSTERARO N., *Una risoluzione del Parlamento europeo sull'uso dell'intelligenza artificiale nel diritto penale*, in Istituto di Ricerche sulla pubblica amministrazione, 23 novembre 2021, https://www.irpa.eu/una-risoluzione-del-parlamento-europeo-sulluso-dellintelligenza-artificiale-nel-diritto-penale/.
[159] *Ibidem*.

In conclusione, il Parlamento chiede all'Agenzia per i diritti fondamentali dell'UE, in collaborazione con il Comitato europeo per la protezione dei dati e il Garante europeo della protezione dei dati, di elaborare raccomandazioni, buone pratiche ed orientamenti completi, per fari sì che siano precisati ulteriormente i criteri e le condizioni per la diffusione, lo sviluppo e l'uso di applicazioni e soluzioni di IA da parte delle autorità di contrasto e giudiziarie. In concomitanza, riserva di condurre uno studio sull'attuazione della direttiva sulla protezione dei dati - Direttiva UE 2016/680 del Parlamento europeo e del Consiglio, del 27 aprile 2016 -, al fine di statuire in che modo la protezione dei dati personali sia stata garantita nelle attività di trattamento da parte delle autorità di contrasto e giudiziarie, in particolare nelle fasi di sviluppo e diffusione delle nuove tecnologie**,** ed esorta la Commissione a valutare se sia doveroso adottare una specifica azione legislativa volta a definire ulteriormente i suddetti criteri e le suddette condizioni**.**

In tal modo, la risoluzione sottolinea come l'IA non possa essere ritenuta fine a sé stessa, ma debba, eventualmente, essere trattata come uno **s**trumento posto al servizio dei singoli, capace di ampliare il benessere degli esseri umani, le capacità umane e la sicurezza. La tecnologia dell'intelligenza artificiale, deve evolversi mettendo al centro le persone: *"solo così potrà essere degna della fiducia dei cittadini e potrà porsi davvero al loro servizio".*[160]

5. Le criticità

La profilazione tramite algoritmi, oltre a compromettere la protezione dei dati personali della collettività, rischia di tradursi in un ulteriore fattore di discriminazione. Si guardi, in proposito, all'esperienza americana con il software PredPol. Alcuni studi hanno evidenziato come il software risenta dei pregiudizi delle forze di polizia. Poiché negli Stati Uniti vengono arrestati per lo più afroamericani e soggetti appartenenti a minoranze etniche e poiché le persone afroamericane sono denunciate più spesso dei bianchi, i quartieri a maggioranza nera tendono ad essere individuati come hotspot criminali in misura maggiore di quanto provino i dati reali. Come rimediare? Oltre ad arginare i bias dell'algoritmo by design, è chiaro che spetterà all'operatore valutare l'output del software e correggerne le storture. L'algoritmo, infatti, per quanto sia un valido ausilio in sede

[160]POSTERARO N., *Una risoluzione del Parlamento europeo sull'uso dell'intelligenza artificiale nel diritto penale*, in Istituto di Ricerche sulla pubblica amministrazione, 23 novembre 2021, https://www.irpa.eu/una-risoluzione-del-parlamento-europeo-sulluso-dellintelligenza-artificiale-nel-diritto-penale/.

investigativa, non può – almeno finora – sostituire tout court l'attività dell'uomo, cui spetta, in ultima istanza, decidere.

Sono stati avanzati diversi dubbi per quanto riguarda l'esattezza dei risultati prodotti dai sistemi di polizia predittiva, del loro impatto nella lotta al crimine e, soprattutto, con riferimento al rischio di discriminazione. In particolare, si evidenzia il pericolo che detti sistemi prevedano il comportamento delle vittime e della polizia piuttosto che dei criminali. Alcune associazioni per i diritti umani considerano pertanto la polizia predittiva come un uso inammissibile dell'intelligenza artificiale[161].

In un'analisi del sistema PredPol realizzata dal Human Rights Data Analysis Group nel 2016, si evidenziava che l'obiettivo di ogni algoritmo di apprendimento automatico è di sviluppare i modelli a partire dai dati che vengono inseriti nel software. Pertanto, se nutrito con i dati della polizia, l'algoritmo apprende sulla base di questi: ciò significa che il sistema non elabora modelli relativi agli illeciti ma piuttosto modelli relativi a come la polizia registra il crimine. Poiché la polizia registra i reati in modo non uniforme in una città, i modelli di criminalità registrata possono differire sostanzialmente dai modelli reali di criminalità. Di conseguenza, il modello che l'algoritmo probabilmente apprenderà è che la maggior parte dei crimini si verifica nei luoghi sovra-rappresentati, che potrebbero non essere necessariamente i luoghi con il più alto livello di criminalità.

Utilizzando i modelli che ha appreso, l'algoritmo può quindi fare previsioni sulla futura distribuzione del crimine e la polizia verrà inviata nei luoghi con il più alto tasso di criminalità previsto. Poiché l'algoritmo ha appreso che i luoghi sovra-rappresentati hanno il maggior numero di reati, più la polizia verrà inviata in quelle aree, più saranno trovati crimini in quei luoghi. Gli illeciti così osservati vengono quindi reinseriti nell'algoritmo per ripetere questo processo nei giorni successivi creando un circolo vizioso in cui la polizia viene inviata in determinati luoghi perché l'algoritmo considera questi luoghi come aventi un più alto tasso di criminalità sulla base del fatto che quelli erano i luoghi in cui sono stati inviati.

Allo stesso modo, molti sono dell'avviso che dai dati basati sugli arresti derivino pregiudizi sui risultati degli algoritmi predittivi perché la polizia arresta più persone nei quartieri dove sono maggiormente presenti le minoranze etniche, portando gli algoritmi a dirigere sempre più le attività di polizia in quelle aree. Ne risulta che gli strumenti

[161] EDRi, "Can the EU make AI "trustworthy"? No – but they can make it just", https://edri.org/our-work/can-the-eu-make-ai-trustworthy-no-but-they-can-make-it-just/

predittivi orientano in modo errato le pattuglie di polizia e che alcuni quartieri sono ingiustamente designati punti caldi della criminalità mentre altri sono sottovalutati.
In questo contesto è stato suggerito che l'apprendimento degli algoritmi sulla base delle denunce da parte delle vittime piuttosto che sui dati di arresto potesse ridurre le distorsioni dei sistemi predittivi. Questo perché, in teoria, le segnalazioni delle vittime dovrebbero essere meno distorte perché meno influenzate dai pregiudizi della polizia o dai circuiti di feedback.
Uno studio su un algoritmo predittivo elaborato da ricercatori dell'Università del Texas evidenzia, però, che le elaborazioni predittive basate sulle segnalazioni delle vittime porta ad altrettanti pregiudizi nei risultati. Confrontando le previsioni del sistema basato su dati delle denunce con i dati reali sulla criminalità per determinate zone, sono apparsi errori significativi. Così, in una zona in cui sono stati segnalati pochi crimini, lo strumento ha previsto solo circa il 20% dei punti caldi effettivi. Al contrario, in una zona in cui sono stati segnalati un numero elevato di crimini, lo strumento ha previsto il 20% in più di hot spot rispetto a quelli realmente esistenti.

Infatti, il primo problema che viene evidenziato con le denunce delle vittime è che i neri hanno maggiori probabilità di essere denunciati per un crimine rispetto ai bianchi. I bianchi, generalmente più ricchi, segnalano tendenzialmente con più probabilità una persona nera più povera, anziché il contrario. Lo studio evidenza come anche i neri hanno maggiori probabilità di denunciare altri neri. Di conseguenza, come per i dati sugli arresti, questo porta i quartieri con prevalenza di popolazione di colore a essere designati come punti caldi della criminalità più frequentemente di quanto dovrebbe essere.
Inoltre, è evidente come la segnalazione delle vittime sia collegata anche alla fiducia della comunità o alla sfiducia di essa nei confronti della polizia. Quindi, se sei in una comunità con un dipartimento di polizia storicamente corrotto o con pregiudizi notoriamente razzisti, ciò influenzerà in modo sostanziale i dati relativi alla criminalità. In questo caso, il sistema predittivo potrebbe sottostimare il livello di criminalità in una determinata area, aumentando quindi il sentimento di impunità e la criminalità stessa.
Sembra dunque che, allo stato attuale, il problema dei pregiudizi degli algoritmi di polizia predittiva non si possa risolvere. Questo a dimostrazione ancora una volta che non sono gli algoritmi a essere discriminatori, ma che essi riproducono le discriminazioni e i pregiudizi presenti nella nostra società. Più che lamentare le discriminazioni frutto degli

algoritmi, occorre quindi intervenire a monte per rimuovere i comportamenti discriminatori dalla lotta alla criminalità.

Questi risultati dimostrano, altresì, che l'utilizzo delle nuove tecnologie nei procedimenti penali deve rimanere uno strumento di assistenza dei diversi attori, ma non può sostituirsi al ruolo centrale delle persone. Infatti, i benefici dell'abbinamento tra sistemi di polizia e l'analisi umana sono in grado di limitare le criticità relative agli algoritmi e, così, rispondere al necessario equilibrio tra lotta al crimine e protezione dei diritti fondamentali nella giustizia penale.

Pertanto, lo sviluppo di strumenti di polizia predittiva non discriminatori passa innanzitutto da politiche volte a rimuovere ogni tipo di discriminazione nella raccolta delle statistiche criminali e nella condotta delle attività di polizia in generale così come dal ristabilire la fiducia dei cittadini verso le forze dell'ordine. Perché non è l'algoritmo a essere discriminatorio ma riproduce semplicemente i pregiudizi ancorati nella società; è dunque su questi che bisogna intervenire in primo luogo.

6. Un modello di polizia predittiva: XLAW

XLAW è un trovato tecnologico e metodologico ideato e implementato per sperimentare, per la prima volta in Italia, l'applicazione della Polizia Predittiva per la Sicurezza Urbana, ovvero un approccio innovativo ai problemi di insicurezza delle comunità che rivoluziona il metodo tradizionale di Prevenzione dell'illegalità diffusa, perchè si basa sulla possibilità di poter Prevedere, con l'impiego d'Intelligenza Artificiale, scippi, rapine, furti, borseggi, truffe e altri delitti di tipo cosiddetto "predatorio" che normalmente avvengono nelle nostre bellissime città.

Il trovato nel suo insieme consiste in un protocollo tecnico e metodologico configurato per generare e impiegare strategicamente allarmi Predittivi georeferenziati di possibili crimini, elaborati secondo un esclusivo modello previsionale di Machine Learning.

L'innovazione consiste nel fatto che rispetto a tutti i sistemi tradizionali di controllo (videosorveglianza ecc.) o allarme (antintrusione, barriere elettroniche ecc.) i quali hanno il limite di poter essere considerati solo dopo che i crimini sono accaduti, il trovato permette di Prevenirli vigilando dove è scientificamente previsto che accadano e non dove si pensi possano accadere o peggio, dove già accaduti, il noto limite del metodo di prevenzione tradizionale.

La lunga e diffusa sperimentazione della Polizia Predittiva per la Sicurezza Urbana secondo il trovato, condotta per dare un valido contributo alla conoscenza e al progresso nel contrasto all'illegalità diffusa, può essere considerata nel suo insieme una ricerca scientifica i cui esiti sono stati valutati e validati indipendentemente dalle più importanti strutture di sicurezza e da due università grazie alla quale, si è potuto stabilire che l'Intelligenza Artificiale, ormai sempre più al fianco delle attività umane, se implementata e impiegata opportunamente anche per l'attività di Sicurezza Urbana, attesa l'alta attendibilità e precisione delle analisi e quindi delle informazioni strategiche che riesce a fornire, può elevare l'attività di prevenzione e controllo e renderla maggiormente dinamica, precisa ed efficace evitando anche lo spreco di risorse e di energie.

«Il suo lungo impiego ha spostato il costrutto strategico dell'azione di controllo da una visione riparatoria del danno ad una visione probabilistica del rischio, quindi da una logica di rincorsa dei problemi e degli effetti che essi generano tipica della permanente emergenza, ad una che lavora sugli schemi della prevenzione»[162].

Il progetto di ricerca, sviluppo e sperimentazione della Polizia Predittiva per la Sicurezza Urbana secondo il trovato, prende il via a valle di uno studio ventennale sui fenomeni di devianza urbana che ha permesso di distinguere con maggiore precisione l'illegalità diffusa nelle città e di comprendere che furti, rapine, scippi, borseggi, truffe e altri delitti, hanno caratteristiche di ciclicità e stanzialità e possono essere previsti qualora si riesca a definire un'appropriata logica di previsione e trasferirla a un modello di apprendimento automatico (Machine Learning) che sulla base di talune informazioni sui fenomeni delittuosi e sulle dinamiche sociali nei contesti in cui essi accadono, ha il compito di decodificare il disegno criminoso all'origine dei singoli delitti e di prevederne la riconfigurazione nel tempo e nello spazio, con il fine di creare la condizione ideale per riuscire a prevenirli più efficacemente rispetto al metodo tradizionale.
Frutto di anni di studio multidisciplinare, il modello si basa su di un evoluto metodo di analisi mai applicato prima per analizzare e affrontare il rischio criminale e che sorpassa tutti gli approcci sinora noti come quello del crime linking, del calcolo probabilistico e della elaborazione su carta topografica di zone di maggiore incidenza criminale (hot spot) i quali, presentano aree grigie e criticità di tipo etico e funzionale.

[162] Prof. Giacomo Di Gennaro Dipartimento di Scienze Politiche Direttore del Master di II livello Criminologia e Diritto Penale Analisi Criminale e Politiche per la Sicurezza Urbana Università Federico II di Napoli

La sperimentazione del trovato da parte di più strutture di sicurezza condotta con il contributo di due università per la supervisione indipendente del lavoro, ha permesso di stabilire trasparentemente che basando le attività sulla selettività e sequenzialità dei controlli in virtù degli allarmi predittivi elaborati secondo il modello d'Intelligenza Artificiale, si ottiene un netto stravolgimento del metodo di lavoro normalmente basato sul pronto intervento pertanto, la prevenzione dei crimini ne esce arricchita grazie ad una migliore razionalizzazione delle risorse e delle energie che in questo modo si evita di sprecare.
Attraverso un articolato framework, il trovato è stato introdotto per la sperimentazione indipendente in divisioni operative di sicurezza di diverse città italiane e impiegato secondo un protocollo esclusivo per favorirne la più agevole e corretta applicazione. Con l'obiettivo di tentare di migliorare l'attività di Prevenzione dei crimini nelle aree urbane secondo il diverso paradigma, dato dalla possibilità di poter prevenire gli illeciti prevedendoli grazie all'Intelligenza Artificiale, sono stati predisposti i controlli sul territorio in maniera selettiva e sequenziale e pertanto con dinamica precisione e puntualità rispetto al reale grado di progressione del rischio nel tempo e nello spazio. Il supporto operativo dell'Intelligenza Artificiale, ha permesso agli operatori di acquisire, in simbiosi con la stessa, maggiore consapevolezza del rischio e capacità decisionale direttamente nello scenario operativo e di partecipare proattivamente a un' appropriata strategia di "risk assessment" che sorpassa il limite di dover lavorare solo sull'emergenza o sulla base di analisi e valutazioni empiriche e obsolescenti.

I risultati ottenuti non sono casuali perchè si situano in quel progressivo itinerario rinvenibile nella copiosa letteratura sulle dinamiche della deterrenza (Teoria dei Giochi). Molti studi dedicati a questo argomento infatti, dimostrano che per questi delitti la punizione del reo non dovrebbe essere il primo obiettivo da perseguire, perchè il tentativo di infliggere la pena è casuale, costoso e produce scarsi risultati. Ciò che dovrebbe essere maggiormente valorizzato nella quotidiana disputa con il reo, è invece la deterrenza perché di fatto è una pena che limita pesantemente il disegno criminoso e maggiori quindi dovrebbero essere gli sforzi per arrivare a infliggerla con sistematicità. Infatti, se i potenziali autori di reato sono sufficientemente influenzabili, come si è riuscito a dimostrare, allora aumentare la probabilità di condizionamento, non solo può ridurre l'ammontare della pena effettivamente inflitta ma ribalta una situazione, dal suo equilibrio di alta violazione, al suo equilibrio di bassa violazione. I risultati, oltre al potenziale che

offrono per ridurre la criminalità e la carcerazione, possono avere importanti implicazioni soprattutto per la corretta gestione del problema, quello dell'Insicurezza Urbana.
La promessa di questa tecnologia è svoltare l'attività degli investigatori, in termini di efficacia operativa e di rapidità. Basti pensare che secondo i risultati delle sperimentazioni, X-Law è in grado di predire l'avvenimento criminoso con una precisione tra l'87 e il 93%. Questo cosa vuol dire? Meno rapine, furti, borseggi; razionalizzazione degli interventi e delle risorse delle forze dell'ordine; città più sicure. Ma qual è il prezzo da pagare?

CAPITOLO III
IA E GIUSTIZIA PREDITTIVA

SOMMARIO: 1. Algoritmi e processi di apprendimento esperti; 1.1.Algoritmi predittivi: una piattaforma sul futuro o un'àncora sul passato?; 2. Intelligenza artificiale e giusto processo; 2.1. Intelligenza artificiale e imparzialità del giudice; 2.2. Intelligenza artificiale e apporto alla rapidità del processo; 2.3. Intelligenza artificiale e uguaglianza dei cittadini; 3 Caso State vs Loomis; 3.1. La sentenza della Corte Suprema; 3.2. Valutazioni conclusive sull'utilizzo dello strumento COMPAS; 3.3. Analisi sul possibile esito discriminatorio di COMPAS; 4. Attività del giudice: uomo o robot?; 4.1. Sussunzione della fattispecie; 4.2. Il giudice robot e il rapporto con i precedenti; 4.3. Errori robotici: un problema di responsabilità giuridica; 5 L'algoritmo in campo probatorio; 5.1. Prove paraconsistenti; 5.2. Il "calcolo" della prova; 5.3. Digital Forensics: la prova digitale; 5.4. Valutazione della prova e legame con le neuroscienze; 6. Giustizia predittiva: articolo 33 della legge n.2019/222

1. Algoritmi e processi di apprendimento esperti

Lo sviluppo della cibernetica e dei processi tecnologici ha portato ad una sempre maggiore evoluzione ed utilizzo degli algoritmi, considerati la soluzione a innumerevoli problemi, utili a facilitare la vita di tutti i giorni, partendo dalle più semplici questioni, nella speranza di poter risolvere un giorno quelle di maggior rilievo. Ma partiamo dal principio: cosa è un algoritmo?

Gli algoritmi sono sequenze matematiche che estrapolano regolarità da basi di dati, codificati in modo tale da essere processabili attraverso formule matematiche.

Esso è, dunque, una sequenza finita di regole formali che rendono possibile il raggiungimento di un risultato a partire da un set iniziale di informazioni in input[163].

Ogni algoritmo, sottostà alle regole matematiche, l'Intelligenza Artificiale, invece, la quale si basa principalmente sulle tecniche di ragionamento e di apprendimento, viene distinta dal mero algoritmo: ciò che distingue l'Intelligenza Artificiale dall'algoritmo è l'autoapprendimento, ovvero il machine learning, il quale permette a sistemi di IA di

[163] C. CASTELLI, D. PIANA, Giusto processo e intelligenza artificiale, Maggioli, Santarcangelo 2019

applicare la logica deduttiva allo scopo di trarre fattori nuovi e ignoti partendo da fatti noti.

Dunque, la macchina non fa che recepire dati e, tramite il machine learning, riesce ad elaborarli e a creare situazioni nuove, che da quelle precedenti hanno preso forma. È tramite la tecnica di apprendimento, invece, che la macchina potrebbe riuscire ad effettuare previsioni anche abbastanza accurate, tramite l'analisi di immagini e ricollegamenti di situazioni simili, questo è il caso della cosiddetta giustizia predittiva[164].

Un sistema esperto, pertanto, è uno dei metodi dell'intelligenza artificiale. È uno strumento capace di rappresentare meccanismi cognitivi di un esperto in un ambito specifico. Più precisamente, si tratta di un software che è capace di rispondere a domande, di elaborare ragionamenti basandosi su fatti noti e su regole, un programma che riproduce le prestazioni di una o più persone esperte in un determinato campo di attività[165].

Il sistema esperto è in grado di effettuare attività che richiedano particolari competenze, tramite un motore che sfrutta la base di conoscenza per risolvere svariati problemi[166].

Due sono le categorie principali entro le quali possiamo oggi suddividere i sistemi esperti:

1) sistemi esperti basati su regole: si tratta di sistemi basati su regole classiche e ben note al mondo dell'informatica nella forma IF (condizione) e THEN (azione). Dati una serie di fatti i sistemi esperti, grazie alle regole di cui sono composti, riescono a dedurre nuovi fatti.

2) sistemi esperti basati su alberi: in questo caso, dato un insieme di dati ed alcune deduzioni, il sistema esperto crea un albero che classifica i vari dati. Di fronte ad un problema, nuovi dati vengono analizzati dall'albero e il nodo finale rappresenta la soluzione.

Un sistema esperto basato su alberi è, in sostanza, un software esperto in grado di riconoscere un problema in base a una sequenza di fatti, decisioni o azioni. A partire da una situazione iniziale, tutte le scelte si ramificano in situazioni e azioni fino a giungere alla conclusione[167].

1.1. Algoritmi predittivi: una piattaforma sul futuro o un'àncora sul passato?

Una delle possibili applicazioni degli algoritmi nel sistema giudiziario riguarda l'uso degli algoritmi predittivi. Difatti, si stanno diffondendo sempre maggiormente i cosiddetti

[164] https://www.altrascienza.it/2019/06/dallintelligenza-artificiale-al-pensiero-sensibile/
[165] CEPEJ, Carta etica sull'uso dell'intelligenza artificiale nei sistemi giudiziari e nel loro ambiente, 2018
[166] G. SARTOR, Intelligenza artificiale e diritto: Un'introduzione, Giuffrè, Milano, 1996
[167] https://www.ai4business.it/intelligenza-artificiale/sistemi-esperti-cosa-sono/

risk assessments tools, strumenti in grado di calcolare il rischio che un prevenuto si sottragga al processo o commetta dei reati.

Questi strumenti analizzano uno svariato numero di dati relativi al passato e individuano delle ricorrenze (ossia dei pattern), caratterizzate da una base statistica anche più solida di quelle che stanno al fondo dei giudizi umani[168]. I primi studi sulla calcolabilità del ragionamento giuridico, riconducibili a Leibniz e alla sua "Ars combinatoria"[169], hanno trovato fondamento e concretezza attraverso l'elaborazione di programmi in grado di riprodurre in maniera automatizzata la logica giuridica.

Leibniz si proponeva di estendere al dominio della giurisdizione l'ideale della calcolabilità universale, che animava tutta la sua riflessione sulla logica. Una calcolabilità che comportava piena prevedibilità[170].

In seguito, lo stesso Max Weber in "Economia e società", nel lontano 1947, asseriva che il diritto è basato su regole scritte e la certezza del diritto altro non è se non la prevedibilità dell'esito giudiziale; in fondo, sarebbe pur sempre un diritto razionale formale[171].

Sebbene il GDPR cerchi di regolare e prevenire la diffusione di dati, fungendo da "scudo protettivo", vi sono casi in cui anche lo stesso GDPR debba essere sottoposto a criteri interpretativi e integrativi. Ad esempio, l'articolo 22 GDPR, vieta di essere sottoposti a decisioni "unicamente" automatizzate, il che vuol dire che, laddove intervenga una decisione fondata su un algoritmo, ma sia comunque intercettabile un intervento umano (come sostenuto appunto dalla Corte del Wisconsin nel caso Loomis), lo scudo dell'art. 22 GDPR non sarà applicabile[172].

L'utilizzo di algoritmi con modelli predittivi, potrebbe, paradossalmente, essere ritenuto come intrinsecamente conservatore, perché fondamentalmente non fa altro che usare dati del passato per prevedere un comportamento futuro. E in certi campi, un comportamento che devia da quello passato potrebbe addirittura essere indice dell'avvio di un percorso delinquenziale o terroristico.95

La neuroscienza spiega ciò descrivendo quella che è la differenza tra la mente umana e la macchina: la nostra mente è flessibile, il nostro pensiero anche lo è.

---

[168] https://www.penalecontemporaneo.it/upload/6903-gialuz2019b.pdf

[169] "Tutte le questioni di diritto puro s[o]no definibili con certezza geometrica" con questo dictum Leibniz ha notato come sia possibile ridurre i concetti complessi ad un piccolo numero di concetti primitivi, ciascuno connotato da un segno.
Dopo aver stabilito una classificazione di concetti primitivi, egli pensa che si possa arrivare a stabilire una sorta di scrittura universale simbolica e con questa risolvere i problemi logici così come si risolvono i problemi algebrici.

[170] M. LUCIANI, La decisione giudiziaria robotica, Rivista N°: 3/2018, data pubblicazione: 30/09/2018

[171] M. WEBER, Economia e società, vol. III, Edizioni di Comunità, Milano, 1980, p. 17. Il diritto deve rispondere a requisiti di prevedibilità e calcolabilità quale strumento al servizio della moderna economia capitalistica.

[172] https://www.dataprotectionlaw.it/2019/05/07/lalgoritmo-che-condanna-i-limiti-della-giustizia- predittiva/

Questa flessibilità determina anche la nostra capacità di adattamento, questo la macchina non ce l'ha, la macchina non ha spirito di adattamento o istinto di sopravvivenza, non ha, quindi, ciò che ci rende "creativi", essa non è creativa.
Questo è il confine: la macchina pensa sempre al passato, mai al futuro, dunque, come già affermato, ha tratti conservatori[173].
Bisogna, dunque, andare a effettuare una valutazione dei pro e dei contro, quanto questi algoritmi possano permettere davvero di predire il rischio di delinquere e arrivare a "sventare" futuri crimini, e quanto il loro utilizzo non rischi di far radicare nella mente delle persone un pregiudizio riservato solo a determinate categorie di individui.

2. Intelligenza artificiale e giusto processo

Le tecniche di elaborazione automatizzata dei dati, come gli algoritmi, non solo consentono agli utenti di Internet di cercare e accedere alle informazioni, ma sono anche sempre più utilizzate nei processi decisionali, che prima erano interamente di competenza dell'uomo. Gli algoritmi possono essere utilizzati per "preparare" le decisioni umane o per prenderle immediatamente con mezzi automatizzati. In effetti, i confini tra il processo decisionale umano e quello automatizzato sono spesso confusi, il che porta alla nozione di processo decisionale quasi o semiautomatico.
È innegabile che l'utilizzo di questo tipo di dati può presentare alcuni vantaggi, grazie alla rapidità di indagine e trattazione di insieme di dati massicci, oltre che alla possibilità di controllare e verificare i rischi in modo più accurato, tuttavia, sono sempre maggiori le preoccupazioni scaturenti da un utilizzo quasi eccessivo di tali tecnologie. A seguito di una serie di attacchi terroristici negli Stati Uniti e in Europa, i politici hanno chiesto alle piattaforme di social media online di utilizzare i loro algoritmi per identificare potenziali terroristi e agire di conseguenza. Alcune di queste piattaforme stanno già utilizzando algoritmi per identificare gli account che generano contenuti estremisti.
Oltre all'impatto significativo che tale applicazione di algoritmi ha per la libertà di espressione, solleva anche rilevanti preoccupazioni per le norme sul giusto processo.
Proprio in riferimento all'articolo 6 della CEDU, in particolare vengono messi in pericolo principi quali la presunzione di innocenza, il diritto di essere informati tempestivamente della causa e della natura di un'accusa, il diritto a un processo equo e il diritto di difendersi

[173] Convegno tenutosi presso l'Università degli Studi di Salerno, il quale ha visto come ospite il professore Jordi Nieva-Fenoll, per la presentazione del suo libro "Intelligenza artificiale e processo", 7 Giugno 2019.

di persona. Possono sorgere preoccupazioni anche per quanto riguarda l'articolo 5 della CEDU, che protegge dalla privazione arbitraria della libertà, e l'articolo 7 (nessuna pena senza legge). Nel campo della prevenzione della criminalità, i principali dibattiti orientativi sull'uso degli algoritmi si riferiscono alla polizia predittiva.

Gli apporti offerti dall'utilizzo degli algoritmi, a volte, possono essere estremamente pregiudizievoli in termini di origine etnica e razziale e richiedono pertanto una supervisione scrupolosa e garanzie adeguate. Spesso i sistemi sono basati su banche dati di polizia esistenti che riflettono intenzionalmente o involontariamente pregiudizi sistemici. A seconda delle modalità di registrazione dei reati, della scelta dei reati da includere nell'analisi e degli strumenti analitici utilizzati, gli algoritmi predittivi possono quindi contribuire a decisioni pregiudizievoli e a risultati discriminatori[174].

Vi sono, difatti, numerosi esempi di discriminazioni derivanti dall'utilizzo degli algoritmi.

Negli Stati Uniti, ad esempio, gli "strumenti di previsione del crimine" hanno dimostrato di discriminare alcune minoranze di individui. Sulla base di eventi passati, l'algoritmo utilizzato ha assegnato un punteggio di rischio più elevato a determinate minoranze etniche e ciò ha fatto che sì la polizia si soffermasse maggiormente su questo gruppo di individui nelle sue ricerche.

Un altro esempio di esiti discriminatori deriva dal processo di selezione effettuato dagli algoritmi. Ad esempio, generalmente, la maggior parte degli insegnanti della scuola primaria è di genere femminile. Dunque, un algoritmo utilizzato per selezionare il personale adeguato per tale lavoro si basa principalmente sui curricula ricevuti in passato. Dato che le scuole primarie impiegano molte più donne che uomini, l'algoritmo svilupperà rapidamente una preferenza per le candidate donne, e, se anche si rendessero i curricula neutrali dal punto di vista del genere, ciò non risolverebbe il problema. L'algoritmo in tal modo tenderà a prediligere determinati hobby femminili, assegnando meno punti ai curricula che elencano passatempi tradizionalmente maschili.

Tali errori, però, non sono errori dell'algoritmo, ma bensì dell'uomo stesso, non è l'algoritmo che è sbagliato, piuttosto sono gli uomini a fare delle discriminazioni e l'algoritmo non fa altro che rilevare questo errore[175].

Alla base dei rischi, difatti, che minacciano la stessa indipendenza del giudice, vi è anche la pressione esercitata su di loro da altri poteri dello Stato.

---

[174] https://www.pandslegal.it/tecnologie-ict/algoritmi-e-giusto-processo-limpatto-dell-ia/

[175] J. NIEVA FENOLL, Inteligencia artificial y proceso judicial, Marcial Pons Ediciones Juridicas y Sociales, Madrid 2018

Nel caso dell'algoritmo bisognerebbe verificare che esso non sia stato alterato in favore di determinati soggetti o interessi, da parte degli stessi programmatori che avevano l'intenzione di favorire alcuni gruppi nel processo[176].

In Virginia, un professore universitario, il quale ha studiato gli algoritmi predittivi da tempo utilizzati in ambito giudiziario, ha riscontrato che l'algoritmo era solito produrre risultati "sessisti", dovuti alle errate associazioni che venivano fatte e che riconducevano sempre alla figura della donna immagini di cucina shopping ecc...

L'errore non proveniva dalla macchina stessa, ma dagli input che le erano rivolti, i quali erano già intrisi di una distorsione di genere, il che, ovviamente, conduceva anche la macchina a risultati alimentati dal pregiudizio[177].

La discriminazione mediante algoritmi è, quindi, un riflesso della discriminazione già in atto "sul campo". Poiché gli algoritmi sono sempre costruiti su dati storici, è praticamente impossibile trovare un set di dati "pulito" su cui un algoritmo possa essere implementato senza essere "privo di pregiudizi"[178].

La stessa neutralità del giudice automatico, infatti, sarebbe illusoria, in quanto i suoi pregiudizi sarebbero insiti nello stesso algoritmo che lo anima e non farebbero che proiettare nel processo decisionale i pregiudizi del compilatore del programma[179].

L'art. 8 d.lgs. n. 51 del 2018, rubricato "Processo decisionale automatizzato relativo alle persone fisiche"[180], in conformità a quanto sancito nel regolamento (UE) 2016/679, al 1 comma sancisce che: "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." Al comma 2 invece: "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."

Ciò significa che, sebbene siano ammesse decisioni basate parzialmente su un trattamento personalizzato, è comunque garantito l'intervento umano. Sul piano costituzionale, sono stati richiamati il principio di soggezione del giudice solo alla legge (ex art 101 Cost.) e

[176] https://www.agendadigitale.eu/sicurezza/decisioni-automatizzate-quale-difesa-da-chi-ci- discrimina-via-algoritmi/

[177] https://www.valigiablu.it/algoritmi-dati-rischi/

[178] https://www.agendadigitale.eu/sicurezza/decisioni-automatizzate-quale-difesa-da-chi-ci- discrimina-via-algoritmi/

[179] https://www.4clegal.com/hot-topic/giudice-automatico-giudizio-penale-intelligenza-artificiale

[180] https://www.gazzettaufficiale.it/eli/id/2018/05/24/18G00080/sg

la previsione dell'accesso in magistratura soltanto per concorso (Art 106 Cost.), norme che presupporrebbero che il giudice sia un essere umano e non una macchina.
In secondo luogo, si è notato che la decisione della macchina difficilmente è trasparente, cioè non assicura la conoscibilità di tutti i passaggi decisionali, con evidente lesione del principio di difesa, di contraddittorio e di trasparenza delle decisioni.
Si è considerato che la "giustizia predittiva" incarna il mito illuminista del "giudice bocca della legge", e si sono evidenziate le anomalie di una giustizia siffatta: alla imparzialità del giudice, si sostituisce la incorporeità e la a-storicità di una macchina che ius dicit al di fuori della storia, cioè lo spazio abitato dagli umani, cioè dai loro corpi.
Per il momento resta la clausola di garanzia finale recata dalla norma sopra indicata, ma limitata: è garantito l'intervento umano, non chiare le sue modalità e certa la sua decisività. Ma resta anche il problema ampiamente evidenziato, del rapporto fra la decisione del giudice e quella della macchina, tutto da verificare e scoprire[181].
Ma come funziona la giustizia predittiva? Tramite l'utilizzo del machine learning, di cui in precedenza si è accennato, esso rientra in un sottocampo dell'Intelligenza Artificiale e si pone l'obiettivo di far apprendere in modo automatico alle macchine attività svolte da noi esseri umani. L'algoritmo in tal caso non è progettato con le informazioni per raggiungere un determinato fine, ma apprende con l'"esperienza" quale sia la scelta migliore per il soddisfacimento del fine stesso.
Tramite il machine learning sarebbe possibile prevedere l'esito di una controversia sulla base dell'analisi dei dati ricavati dai precedenti o fornire comunque dati di conoscenza sul contenuto della possibile decisione[182]. Sono ammissibili apporti della elaborazione informatica di dati, ma in campi limitati e circoscritti, connotati da ampia discrezionalità giudiziale ma anche da immediata misurabilità, nei quali è essenziale, ottenere sintesi di innumerevoli e disparate decisioni, ma dove l'interferenza con l'attuazione di regole e di valori è evitata perché la stessa valutazione quantitativa è già stata operata dal giudice a monte: per esempio come per la quantificazione del danno biologico.
Il giurista avveduto sa bene che la storia si nutre anche di discontinuità, e che talvolta sono state proprio le interpretazioni di rottura a innescare spirali volte al rafforzamento di interessi sino a quel momento privi di tutela. Tutto questo l'algoritmo non è in grado di

[181] C. CASTELLI, D. PIANA, Giusto processo e intelligenza artificiale, Maggioli, Santarcangelo 2019
[182] http://www.questionegiustizia.it/articolo/intelligenza-artificiale-e-processo-penale-tra-norme- prassi-e-prospettive-_18-10-2019.php

comprenderlo, se non registrando prima, e mimetizzando poi, una sequenza di discontinuità che in tal modo diventerebbero a loro volta continue[183]!

2.1. Intelligenza artificiale e l'imparzialità del giudice

Per il giudice, la libertà può essere descritta così: in seguito all'ascolto delle parti, in osservanza del principio del contraddittorio, dovrà condurre la sua analisi, rispettando regole e valori quali, ad esempio, quelli derivanti da convenzioni internazionali come la Convenzione europea per la protezione dei diritti umani e delle libertà fondamentali, la Carta dei diritti fondamentali dell'Ue e altri principi dell'ordinamento europeo. Tale libertà presuppone, per statuto, l'indipendenza del giudice (giudice terzo e imparziale), il quale dovrà sempre collocarsi alla giusta distanza dal contenzioso a lui sottoposto, senza alcuna influenza esterna: per queste stesse ragioni, il giudice è libero[184].

L'imparzialità, dunque, prevede che il giudice sia estraneo sia all'oggetto del giudizio che all'interesse delle parti[185].

Nell'esercizio delle sue funzioni, il giudice deve assicurare il rispetto della "parità delle armi", evitando qualunque tipo di asimmetria tra le parti, garantendo per ciascuna di esse la possibilità di difendersi e agire nei propri interessi. È possibile attribuire la medesima funzione a un algoritmo?

Al riguardo, dovrà tenersi in ampia considerazione il ruolo dei principi fondamentali, sia nazionali che internazionali. Al giudice spetta l'utilizzo di tutte le risorse giuridiche, compreso l'insieme dei diritti fondamentali.

In un documento redatto dalla Rete europea dei Consigli di giustizia (ENCJ), che illustra gli obblighi deontologici dei giudici, è scritto:

«Il senso di umanità del giudice si manifesta attraverso il rispetto delle persone e della loro dignità, in tutte le circostanze della sua vita professionale e privata.

La sua condotta si basa sul rispetto della persona umana, considerandone tutte le caratteristiche: fisiche, culturali, intellettuali, sociali, nonché relative alla razza e al genere. Il giudice mostra rispetto nel suo rapporto con le parti in causa, ma anche nei confronti di coloro che compongono il suo ambiente professionale, gli avvocati, il personale amministrativo, etc.

Questa umanità, che comprende una sensibilità per le situazioni sottoposte al suo giudizio, permette al giudice di considerare la dimensione umana delle sue decisioni. Spetta a lui,

[183] http://www.medialaws.eu/algoritmo-in-tribunale-giudice-o-imputato/
[184] http://questionegiustizia.it/rivista/2018/4/liberta-e-umanita-del-giudice-due-valori- fondament_593.php
[185] Art. 111 Cost. co2

nella valutazione dei fatti come nella fase decisionale, trovare un equilibrio tra empatia, compassione, comprensione, rigore e severità, in modo che la sua applicazione del diritto sia avvertita come legittima e giusta».

Il rispetto di questo requisito è importante per stabilire o restaurare la fiducia tra i cittadini e la giustizia[186].

A volte il giudice non si attiene completamente alla volontà del legislatore, creando la legge del caso concreto. Tutto ciò è frutto dell'attività discrezionale e creativa del giudice, il quale non è contemplato nel caso dell'algoritmo.

I sistemi di intelligenza artificiale si occupano di applicare in modo inflessibile la legge, senza alcun margine di apprezzamento basato su sensazioni personali, che di fatto la macchina non ha[187].

Quando si parla di imparzialità e terzietà, però, si chiede al giudice di prendere le distanze dalle emozioni?

L'amore (ma anche l'odio, il risentimento, l'inimicizia), valgono come cause di ricusazione, quindi «è un bene o un male» che il giudice decida seguendo le proprie emozioni?

Se fossimo portati a rispondere che è un bene, perché le emozioni denotano umanità e una decisione umana appare più giusta, allora dovremmo rifiutare le "decisioni robotiche", in quanto prive di emozioni, se, invece, fossimo indotti a rispondere che è un male, perché il giudizio non sarebbe del tutto imparziale, le "decisioni robotiche" possono risultare la migliore soluzione[188].

Numerose sono le contraddizioni, in quanto è lo stesso pensiero umano ad essere fonte di dubbi, non essendo ancora del tutto chiaro come esso funzioni, l'imitarlo diventa un problema secondario[189].

Emmanuel Jeuland113, in uno studio sul giudice e l'emozione, si chiede: «Non è pericoloso mettere le proprie emozioni a distanza fino a chiamarsi fuori dalla loro sfera per raggiungere il giudizio?».

---

[186] http://questionegiustizia.it/articolo/stupidita-non-solo-artificiale-predittivita-e-processo_03-07-2019.php#_ftn7

[187] J. NIEVA FENOLL, Inteligencia artificial y proceso judicial, Marcial Pons Ediciones Juridicas y Sociales, Madrid, 2018.

[188] http://questionegiustizia.it/articolo/stupidita-non-solo-artificiale-predittivita-e-processo_03-07-2019.php#_ftn7

[189] Cfr. A. GARAPON e J. LASSEGUE,, Justice digitale, Presses Universitaires de France, Paris, 2018, che, nel dedicare un capitolo diquestointeressante volume, dove si analizzano molteplici aspetti dell'avvento del digitale nei sistemi giuridici e giudiziari e delle trasformazioni che esso può indurre, al tema delle "emozioni" dei giudici robot, si domandano: «Juges inanimes, aves-vouz une âme?» (pp. 195 ss.).

«Sembra acquisito», prosegue, «alla luce delle scoperte delle neuroscienze, che Descartes abbia sbagliato nel distinguere tra corpo e mente, ragione ed emozione (...). Una ragione senza emozione porta a cattive decisioni, laddove una gestione incontrollata delle emozioni porterà a cattive intuizioni»

Antonio Damasio[190], afferma che l'emozione partecipa della ragione e può aiutare il processo di ragionamento, anziché necessariamente disturbarlo, come a lungo si è pensato, le emozioni sono risposte fisiologiche che mirano ad ottimizzare le azioni intraprese dall'organismo nel mondo che lo circonda.

Sostenendo questo, sembra alquanto difficile la possibilità di costruire un computer dotato di coscienza[191].

Bisogna ammettere però, che vi sono casi in cui la stessa macchina potrebbe arrivare a "simulare" le emozioni umane: rilevando, ad esempio, informazioni che denotano debolezza in una delle parti e cambiando la propria decisione, se pur con freddezza e sistematicità, sempre sulla base di dati che normalmente inducono pietà negli esseri umani[192].

Ma in che misura la IA può imitare il nostro pensiero?[193] La IA non pensa e non decide, può fare finta di sentire emozioni, ma non sente emozione.

Per sapere se la macchina può imitarci, bisogna capire come pensiamo noi: in tal ambito vi è già uno studio molto sviluppato dalla psicologia cognitiva che svolge questo lavoro: essa ci dice che noi umani effettuiamo le nostre scelte mediante calcoli statistici imprecisi che facciamo nello stesso momento in cui prendiamo la decisione, per esempio: nello scegliere cosa mangiare a cena decido in base a cosa ho mangiato a pranzo, in tal caso, effettuo un calcolo statistico.

Questo criterio definito "euristico di rappresentatività", ci spiega come effettuiamo le nostre decisioni nella maggior parte dei casi, oltre quest'ultimo, ci sono altri 3 tipi di euristiche:

- La prima, è detta "euristica della disponibilità". Io decido, secondo quello che mi sembra più frequente, perché lo ricordo meglio, è più "disponibile" per me.

[190] Direttore presso l'istituto di ricerca legale alla Sorbona, ricordando un seminario tenuto all'Università di Nanterre nel febbraio 2017.

[191] Medico neurologo, ideatore del saggio "L'errore di Cartesio", Adelfi (Collana "Biblioteca Scientifica" n 22),1995

[192] http://questionegiustizia.it/rivista/2018/4/liberta-e-umanita-del-giudice-due-valori- fondament_593.php

[193] J. NIEVA FENOLL, Inteligencia artificial y proceso judicial, Marcial Pons Ediciones Juridicas y Sociales, Madrid, 2018

Esempio: tutti sappiamo che viaggiare in macchina è più pericoloso che viaggiare in aereo, eppure sentiamo maggiormente paura quando andiamo in aereo.

- In seguito, abbiamo l' "euristica dell'ancoraggio e aggiustamento". Questa provoca un "errore di conferma", ovvero, quando una persona crea un parere su un determinato argomento, fargli cambiare parere diventa molto difficile, anzi, in tal caso il soggetto ascolta gli argomenti della persona e li modella per sostenere il suo stesso parere.

Cambiare parere è sempre difficile, in quanto sembra di dover rimodulare il proprio pensiero, in materia giudiziaria, quando il giudice forma questo criterio, farlo cambiare risulta ancora più faticoso.

- Infine, abbiamo l'"euristica della affezione": Quest'ultima ci spiega come decidiamo quando ci troviamo di fronte a cose che ci piacciono, oppure no, e come questo cambi il nostro approccio in seguito. Esempio: qualcuno entra nell'aula e dice la parola "cioccolato", ovviamente questo porterà ad una reazione positiva, la persona in questione attirerà l'attenzione più facilmente, al contrario, se la stessa persona entra in aula e dice la parola "cancro", la reazione sarà diversa, anche ciò che verrà detto successivamente porterà ad un approccio differente.

La domanda da porsi è: può la macchina imitare tutti questi esempi di euristica?

La risposta è sì, nel caso della rappresentatività la macchina fa continuamente calcoli statistici, ed anche essi non sono privi di margini di errore, nel caso della disponibilità, non sbaglierebbe, come gli uomini nel caso dell'aereo, però potrebbe rilevare lo sbaglio umano e decidere di sbagliare anche essa. Nel caso dell'ancoraggio e aggiustamento, questa è la situazione basica di come funziona la IA, la quale non muta parere.

Per quanto riguarda le emozioni, la macchina può fare finta di amare, gioire, può fingere ed imitare le emozioni, imitare ciò che odiamo o amiamo noi. Tutti noi sappiamo cosa provoca gioia o tristezza negli altri, possiamo essere noi stessi umani ad imitare determinate emozioni, e come siamo in grado di farlo noi, possono essere in grado di farlo anche le macchine. Una macchina potrà applicare queste "emozioni" per una decisione concreta. Questa visione è affermata dall'esperto di psicologia Daniel Kahneman[194], il quale disse di non vedere alcuna ragione per cui il pensiero umano non possa essere imitato da una macchina[195].

[194] Psicologo israeliano, vincitore, insieme a Vernon Smith, del Premio Nobel per l'economia nel 2002.

[195] Convegno tenutosi presso l'Università degli Studi di Salerno, il quale ha visto come ospite il professore Jordi Nieva-Fenoll, per la presentazione del suo libro "Intelligenza artificiale e processo", 7 Giugno 2019.

Sembra essere una conferma di ciò, lo sviluppo della chatbot LAILA[196], la quale ha sviluppato un sistema di interazioni tra più sistemi di intelligenza artificiali, inoltre è dotata di un livello discorsivo notevole, comprendendo il linguaggio naturale e arrivando addirittura ad intuire emozioni ed esigenze del proprio interlocutore, essa è stata definita "empatica"[197].

2.2. Intelligenza artificiale e apporto alla rapidità del processo.

Il 2° comma dell'articolo 111 della nostra Costituzione, sancisce il principio della "ragionevole durata del processo", garantendo la parte affinché non resti "vittima" delle eccessive lungaggini processuali.

Vi è inoltre da considerare, anche il costo delle operazioni attinenti allo svolgimento e alle conseguenze dei provvedimenti giudiziari, il costo dell'apparato giudiziario, di quello repressivo e di vigilanza[198]. L'applicazione della intelligenza artificiale potrebbe non essere altro che un ulteriore strumento utile a condurre alla soluzione di uno dei problemi che più affliggono il sistema giudiziario: processi interminabili e dispendio di risorse e denaro.

Una recente critica al modello processuale spagnolo afferma che parte del lavoro dei giudici è meccanica. Buona parte dei funzionari giudiziari utilizza il metodo del "curta y paga" (copia e incolla), per copiare risoluzioni anteriori o integrarle con dati del caso trattato.

Poche sentenze vengono redatte completamente ex novo.

Questa riflessione può essere sviluppata dalla intelligenza artificiale, con essa non solamente si ottiene una maggior varietà di creazioni di documenti e di copie degli stessi, ma si consegue anche una maggior capacità di analisi di detti documenti.

Dalla combinazione di queste tre funzionalità si consegue che l'applicazione dell'intelligenza artificiale potrà essere incredibilmente più rapida di un giudice nella risoluzione di procedimenti prevedibili, soprattutto nell'analisi della documentazione[199].

[196] Chatbot ideata dalla Mazer Srl.

[197] Per maggiori informazioni https://www.laila.tech/

[198] V. FROSINI, Cibernetica: Diritto e società, Ed. di Comunità (Collana diritto e Cultura moderna n 6), Milano 1968

[199] J.NIEVA FENOLL, Inteligencia artificial y proceso judicial, Madrid,2018. Per risparmiare i tempi, basterebbe che le parti sollevassero le loro pretese e le loro difese attraverso un'applicazione, per richieste semplici come la richiesta di una somma indiscussa di denaro o uno sfratto per mancato pagamento.

Sicuramente le lungaggini giudiziarie sono uno dei principali problemi, a cui l'intelligenza artificiale può trovare pronta soluzione, pian piano, partendo da semplici richieste, fino a raggiungere le grandi cause.
Ogni sistema di IA ha per oggetto e presupposto un "complesso" di dati e informazioni. Chi progetta, realizza e infine utilizza un sistema sa che deve immettervi un insieme di dati e informazioni, destinati a essere elaborati. La più complessa, raffinata, veloce forma di IA diventa fallibile e mediocre se non "alimentata" in termini adeguati[200].
Il sistema giudiziario è intasato da cause di minima entità e ripetitive tra loro, l'utilizzo della IA può essere la più efficiente soluzione al loro smaltimento.
Un esempio lampante di ciò è offerto dall'Estonia, che ha sperimentato l'utilizzo di robot per lo svolgimento di funzioni giudiziarie, per ciò che concerne le cause di lieve entità.

2.3. Intelligenza artificiale e uguaglianza dei cittadini

L'idea di eguaglianza dinanzi alla legge e la conseguente necessaria indipendenza dell'organo che in prima istanza è chiamato ad applicare le leggi in uno stato costituzionale, è così radicata nella cultura politica delle democrazie avanzate che, quando si verificano delle condizioni per cui l'eguaglianza non viene rispettata, la legittimità stessa delle istituzioni è messa in discussione.
In fondo, nessun cittadino sottoscrive ogni giorno il patto democratico, ma lo fa in silenzio, implicitamente accettando e seguendo le regole del vivere civile.
La legittimità stessa dell'azione della magistratura dipende da come il cittadino la "percepisce": il cittadino non ha le conoscenze tecniche e specialistiche per apprezzare quanto accade nel lungo e complesso iter che caratterizza un processo, quindi si affida ad un legale. Perché affidarsi ad un giudice quindi? Perché il cittadino chiede giustizia e la chiede all'istituzione da cui si attende una giustizia giusta.
Cosa si intende però, per "Giustizia giusta"? essa non è la mera applicazione della legge, il cittadino chiede una giustizia giusta che non è determinabile in modo automatico. Se il punto di partenza normativo è uguale per tutti, per valutare se la legge sia uguale per tutti l'osservazione delle norme giuridiche non basta. Occorre tenere conto degli altri fattori, quali ad esempio la esistenza di un buon dialogo tra il foro e la magistratura del circondario, oppure la disponibilità di servizi di orientamento e informazione per il pubblico ecc...

[200] https://www.penalecontemporaneo.it/upload/6695-parodisellaroli2019a.pdf

Dunque, bisogna studiare il contesto in cui viene svolta questa delicata funzione che è "rendere giustizia".

La variabile spazio-tempo si sta rivelando sempre più importante dal punto di vista dei servizi resi al cittadino, compreso nel settore giudiziario. Ammesso, quindi, che la giustizia sia resa in modi e tempi diversi a seconda dei territori dove il cittadino vive, questa diversità è prevedibile dal cittadino?

Disomogeneità e imprevedibilità hanno un effetto di distorsione sulla uguaglianza dinanzi alla giustizia, nel modo in cui essa è resa e rispetto alle risorse richieste per l'accesso.

L'utilizzo della tecnologia, la quale è intesa come un sistema di regole che si sovrappongono e si giustappongono a quelle della organizzazione giudiziaria e quasi come per vantaggio comparativo tendono a sostituirla laddove i casi sono semplici e seriali, può essere vista come una sorta di garanzia funzionale di uguaglianza di trattamento?

Se un caso viene trattato sulla base di una serie di procedure che sono fissate ex ante attraverso piattaforme di scrittura di documenti e meccanismi di elaborazione dati, ovvero processo di calcolo che trattano il caso N+1 esattamente come sono stati trattati i casi M-N, da cui si è estrapolata, in via matematica, una regola, non viene dunque trattato in modo eguale a quelli precedenti e in modo perfettamente prevedibile?[201]

Come già accennato, numerosi sono stati i casi di discriminazione avvenuti a causa dell'impiego di algoritmi a scopo di selezione.

Merita di essere qui citata la sentenza 8 aprile 2019 n. 2270, in seguito ad un appello dinanzi al Consiglio di Stato nei confronti del Ministero dell'Istruzione, poiché quest'ultimo aveva gestito la formazione delle graduatorie tramite uno specifico processo algoritmico, il quale avrebbe prodotto una erronea lavorazione delle informazioni.

In particolare, docenti con diversi anni di servizio si sono ritrovati destinatari in ambiti territoriali non richiesti, mentre altri docenti con minori anni di servizio e con meno qualifiche, hanno beneficiato della classe di concorso prescelta.

Gli stessi Giudici del Tar Lazio avevano affermato: "gli istituti di partecipazione, di trasparenza e di accesso, in sintesi, di relazione del privato con i pubblici poteri non possono essere legittimamente mortificate e compresse soppiantando l'attività umana con quella impersonale, che poi non è attività, ossia prodotto delle azioni dell'uomo, che può essere svolta in applicazione di regole o procedure informatiche o matematiche"[202].

[201] C. CASTELLI, D. PIANA, Giusto processo e intelligenza artificiale, Maggioli, Santarcangelo 2019.

[202] Il Tribunale Amministrativo Regionale per il Lazio, (Sezione Terza Bis), n. 10964/2019 pubblicato il 13/09/2019.

Diversa la posizioni dei Giudici di Palazzo Spada, sede del Consiglio di Stato: "l'assenza di intervento umano in un'attività di mera classificazione automatica di istanze numerose, secondo regole predeterminate, e l'affidamento di tale attività a un efficiente elaboratore elettronico appaiono come doverose declinazioni dell'art. 97 Cost.[203] coerenti con l'attuale evoluzione tecnologica"[204].

L'esclusione dell'intervento umano, in particolare per operazioni meramente ripetitive e prive di discrezionalità, serve ad evitare interferenze dovute a negligenza (o peggio dolo) del funzionario (essere umano), in tal modo garantendo maggiore garanzia di imparzialità della decisione automatizzata.

Innegabile, però, è che l'operato dell'algoritmo, sebbene sia autorizzato e talune volte preferito, debba essere conforme ai principi dell'ordinamento, ed in quanto tale, sindacabile[205].

In un'intervista rivolta ad Andrea Mascherin[206], gli viene chiesto come veda l'utilizzo della tecnologia, impiegata in ambito giuridico, alla luce dei principi che sorreggono il nostro ordinamento quale lo stesso principio di eguaglianza.

Il presidente si è trovato scettico sul punto, pur non negando che senza dubbio la rivoluzione tecnologica investe numerosi settori, tra cui lo stesso settore giuridico, forti contrasti sorgono intorno alla possibilità di applicare una giustizia predittiva, in relazione ai limiti posti dall'evoluzione dinamica della giurisprudenza e del diritto. Nonostante ciò, egli non nega che l'innovazione debba essere sfruttata al meglio, e che sia necessario che giudici e avvocati siano dotati di elevate capacità tecniche e professionali in modo da riuscire ad utilizzare l'intelligenza artificiale con equilibrio.

3. Caso State vs Loomis

La intelligenza artificiale può avere un notevole impatto in materia di misure cautelari. Come affermato da Piero Calamandrei, per poter applicare le misure cautelari, si necessita di due presupposti: il fumus boni iuris e il periculum in mora. Il fumus boni iuris consiste nella verosimiglianza o la probabilità dell'esistenza di un diritto, ma è il periculum in mora, quale elemento probabilistico per eccellenza, a dover essere valutato per una

[203] Art. 97 co 1: "I pubblici uffici sono organizzati secondo disposizioni di legge, in modo che siano assicurati il buon andamento e l'imparzialità dell'amministrazione."
[204] Consiglio di Stato, sentenza 8 aprile 2019 n. 2270
[205] https://www.4clegal.com/settore-pubblico/lintelligenza-artificiale-tinge-diritto-0
[206] Presidente del Consiglio Nazionale Forense, intervista rilasciata il 27 novembre 2018. C.Castelli, D.Piana, Giusto processo e intelligenza artificiale, Maggioli, Santarcangelo 2019.

possibile recidiva: e questo risulta essere l'elemento più dibattuto al fine di una possibile applicazione dell'intelligenza artificiale.

L'utilizzo degli algoritmi predittivi, volti a calcolare il rischio di recidiva, trova il suo caso più emblematico negli Stati Uniti, con lo strumento COMPAS, il quale è stato davvero innovativo da questo punto di vista. Esso funziona tramite delle domande poste al soggetto: sulla sua storia precedente, su possibili condanne a suo carico, se sia stato in prigione, se vi siano stati precedenti per consumo di droga o alcol, se abbia cambiato domicilio (la qual cosa non dovrebbe essere nemmeno ritenuta come un fattore rilevante), quale sia la situazione familiare e se qualche membro della sua famiglia abbia commesso qualche tipo di reato in passato, oltre che il livello di criminalità nella zona (come se una persona fosse indotta a delinquere solo perché circondata da chi lo ha già fatto), inoltre, viene richiesto anche il livello di studi (domanda ritenuta classista) e la situazione emozionale (se sia una persona normalmente triste o distratta, ecc...). Le domande che vengono poste, insomma, tendono a versare particolarmente sul privato, oltre che a richiedere spesso informazioni che possano evidenziare un contenuto classista o discriminatorio, in particolare, una delle domande più frequenti che viene posta è la "propensione ideologica al reato", la domanda rivolta, con precisione, è: "Lei pensa che il numero dei reati in una determinata città ha a che fare con il livello economico del numero delle persone che vivono lì?". Se la persona in questione risponde "sì", affermando che la povertà ovviamente può provocare delinquenza, questa risposta può essere ritenuta come una risposta classista e portare la persona a divenire sospettata[207].

L'algoritmo elabora i dati ottenuti dal fascicolo dell'imputato e dalle risposte fornite nel colloquio con lo stesso. Per quanto riguarda la valutazione del rischio, l'elaborato consiste in un grafico di tre barre che rappresentano in una scala da 1 a 10 il rischio di recidiva preprocessuale, il rischio di recidiva generale ed il rischio di recidiva violenta, in modo da assegnare un punteggio alla possibilità che individui con una storia criminosa simile siano più o meno propensi a commettere un nuovo reato una volta tornati in libertà. Difatti, bisogna considerare che COMPAS non prevede il rischio di recidiva individuale dell'imputato, bensì elabora la previsione comparando le informazioni ottenute dal singolo con quelle relative ad un gruppo di individui con caratteristiche assimilabili[208].

[207] Convegno tenutosi presso l'Università degli Studi di Salerno, il quale ha visto come ospite il professore Jordi Nieva-Fenoll, per la presentazione del suo libro "Intelligenza artificiale e processo", 7 Giugno 2019

[208] S. CARRER, Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin, in Giurisprudenza Penale Web, 2019, 4

In una discussa sentenza del 2016[209], la Corte Suprema del Wisconsin si è pronunciata sull'appello del sig. Eric L. Loomis, la cui pena a sei anni di reclusione era stata sancita sulla base di un calcolo effettuato dal sistema di intelligenza artificiale denominato COMPAS[210].

Inizialmente gli strumenti di valutazione del rischio venivano utilizzati solo dai dipartimenti di libertà vigilata per aiutare a determinare le migliori strategie di supervisione e trattamento per i trasgressori, focalizzandosi a livello nazionale sulla necessità di ridurre la recidiva e l'importanza delle pratiche basate sull'evidenza[211], l'uso di tali strumenti si è ora esteso alla condanna[212].

Passiamo alla descrizione del fatto:

Nel 2013 Loomis era alla guida di un'automobile precedentemente usata per una sparatoria nello stato del Wisconsin, USA. Il soggetto viene fermato dalla polizia, e gli vengono addebitati cinque capi d'accusa: 1) messa in pericolo della pubblica sicurezza, 2) tentativo di fuga od elusione di un ufficiale del traffico, 3) guida di un veicolo senza consenso del proprietario, 4) possesso di arma da fuoco da parte di un pregiudicato, 5) possesso di fucile a canna corta o pistola[213].

Si comincia con l'avvisare il signor Loomis, prendendo come modello una precedente causa (State vs. Frey)[214], su cosa consista la procedura di read- in, la quale espone il soggetto ad una pena più elevata nel raggio di condanna per i reati che vengono "letti" in udienza.

Quindi, viene chiesto al signor Loomis se acconsente definitivamente affinché le accuse siano lette in udienza e considerate, perché in tal caso si esporrà alla possibilità di essere condannato al massimo della pena per le accuse di cui si dichiari colpevole.

---

[209] Supreme Court of Wisconsin, State of Wisconsin v. Eric L. Loomis, Case no. 2015AP157-CR, 5 April – 13 July 2016

[210] S. CARRER, Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin, in Giurisprudenza Penale Web, 2019, 4

[211] Tale pratica richiede l'uso delle migliori prove scientifiche disponibili e dei dati raccolti sistematicamente, se disponibili, dalle legislature come base per la loro formulazione e scrittura della legge.

[212] Pamela M. CASEY,Roger K. WARREN, Jennifer K. ELEK., Using Offender Risk and Needs Assessment Information at Sentencing: Guidance for Courts from a National Working Group, at 1 (2011), http://www.ncsc.org/~/media/Microsites/Files/CSI/RNA%20Guide%20Final.ashx.

[213] S. CARRER, Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin, in Giurisprudenza Penale Web, 2019, 4

[214] State v. Michael L. Frey, 2012 WI 99. Questa sentenza aiuta a capire in cosa consista la procedura di read-in, in quanto le accuse ammesse hanno un differente significato da quelle non ammesse. Dovrebbe spettare in ogni caso al procuratore distrettuale decidere se perseguire o no una denuncia respinta

Il signor Loomis acconsente, dichiarandosi colpevole di due delle accuse meno gravi: "guida di un autoveicolo senza il consenso del proprietario" e "fuga da un ufficiale del traffico".

Loomis, quindi, negava la partecipazione alla sparatoria, ma ammetteva di aver guidato la stessa macchina coinvolta, più tardi, quella sera[215].

In seguito all'ammissione di colpevolezza, nella preparazione della sentenza, il tribunale circondariale ordinava un Presentence Investigation Report (PSI), ossia una relazione dei risultati delle investigazioni condotte sulla storia personale dell'imputato, preliminare alla sentenza sulla determinazione della pena, finalizzate a verificare la presenza di circostanze utili a modulare la severità della stessa. Il PSI includeva i risultati sulla valutazione del rischio elaborati dal software COMPAS[216].

I risultati di COMPAS attestavano un alto rischio di recidiva in tutti e tre gli ambiti di recidiva (recidiva generale, recidiva preprocessuale e recidiva violenta), affermando che Loomis potesse essere un individuo potenzialmente pericoloso per la comunità. A fronte di queste analisi, il tribunale circondariale, tenendo conto di tali fattori decideva, quindi, di non concedere la libertà vigilata, condannando Loomis a sei anni di reclusione e cinque anni di supervisione estesa.

Loomis ha presentato una richiesta di risarcimento post-condanna in tribunale e di essere risentito in una nuova udienza, sostenendo che la dipendenza del tribunale da COMPAS avesse violato i suoi diritti processuali, nonché i principi del giusto processo, appellandosi, in seguito, alla sentenza del tribunale, il quale negò la sua richiesta. Inoltre, Loomis affermava che il tribunale avesse erroneamente esercitato la propria discrezione, supponendo che le basi fattuali per le accuse delle quali non si fosse dichiarato colpevole, fossero vere, inoltre, lamentava, come già accennato, la lesione dei suoi diritti ad un giusto processo, in particolare: 1) il diritto ad essere condannato ad una determinata pena sulla base di informazioni accurate, delle quali non si poteva disporre in quanto coperte da diritti di proprietà industriale; 2) il diritto di essere condannato ad una pena individualizzata; 3) l'uso improprio del dato di genere nella determinazione della pena[217].

Il tribunale ha tenuto due udienze, sulla questione, in seguito alle richieste del signor Loomis:

[215] https://caselaw.findlaw.com/wi-supreme-court/1742124.html#footnote_12

[216] Harvard Law Review, https://harvardlawreview.org/2017/03/state-v-loomis/

[217] S. CARRER, Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin, in Giurisprudenza Penale Web, 2019, 4

- Durante la prima udienza, il tribunale ha affrontato l'affermazione di Loomis circa l'erroneo esercizio della propria discrezionalità e l'erronea valutazione delle accuse lette in udienza, come ha considerato le accuse di lettura. Considerando la giurisprudenza e le norme giuridiche pertinenti, il tribunale circondariale ha concluso di aver applicato le norme in modo corretto e ha respinto la mozione di Loomis su tale questione, in quanto la spiegazione ricavata dal signor Loomis era incompatibile con i fatti, mentre quella dello Stato era più coerente. Inoltre, sempre richiamando la sopracitata sentenza State v. Frey, un esercizio errato di discrezionalità si verifica quando un tribunale circoscrizionale impone una sentenza "senza le basi di un processo di ragionamento giudiziario motivato", oltre al caso in cui la decisione non sia supportata da fatti registrati o avvenga in seguito ad una erronea applicazione della legge (in caso di fattori irrilevanti o impropri).
- Nella seconda udienza, il signor Loomis chiama a testimoniare un esperto, il dott. David Thompson, in quanto si riteneva che l'utilizzo dello strumento COMPAS non fosse adeguato ad avere valenza nelle decisioni di incarcerazione, in quanto, i dati analizzati dall'algoritmo non sono in grado di stimare efficacemente il rischio di recidiva, rischiando di condannare un soggetto sulla base di fattori ininfluenti. Il dott. Thompson continua affermando che la Corte non sa come i dati ricevuti siano analizzati dall'algoritmo, in quanto quest'ultimo deve comparare le informazioni ottenute sull'individuo in questione e le informazioni su soggetti involti in questioni simili, però non si sa se gli individui in questione facciano parte della popolazione del Wisconsin, o di New York, o della California. Le informazioni ricavate potrebbero essere erronee o fuorvianti, il che andrebbe a pregiudicare l'intero risultato.

Il tribunale, rigettando la mozione, afferma che, sebbene la decisione si sia basata anche sui dati verificati e ricavati dallo strumento COMPAS, questi non fossero determinanti, essendo stati utili solo a confermare un risultato già ottenuto dalla valutazione di altri elementi e che la decisione sarebbe stata quella indipendentemente dal risultato ottenuto dalla macchina. Quindi, i risultati dell'algoritmo non sono stati altro che un'ulteriore conferma.

3.1. La sentenza della Corte Suprema

La questione viene portata, dunque, all'attenzione della Corte Suprema, la quale analizza, in particolare:

- Per quanto riguarda il punto (a), ovvero l'impossibilità, a causa dei diritti di proprietà industriale e la copertura del brevetto, di confutare scientificamente l'uso dei dati

processati da COMPAS, poiché essi avrebbero impedito alla difesa di accedere alle informazioni ed ai meccanismi di rielaborazione attuati dal software, la Corte ha affermato che Loomis aveva comunque la possibilità di contestare i risultati finali di calcolo del rischio. Nonostante i processi di funzionamento rimangano segreti, il manuale di COMPAS spiega che i punteggi sono basati in gran parte su dati statistici, di cui Loomis poteva verificare l'accuratezza senza alcun problema.

- Per quanto riguarda il punto (b), ovvero al fatto che COMPAS non sia in grado di individuare un particolare individuo, ma solo determinati gruppi, con la conseguenza che un imputato incensurato potrebbe essere marcato dal software come possibile futuro criminale sulla base della propria condizione sociale e di altri fattori, la Corte ha indicato che in tal caso funge da elemento determinante il fattore discrezionalità, che i tribunali circondariali devono esercitare, tenendo conto dei risultati COMPAS con riguardo peculiare ad ogni specifico individuo, considerando anche tutti gli altri fattori a disposizione.
- Infine, per quanto riguarda il punto (c), ovvero al rischio che COMPAS attribuisca importanza sproporzionata ad alcuni fattori, come l'appartenenza ad un determinato gruppo etnico, o il livello di educazione o la situazione familiare, la Corte ha ribadito che per garantirne l'accuratezza della valutazione, COMPAS debba essere costantemente monitorato e aggiornato sulla base dei cambiamenti sociali, nonché usato con accortezza e seguendo specifiche cautele[218].

In sostanza, alla luce delle considerazioni svolta, la Corte ha concluso che l'uso di COMPAS non aveva violato il diritto di Loomis all'equo processo, confermando la pena di sei anni di reclusione.

Inoltre, l'utilizzo di strumenti per la valutazione dei rischi era stato già ampiamente utilizzato da varie giurisdizioni all'interno dello stato, in particolare viene fatto riferimento al caso State v. Samsa[219], il quale va a sottolineare il potere di discrezione del

[218] S. CARRER, Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin, in Giurisprudenza Penale Web, 2019, 4

[219] State v. Samsa, 2015 WI App 6, 359 Wis.2d 580, 859 N.W.2d 149. Jordan Samsa impugnava una sentenza di condanna per aggressione sessuale di terzo grado e un ordine che nega la sua mozione post-condanna. Samsa cerca di modificare la sentenza o di essere risentito in udienza. Egli sostiene che il tribunale di circuito ha interpretato e applicato erroneamente una valutazione criminogena sulla base dei risultati elaborati dallo strumento COMPAS e che il tribunale avesse erroneamente dedotto che era più pericoloso di quanto indicato dalla valutazione del rischio COMPAS. Le sue argomentazioni erano state respinte in quanto la decisione si basava sulla stessa testimonianza della vittima, una ragazzina di 14 anni che aveva affermato di essere stata aggredita sessualmente da Samsa, oltre al fatto che il soggetto era stato precedentemente citato per possesso di mariujana e altri oggetti. La corte ha anche notato Samsa aveva alcune lievi limitazioni cognitive, potenzialmente derivanti dall'esposizione al piombo. La corte ha concordato con lo scrittore PSI e lo Stato che Samsa non si è assunto la responsabilità delle sue azioni. Il sistema COMPAS, dunque, è semplicemente uno strumento a disposizione di un tribunale al momento della

tribunale, il quale può anche decidere di discostarsi dalle valutazioni effettuate dal COMPAS.

3.2. Valutazioni complessive sull'utilizzo dello strumento COMPAS

La Corte Suprema, in seguito alla affermazione dell'utilizzo di COMPAS per la determinazione della pena, ne stabilisce anche i limiti, in quanto, tali strumenti possono essere considerati fattori rilevanti in questioni quali:

1) la comminazione di misure alternative alla detenzione per gli individui a basso rischio di recidiva;

2) la valutazione della possibilità di controllare un criminale in modo sicuro all'interno della società;

3) l'imposizione di termini e condizioni per la libertà vigilata, la supervisione e per le eventuali sanzioni alle violazioni delle regole previste dai regimi alternativi alla detenzione.

Dunque, la Corte ha confermato che lo scopo di COMPAS è quello di individuare le esigenze del soggetto che deve scontare la pena e di valutare il rischio di reiterazione del reato[220].

3.3. Analisi sul possibile esito discriminatorio di COMPAS

Sebbene le valutazioni effettuate sullo strumento COMPAS fossero per lo più positive, uno studio pubblicato da ProPublica, una ONG statunitense con sede a Manhattan, che mira a produrre giornalismo investigativo di interesse pubblico, ha constatato che i risultati prodotti da COMPAS tendono spesso a ricondurre nella fascia più elevata di rischio soggetti di colore o appartenenti a determinate zone, risultando discriminatori e non imparziali[221].

Un esempio pratico fu dato dal caso di Brisha Borden, una ragazza di 18 anni che, in ritardo per prendere sua sorella a scuola, ruba una bicicletta ad un ragazzino di 6 anni.

La ragazza fu presa e accusata di furto con scasso e piccoli furti di oggetti, che furono valutati per un totale di $ 80.

sentenza e un tribunale è libero di fare affidamento su parti della valutazione respingendo altre parti, in forza del suo potere discrezionale.

[220] http://www.giurisprudenzapenale.com/2019/04/24/lamicus-curiae-un-algoritmo-chiacchierato- caso-loomis-alla-corte-suprema-del-wisconsin/

[221] https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

L'estate precedente, il 41enne Vernon Prater fu arrestato per furto di strumenti per un valore di 86,35. Prater, però, era già stato condannato per rapina a mano armata, per la quale aveva scontato cinque anni di carcere, oltre il fatto di essere stato accusato per un'altra rapina a mano armata. L'algoritmo per la valutazione dei rischi aveva calcolato per Borden, la quale era nera, un rischio di recidiva maggiore rispetto a Prater, il quale era bianco. Due anni dopo, Borden non è stata accusata di alcun nuovo crimine, mentre Praterha scontato una pena detentiva di otto anni per essere successivamente entrato in un magazzino e aver rubato migliaia di dollari di elettronica[222].

L'analisi del codice sorgente aveva fatto stabilire alla Corte Suprema che COMPAS fosse costituzionale, tuttavia, dall'analisi è anche risultato che sia altamente improbabile che l'esame del codice riveli una discriminazione esplicita.

Difatti, l'analisi dettagliata di ProPublica dello strumento COMPAS ha rivelato che l'algoritmo non è completamente neutrale rispetto alla razza. Si prevede spesso che gli imputati neri abbiano un rischio di recidiva più elevato di quanto non sia effettivamente, e che gli imputati bianchi abbiano spesso un rischio più basso. Sebbene la macchina, infatti, non possa avere tendenze razziste, ed essere quindi garante, in un certo senso, di imparzialità e oggettività, è pur vero che gli algoritmi di apprendimento automatico, sono semplicemente riflessi dei dati input che li addestrano. E se quei dati di input sono veramente rappresentativi del nostro mondo distorto, anche il computer sarà distorto. I computer sono intelligenti, ma non sono saggi[223].

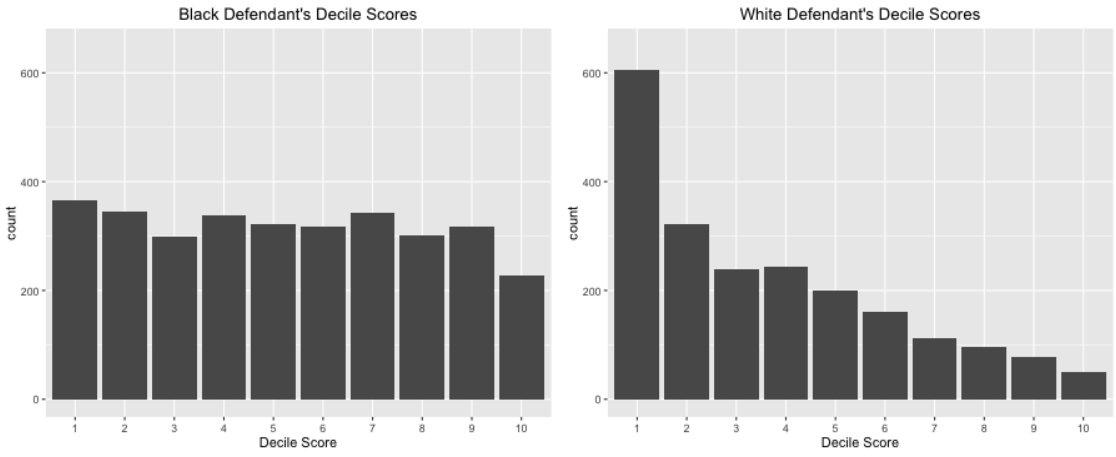
A dimostrazione di quanto affermato, la ProPublica ha sviluppato un grafico che dimostra quanto il fattore della razza influisca sulle scelte e i risultati di COMPAS, in quanto spesso altri fattori che potrebbero avere la medesima rilevanza, non vengono analizzati:

Grafico n **1**- analisi dei fattori razziali nelle valutazioni dello strumento COMPAS[224]

[222] https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

[223] 147http://jolt.law.harvard.edu/digest/algorithmic-due-process-mistaken-accountability-and- attribution-in-state-v-loomis-1

[224] Fonte: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm



Da questo grafico si può chiaramente evincere che i punteggi per gli imputati bianchi siano concentrati verso le categorie a basso rischio, mentre per gli imputati neri i rischi siano equamente distribuiti.

Sebbene vi sia una chiara differenza tra le distribuzioni dei punteggi COMPAS per gli imputati bianchi e neri, la semplice osservazione delle distribuzioni non tiene conto di altri fattori demografici e comportamentali. Per testare le disparità razziali nel punteggio che controlla altri fattori, è stato creato un modello di regressione logistica che ha considerato la razza, l'età, la storia criminale, la recidiva futura, il grado di carica, il genere e l'età.

Il nuovo modello logistico ha scoperto che il fattore maggiormente predittivo di un punteggio più alto di rischio era l'età: gli imputati di età inferiore ai 25 anni avevano 2,5 volte più probabilità di ottenere un punteggio più alto rispetto ai delinquenti di mezza età, anche quando il controllo per i crimini precedenti, criminalità futura, razza e genere (quindi sulla base di altri fattori).

Sorprendentemente, dati i loro bassi livelli di criminalità in generale, le donne imputate erano 19,4 % più probabilità di ottenere un punteggio più alto rispetto agli uomini[225].

In particolare, l'analisi di ProPublica ha rilevato che:

• Gli imputati neri erano spesso previsti a maggior rischio di recidiva di quanto non fossero in realtà. Gli imputati neri che non avessero ricommesso un reato per un periodo di oltre due anni erano stati ritenuti avere il doppio delle probabilità di essere classificati erroneamente con il rischio più elevato rispetto ai loro omologhi bianchi (45 % contro 23 %).

[225] Per maggiori informazioni sui risultati di tale analisi, si consiglia di leggere l'articolo di J. Larson, S. Mattu, L. Kirchner e J. Angwin, How We Analyzed the COMPAS Recidivism Algorithm, May 23, 2016.

• Gli imputati bianchi erano spesso previsti meno rischiosi di quanto non fossero. Gli imputati bianchi che hanno recidivato nei due anni seguenti erano stati erroneamente etichettati a basso rischio quasi due volte più spesso dei rei offensori neri (48 % contro 28%).
• L'analisi ha anche mostrato che anche quando si controllano reati precedenti, rischi recidiva futura, età e genere, gli imputati neri hanno il 45% in più di probabilità di ottenere punteggi di rischio più alti rispetto agli imputati bianchi.
• Gli imputati neri avevano anche il doppio delle probabilità che gli imputati bianchi venissero classificati avere erroneamente un rischio maggiore di recidiva violenta. E i recidivi violenti bianchi avevano il 63 % in più di probabilità di essere stati classificati erroneamente come un basso rischio di recidiva violenta, rispetto ai recidivi violenti neri.
• L'analisi della recidiva violenta ha anche mostrato che anche quando controllano i precedenti crimini, la recidiva futura, l'età e il genere, agli imputati neri è stato assegnato il 77 % in più di probabilità di ottenere punteggi di rischio più alti rispetto agli imputati bianchi[226].
Termino questa osservazione con una frase pronunciata da Antoine Garapon[227]: "COMPAS proietta sul futuro le accuse all'imputato, riproducendo i pregiudizi e aumentando la stigmatizzazione razziale della società americana. Ciò si spiega scientificamente, poiché gli algoritmi, nella maggior parte dei casi, prevedono il futuro come se fosse la continuazione del passato".

4. Attività del giudice: uomo o robot?

Ancora molti sono i dubbi e gli ostacoli che si pongono alla completa affermazione e accettazione di un giudice robotico, ritenendo che l'attività compiuta dal giudice possa difficilmente trovare attuazione da parte di una macchina, che, in quanto tale, non è pensante e non può discernere ciò che è giusto da ciò che non lo è.
Lo stesso articolo 192 del codice di procedura penale, descrive il procedimento di valutazione della prova, basata sul libero convincimento del giudice, che però: "valuta la prova dando conto nella motivazione dei risultati acquisiti e dei criteri adottati".

[226] https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm
[227] A cura di E. FRONZA e C. CARUSO, Ti faresti giudicare da un algoritmo? Intervista ad Antoine Garapon, Rivista trimestrale, Fascicolo 4/2018

Il giudice, quindi, ha l'obbligo di motivare la sua decisione, ricostruendo il percorso logico-conoscitivo che lo ha condotto a valutare in un determinato modo gli elementi acquisiti.

In questo caso, un giudice automatico avrebbe bisogno di assegnare un valore numerico definito (un punteggio di efficacia dimostrativa) ad ogni elemento di prova, a favore o contro, reintroducendo necessariamente un sistema di prove legali simile a quello adottato nei sistemi giudiziari dell'ancien régime, il che ovviamente collliderebbe con i principi del nostro ordinamento. Infatti, la macchina, si trova ad affrontare un tipo di giudizio non strettamente deterministico (forse troppo umano) che la metterebbe in serie difficoltà[228].

Gli ostacoli all'operare di un giudice robotico non si limitano a questo, anzi, sono molteplici, ma numerose potrebbero esserne anche le soluzioni.

4.1. Sussunzione delle fattispecie

La parte preponderante dell'attività del giudice è la sussunzione del fatto nella norma, la riconduzione della fattispecie concreta alla fattispecie astratta, quindi sono fondamentali i fatti che al giudice sono "rappresentati". Come è evidente, questi fatti sono l'inevitabile risultato di una selezione, essendo impossibile rappresentare al giudice l'universalità dei fatti che egli dovrà valutare. Quindi, i fatti rilevanti non sono altro che il frutto di una selezione della molteplicità dei dati materiali disponibili, che trova il proprio fondamento nell'identificazione della norma che va loro applicata[229]. L'attività del giudice, quindi, si pone come uno strumento che adopera una catena di sillogismi[230], che si basano sulla individuazione della norma astratta e dal riscontro che si siano verificati casi concreti previsti dalla norma, sulla base dello schema dei postulati enunciati da Weber[231]. Questo processo non muterebbe nel caso si decidesse di "sostituire" il giudice umano con un giudice robot, in quanto, anche esso dovrebbe effettuare una cernita dei dati da analizzare,

[228] https://www.4clegal.com/hot-topic/giudice-automatico-giudizio-penale-intelligenza-artificiale

[229] M. LUCIANI, La decisione giudiziaria robotica, Rivista N°: 3/2018, data pubblicazione: 30/09/2018

[230] Sulle decisioni giudiziali ritenute come catene di sillogismi P. Calamandrei,in Opere giuridiche, vol. 1, Morano, 1965, pag 15 scrive: "Il giudice deve fare prima di giungere alla pronuncia della sentenza una serie complicata di deduzioni concatenate, è da notare che quando si raffigura la sentenza come un sillogismo, si ha di mira soltanto il momento finale dell'attività del giudice, in cui, avendo questi preparato ormai dinanzi a sé la norma di legge da applicare e il giudizio sui caratteri giuridici del fatto accertato, più non gli resta che trarre a rigor di logica la conseguenza di queste due premesse".

[231] M. Weber, Economia e società, vol. III, Edizioni di Comunità, Milano, 1980. 1) la giurisdizione è applicazione di uno schema astratto a una fattispecie concreta; 2) per ogni fattispecie concreta deve essere sempre ricavabile una decisione dal sistema delle fattispecie astratte; 3) l'ordinamento giuridico deve essere, o deve venir trattato come se fosse, privo di lacune; 4) ciò che non è sussumibile nella forma astratta è anche irrilevante per il diritto; 5) ogni agire umano deve essere sussumibile in una regola precostituita.

una cernita effettuata nel momento in cui è necessaria una semplificazione che riconduca il fatto ad un dato numerico, facilmente leggibile dalla macchina[232].

4.2. Il giudice Robot e il rapporto con i precedenti

Mentre la sussunzione della fattispecie concreta nella fattispecie astratta è attività di tipo sillogistico, non lo è la ricostruzione della fattispecie astratta e spesso viene analizzata la possibilità di dare istruzioni alla macchina al fine di formulare proposizioni allo stesso modo in cui lo fanno i giudici (umani).

Spesso, però, questo lavoro viene tacciato di poca trasparenza, e molti affermano che l'istruzione del robot spesso si riduca ad una decisione politica.

Una soluzione a tale inconveniente potrebbe essere quella di istruire il robot inserendo nella sua memoria tutti i possibili repertori giurisprudenziali, assicurando un dato oggettivo, ma questo comporterebbe alcuni problemi, quali:

- Violazione articolo 101 Cost. (subordinazione del giudice soltanto alla legge), in quanto il giudice sarebbe in tal caso subordinato alla giurisprudenza.
- Possibili conflitti di istruzioni e contrasti giurisprudenziali, relativi alla decisione su quale tipo di repertorio giurisprudenziale considerare.
- Effetto conservatore (di cui già accennato), relativo all'analisi di situazioni precedenti precludenti le evoluzioni e gli sviluppi futuri. In un'intervista rilasciata da Pasquale Liccardo[233], si è fatto riferimento al fatto che il rapporto tra presente e passato sia un elemento fondamentale per il sistema giudiziario: la scienza giuridica guarda al passato per leggere il presente, essendo il precedente un elemento fondamentale e giustificatore del dictum giudiziario.

L'utilizzo delle tecnologie, secondo lui, potrebbe aiutare e collaborare come raccordo tra norma astratta e fatto concreto, tramite rappresentazioni cognitive e processi di accumulazione di dati, dando un oggettivo apporto che permetta tramite l'utilizzo dei big data e di analisi numeriche, di sopperire alle equivocità delle statistiche giudiziarie.

4.3. Errori robotici: un problema di responsabilità giuridica

In uno scenario dove l'IA e la robotica diventano sempre più rilevanti ed evolute c'è da chiedersi come si evolverà il concetto di responsabilità.

[232] M. LUCIANI, La decisione giudiziaria robotica, Rivista N°: 3/2018, data pubblicazione: 30/09/2018
[233] Presidente del Tribunale di Modena, C.Castelli, D.Piana, Giusto processo e intelligenza artificiale, Maggioli, Santarcangelo 2019.

La questione può essere analizzata sotto due diversi punti di vista: da una parte considerando l'IA come un nuovo mondo da regolare, dall'altra come uni strumento che incide sulla modalità del diritto di dare regole.

Uno degli aspetti più interessanti rimane legato alla "soggettività giuridica" dei sistemi intelligenti. Sappiamo che il soggetto di diritto è un portatore di interessi giuridicamente tutelati, centro di imputazione di situazioni soggettive. Quando parliamo di una macchina, possiamo parlare di un soggetto di diritto? Sembra che per alcuni paesi la risposta a tale quesito possa essere affermativa: basti pensare che l'Arabia Saudita ha addirittura concesso la cittadinanza a Sophia, un evoluto robot umanoide prodotto dall'azienda americana Hanson Robotics. L'installazione di due microcamere negli occhi le permette di vedere il suo interlocutore, di percepirne lo stato d'animo, riuscendo ad intrattenere con esso anche dei dialoghi, imitando perfettamente i gesti umani e le espressioni facciali, migliorando di giorno in giorno la qualità della conversazione, grazie al ricordo dei dialoghi precedenti[234].

Il problema della possibile configurabilità di soggetti robotici quali attori criminosi potrebbe essere così risolto: machina delinquere (et puniri) non potest. Potremmo riassumere in tal modo un dogma inespresso del diritto penale che fino ad oggi abbiamo conosciuto: macchine, robot e soggetti artificiali di ogni genere non possono essere direttamente responsabili per la commissione di reati[235].

Ci troviamo di fronte al recupero in chiave tecnocratica di una teocratica "Giustizia Eterna"? O forse solo a un alibi de-responsabilizzante?[236]

In questo scenario, osserva il Parlamento, lo sviluppo della tecnologia robotica dovrebbe mirare a integrare le capacità umane, e non a sostituirle, ritenendo fondamentale, nello sviluppo della robotica e dell'intelligenza artificiale, garantire che gli uomini mantengano in qualsiasi momento il controllo sulle macchine intelligenti e sottolineando altresì il pericolo che nasca un attaccamento emotivo tra gli uomini e i robot, in particolare per i gruppi vulnerabili (bambini, anziani e disabili), con il conseguente grave impatto emotivo e fisico che un tale attaccamento potrebbe avere sugli uomini[237].

Il Parlamento è quindi consapevole che «è necessaria una serie di norme che disciplinino in particolare la responsabilità civile per i danni causati da robot», che è una questione

[234] http://www.notaiocamilloverde.it/mc/475/intelligenza-artificiale-e-diritto-tra-etica-delle- robotica-e-tutele-giuridiche

[235] https://discrimen.it/wp-content/uploads/Cappellini-Machina-delinquere-non-potest.pdf

[236] C. CASTELLI, Giusto processo e intelligenza artificiale, Maggioli, Santarcangelo di Romagna 2019.

[237] http://questionegiustizia.it/stampa.php?id=1658

fondamentale, ed invita ad adottare sia uno strumento legislativo, che strumenti para-legislativi, quali linee guida e codici di condotta.

In altri termini, si domanda se, in base all'attuale quadro giuridico, la responsabilità da prodotto (secondo la quale il produttore di un prodotto è responsabile dei malfunzionamenti) e le norme che disciplinano la responsabilità per azioni dannose (in virtù delle quali l'utente di un prodotto è responsabile di un comportamento che conduce al danno) siano applicabili ai danni causati dai robot e dall'intelligenza artificiale, ponendosi poi il problema, nel caso i robot acquisissero la capacità di assumere decisioni autonome, se siano sufficienti le norme tradizionali.

Nel caso in cui un soggetto artificiale commetta materialmente, di propria mano, un fatto qualificabile come reato, si potrà parlare di un modello imputativo più tradizionale e familiare al penalista, ovvero di una responsabilità vicaria dell'uomo. In esso, la macchina è da considerarsi come un mero strumento nelle mani del vero autore alle sue spalle, che la utilizzi direttamente o la predisponga ad atteggiarsi in base ai propri fini[238]. Proprio uno dei Considerando (considerando "AG")[239] della Risoluzione del Parlamento europeo del 16 febbraio 2017, fa emergere come necessaria la predisposizione di apposite norme che disciplinino le caratteristiche della responsabilità e i criteri di assunzione della stessa da parte di un robot agente, nel rispetto di valori universali e umanistici del tutto conformi, oltretutto, a principi intrinsecamente europei (considerando "U")[240].

Il Parlamento osserva, inoltre, che la direttiva 85/374/CEE (Considerando AH), 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, ma questo potrebbe non essere sufficiente a coprire i danni causati dalla nuova generazione di robot, in quanto questi possono essere dotati di capacità di adattamento e di apprendimento che implicano un certo grado di imprevedibilità nel loro comportamento, dato che «imparerebbero in modo autonomo, in

[238] https://discrimen.it/wp-content/uploads/Cappellini-Machina-delinquere-non-potest.pdf

[239] Considerando AG: "considerando che sono palesi le carenze dell'attuale quadro normativo anche in materia di responsabilità contrattuale, dal momento che le macchine progettate per scegliere le loro controparti, negoziare termini contrattuali, concludere contratti e decidere se e come attuarli rendono inapplicabili le norme tradizionali;considerando che ciò pone in evidenza la necessità di norme nuove, efficaci e al passo con i tempi che corrispondano alle innovazioni e agli sviluppi tecnologici che sono stati di recente introdotti e che sono attualmente utilizzati sul mercato".

[240] Considerando U: considerando che è necessaria una serie di norme che disciplinino in particolare la responsabilità, la trasparenza e l'assunzione di responsabilità e che riflettano i valori intrinsecamente europei, universali e umanistici che caratterizzano il contributo dell'Europa alla società; che tali regole non devono influenzare il processo di ricerca, innovazione e sviluppo nel settore della robotica;

base alle esperienze diversificate di ciascuno», e interagirebbero con l'ambiente in modo unico e imprevedibile.

Sul punto ci si permette di essere critici e di sottolineare come tale imprevedibilità faccia parte di una "imprevedibilità prevedibile", in quanto anche nella creazione dell'algoritmo più sofisticato, l'ideatore, in base a processi matematici, ha un range di prevedibilità dell'operato del robot, in quanto, se è vero che gli input non vengono tutti inseriti al momento della creazione, ma vengono acquisiti anche durante la "vita" del robot (cd. "autoapprendimento"), il processo matematico che collega gli input, anche se via via acquisiti durante il percorso di attivazione del robot, è comunque sempre immesso dal fattore umano, dall'ideatore[241].

In molti casi, addirittura, la IA è in grado di apprendere non solo dalla propria esperienza, ma anche da quella dei suoi "simili", tramite il ricorso a tecnologie di cloud computing[242], che permettono di collegare tra loro più soggetti artificiali e sommarne le "esperienze di vita". L'effetto ottenuto è che in tal modo la rapidità di apprendimento dei soggetti robotici è di gran lunga aumentata, diminuendo, in parallelo, il grado di controllo che l'uomo può esercitare nei loro confronti.

Proseguendo nell'analisi, il Parlamento evidenzia altri problemi aperti, quali ad esempio, le carenze dell'attuale quadro normativo anche in materia di responsabilità contrattuale, dal momento che le macchine progettate per scegliere le loro controparti, negoziare termini contrattuali, concludere contratti e decidere se e come attuarli, rendono inapplicabili le norme tradizionali in materia di responsabilità contrattuale, con necessità di nuove norme.

Il Parlamento formula alcune proposte e suggerimenti per la redazione della successiva normazione da parte della Commissione, ricordando, in primo luogo, che nella scelta tra un regime di responsabilità oggettiva e un approccio della gestione dei rischi, a favore della scelta per un regime di responsabilità oggettiva, milita il fatto che per tale tipo di responsabilità viene richiesto al danneggiato solo la prova del danno avvenuto e l'individuazione di un nesso di causalità tra il funzionamento lesivo del robot e il danno subito dalla parte, ma che a favore dell'approccio di gestione dei rischi, milita il fatto che la responsabilità non si concentra sulla persona "che ha agito con negligenza" in quanto

[241] http://questionegiustizia.it/articolo/iot-e-intelligenza-artificiale-le-nuove-frontiere-della- responsabilita-civile-e-del-risarcimento_13-06-2018.php

[242] http://www.treccani.it/enciclopedia/cloud-computing/Letteralmente "nuvola informatica", termine con cui ci si riferisce alla tecnologia che permette di elaborare e archiviare dati in rete. In altre parole, attraverso internet il c.c. consente l'accesso ad applicazioni e dati memorizzati su un hardware remoto invece che sulla workstation locale.

responsabile a livello individuale, bensì sulla persona che, in determinate circostanze, è in grado di minimizzare i rischi e affrontare l'impatto negativo.

Le indicazioni fornite dal Parlamento riguardano:

1) I legittimati passivi. La loro responsabilità dovrebbe essere

proporzionale all'effettivo livello di istruzioni impartite al robot e al grado di autonomia di quest'ultimo, di modo che quanto maggiore è la capacità di apprendimento o l'autonomia di un robot e quanto maggiore è la durata della formazione di un robot, tanto maggiore dovrebbe essere la responsabilità del suo formatore; osservando peraltro, 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 «che, almeno nella fase attuale, la responsabilità deve essere imputata a un essere umano e non a un robot»

2) Regime di assicurazione obbligatorio. Per una soluzione alla complessità del problema «di chi la colpa», il Parlamento propone un regime di assicurazione obbligatorio (che tenga conto delle responsabilità non solo dell'ultimo anello della catena, ma di tutte le potenziali responsabilità lungo la catena), in virtù del quale, come avviene già per le automobili, venga imposto ai produttori e i proprietari dei robot di sottoscrivere una copertura assicurativa per i danni potenzialmente causati dai loro robot.

Nella risoluzione del Parlamento europeo del 16 febbraio 2017 viene anche auspicata «l'istituzione di uno status giuridico specifico per i robot nel lungo termine, di modo che almeno i robot autonomi più sofisticati possano essere considerati come persone elettroniche responsabili di risarcire qualsiasi danno da loro causato, nonché eventualmente il riconoscimento della personalità elettronica dei robot che prendono decisioni autonome o che interagiscono in modo indipendente con terzi».

Questa scelta potrebbe incentivare una totale deresponsabilizzazione del produttore e dell'ideatore, un pericoloso laissez faire, considerato che i produttori e ideatori agirebbero su un mercato globale, con regole non uniformi e pericolosi incastri di "scatole cinesi" a livello mondiale, che minerebbero in ultima analisi il concreto risarcimento, per problemi, non ultimi, legati all'effettiva solvibilità del debitore[243].

Per ciò che concerne la responsabilità penale, bisogna distinguere fra actus reus e mens rea.

[243] http://questionegiustizia.it/articolo/iot-e-intelligenza-artificiale-le-nuove-frontiere-della- responsabilita-civile-e-del-risarcimento_13-06-2018.php

Sicuramente una macchina è in grado di un'azione penalmente rilevante (si pensi al movimento del braccio robotico), il problema concerne la mens rea. Le intelligenze artificiali recepiscono i dati dal mondo esterno, quindi esse "conoscono" la realtà esterna, se la "rappresentano" e "prevedono" un certo risultato.

Tutto ciò farebbe pensare che il loro agire potrebbe integrare requisiti quali la "negligenza", "la conoscenza", "l'intenzione", ecc...[244]

Gabriel Hallevy[245], ha affermato che il fatto che tutti neghino la responsabilità penale delle macchine riecheggia un po' la storia del diffuso scetticismo circa la responsabilità da reato degli enti. Le corti di common law hanno dovuto affrontare un lungo cammino verso il superamento dell'assioma del societas delinquere non potest. Alcuni ritenevano la responsabilità delle società da escludere, in quanto non vi era un vero e proprio "corpo" su cui infliggere la pena. Queste teorie sono state combattute e "sconfitte", però, da esigenze di regolazione sociale.

Perché per le macchine ciò dovrebbe essere diverso? A dirla tutta, le macchine hanno in più delle società, anche un "corpo" da castigare: qui non si parla solo del corpo fisico, quale, come già accennato, il braccio robotico, ma anzi, anche i soggetti artificiali immateriali, quali i software, possono essere eliminati sia nel mondo digitale (cancellati dalla memoria in cui sono contenuti), sia nel mondo fisico (tramite la distruzione dell'hardware in cui sono scritti)[246].

In base a quanto prima affermato, sorgono 3 importanti critiche:

1) La colpevolezza: sebbene il ricevere dati dalla realtà esterna faccia presupporre una qualche "conoscenza" da parte delle macchine, e che la "previsione" di un risultato da raggiungere faccia pensare ad una "volontà" di queste ultime, tutto ciò sembra essere lontano dalla definizione vera e propria di "dolo". In tal caso il dolo in questione sarebbe solo apparente, esteriore, ma l'elemento esteriore non basta, essendo necessaria, oltre alla tipicità dolosa, anche la colpevolezza dolosa. Elemento fondamentale che caratterizza la colpevolezza dolosa è la scelta: l'uomo sceglie un comportamento antigiuridico, e con tale scelta, diviene anche colpevole. Le azioni delle macchine, sebbene dotate di tipicità dolosa, mancano della colpevolezza, e ciò è dovuto alla mancanza di capacità di autodeterminarsi al pari dell'uomo.

[244] https://discrimen.it/wp-content/uploads/Cappellini-Machina-delinquere-non-potest.pdf

[245] Professore di diritto penale presso la Facoltà di Giurisprudenza, Ono Academic College, la più grande facoltà di giurisprudenza in Israele, nel suo articolo "The Criminal Liability of Artificial Intelligence Entities".

[246] G. HALLEVY, The Criminal Liability of Artificial Intelligence Entities, in "Akron intellectual property journal", 2010.

2) Funzione della pena: una pena irrogata ad una macchina intelligente, non potrebbe mai svolgere nessuna delle classiche funzioni che la dottrina riconosce alla sanzione criminale[247]:

a) Funzione retributiva; non può essere attuata in quanto le macchine sono insuscettibili di un rimprovero colpevole.

b) Funzione rieducativa; rieducare presuppone la possibilità di apprendere dalla sanzione irrogata, ed una macchina, non può essere rieducata attraverso una pena, a meno che non la si programmi appositamente per farlo.

c) Funzione intimidatoria; questa presupporrebbe emozioni quali il timore o la paura, di cui le macchine sono prive.

3) Parallelo con la corporate liability: la corporate liability è diretta solo apparentemente all'ente, in realtà essa si rivolge a soggetti fisici, gli uomini, suoi rappresentanti, veri destinatari della pena[248].

Diverso è il caso delle intelligenze artificiali, che esistono davvero nel mondo fisico, l'uomo si limita a crearle e programmarle. Una sanzione rivolta alla IA, si ripercuoterà, in tal modo, su nessun altro all'infuori di loro medesime.

Inoltre, è da tenere in conto che il controllo dell'apparecchio, essendo quest'ultimo programmato appunto per essere autonomo, una volta perso, diventa complesso da recuperare[249].

Sebbene, dunque, lo scetticismo sulla commissione di reati da parte dei soggetti artificiali sia sempre più concreto, è innegabile l'ingrandimento dello spazio vuoto di protezione penale generato dallo sviluppo delle nuove tecnologie, rispetto a quelle offese non attribuibili né al programmatore, né alla macchina stessa. Questi ultimi fatti, sono destinati ad essere ricondotti all'interno della categoria del caso fortuito, in quanto realizzazioni di un rischio imprevedibile e ineliminabile[250].

5. L'algoritmo in campo probatorio

[247] P.M ASARO, Abody to Kick, but Still No Soul to Damn, MIT PRESS, 2011, cit.,181; Pagallo U., Killers, fridges, and slaves,cit., 167.

[248] P.M. ASARO, A body to Kick, but Still No Soul to Damn, MIT PRESS, 2011 cit.,182.

[249] https://www.altalex.com/documents/news/2019/03/12/intelligenza-artificiale-gdpr- produttore#_ftnref10Si parla in tal caso di perdita di controllo locale, se l'apparecchio non è più controllato dall'umano che lo dovrebbe fare, o generale, qualora non sia più recuperabile in toto. È chiaro che la seconda prospettiva, più plausibile vista l'autonomia con cui deve agire l'intelligenza artificiale, può portare ad un rischio pubblico enorme, nonché ad eventuali danni inestimabili, conditi con un'incertezza normativa in materia di responsabilità che aumenta il rischio di allocazione inefficiente dei costi provenienti dai danni. (Si veda Scherer, op. cit., ibidem; W. Wallach, C. Allen, Moral machines: teaching robots right from wrong, Indiana University Press, Indianapolis, 2009, pp. 197–214)

[250] https://discrimen.it/wp-content/uploads/Cappellini-Machina-delinquere-non-potest.pdf

Una prova non è altro che una serie finita di parole di un linguaggio, la quale rispetta un certo numero di regole ben stabilite.

Se si parte dal presupposto che la prova altro non sia che il frutto di caratteristiche logiche e di una contrapposizione tra due elementi quali: un rigido formalismo e pertinenza e autonomia, gli stessi elementi li ritroviamo nell'intelligenza artificiale[251]. L'utilizzo della logica come mezzo per permettere alle macchine di essere applicate al diritto può essere esteso anche al campo probatorio, in quanto anche quest'ultimo si serve sempre più spesso dell'utilizzo della logica, ergo, l'applicazione dell'intelligenza artificiale in campo probatorio è più che probabile e potrebbe ottenere risultati ottimali.

5.1. Prove paraconsistenti

La logica paraconsistente[252] è quella logica che pone una eccezione al principio di non contraddizione[253], in quanto essa afferma che, in tal caso, una proposizione A e la sua negazione -A possono essere entrambe vere.

Vi sono vari esempi di teorie probatorie a cui viene applicata la logica paraconsistente, un esempio è fornito dalla regola di deduzione naturale, la quale si ricava dall'aggiunta di una negazione forte a quella già posta dalla logica classica, dunque la logica paraconsistente non risulterà altro che un'estensione della classica, la quale può finire con l'essere tradotta in logica paraconsistente[254].

Altro esempio è fornito dal calcolo dei sequenti[255], tramite il quale partiamo da un insieme noto di connessioni premesse-conseguenze e andiamo a derivarne di nuove, partendo dalle stesse regole della logica classica ma in seguito sostituendole con regole nuove. Ciò dimostra che vi sono delle logiche in cui la logica classica può essere immersa, quindi che la logica paraconsistente può essere adattata alle estensioni della classica. L'agire della intelligenza artificiale è guidato dal "principio di contraddizione", in quanto essa si basa sulle informazioni che vengono rilevate, i quali, a loro volta, potrebbero essere contraddittorie tra di loro. Ecco che la logica consistente può essere davvero d'aiuto in tale caso: quando si presentano due informazioni contraddittorie, scegliere quale delle due sia giusta può risultare alquanto difficile,

[251] J. SALLANTIN, J. JACQUES, Il concetto di prova alla luce dell'intelligenza artificiale, Giuffrè editore (Collana "Epistemologia giudiziaria"), Milano, 2005.

[252] Termine introdotto da F.Mirò Quesada nel 1976.

[253] «È impossibile che il medesimo attributo, nel medesimo tempo, appartenga e non appartenga al medesimo oggetto e sotto il medesimo riguardo», Aristotele, Metafisica, Libro Gamma, cap. 3.

[254] J. SALLANTIN, J. JACQUES, Il concetto di prova alla luce dell'intelligenza artificiale, Milano, 2005.

[255] Sviluppato da J.-Y BEZIAU, Noveaux résultats et noveau regard sur la logique paraconsistante C1, in Logique et analyse, 1993, p.45 ss.

poiché si ritiene difficile che lo siano entrambe. Se si sceglie l'informazione sbagliata, però, questo potrebbe compromettere l'intero ragionamento e risultato. Ecco perché la scelta maggiormente ottimale è quella di analizzare simultaneamente entrambe le informazioni, anche se contraddittorie, e ricavare un risultato comune da tale analisi[256].

5.2. Il "calcolo" della prova

Negli ultimi anni, le macchine sono chiamate a manipolare un complesso sistema logistico formale come strumento per implementare la soluzione di un problema. La soluzione al problema, può essere incarnata in un teorema, ovvero, un meta-teorema generale, partendo dall'enunciato del problema, si possono ricavare gli elementi che hanno il medesimo ruolo. L'insieme di tutte le simmetrie sintattiche per un dato insieme di formule è costruito e visualizzato in una forma favorevole alla programmazione del minimo sforzo per un computer[257].

Il matematico non si limita a provare dei teoremi, ma crea nuovi concetti che potrebbero aiutarlo a creare nuove congetture, è il caso di AM e EURISKO, due sistemi di intelligenza artificiale realizzati da D. Lenat[258], capaci di creare nuovi concetti, di scoprire numerosi risultati e partire da questi per nuove congetture, Sebbene i passi in avanti siano notevoli, però i sistemi di intelligenza artificiale non riescono ancora completamente ad eguagliare il lavoro di un vero matematico, il che può risultare utile al fine dell'analisi delle prove, nonché alla loro individuazione[259].

5.3. DigitalForensics: la prova digitale

Con la digital forensics, prassi e teorie proprie della scienza forense vengono estese anche al mondo delle nuove tecnologie. Essa è una scienza che studia i dati informatici e il valore che essi possono avere in differenti ambiti sociali, giuridici o legali. La digital forensics sancisce la nascita della cosiddetta fonte di "prova digitale".

[256] J. SALLANTIN, J. JACQUES, Il concetto di prova alla luce dell'intelligenza artificiale, Milano, 2005. Un esempio è dato dal fallimento della impresa SCHLOK. Il gruppo di esperti STRATÈGE ne annuncia il fallimento, mentre il gruppo MANÈGE annuncia che l'impresa sopravviverà. I due gruppi non sono d'accordo sul futuro dell'impresa ma si accordano per affermare che non è possibile che: <<SCHLOK non fallisca>> e che << tutti gli impiegati di SCHLOK saranno licenziati>>, da ciò si può dedurre che << se SCHLOK non fallisce, alcuni suoi impiegati non saranno licenziati>>.

[257] H. GELERNTER, A Note on Syntactic Symmetry and the Manipulation of Formal Systems by Machine, in Information and Control, New York, 1959.

[258] R. DAVIS-D. LENAT, Knowledge Based Systems In Artificial Intelligence, McGraw Hill International Book Co.,New York, 1982.

[259] J. SALLANTIN, J. JACQUES, Il concetto di prova alla luce dell'intelligenza artificiale, Milano, 2005

Quest'ultima è particolarmente difficile da identificare dato il suo continuo mutare e viene raccolta tramite degli strumenti (tool) regolati da precisi metodi[260].

Il 4 dicembre 2017 è stata pubblicata sul sito della Guardia di Finanza la Circolare 1/2018 intitolata "Manuale operativo in materia di contrasto all'evasione e alle frodi fiscali", nella quale viene introdotta la figura del personale qualificato "CFDA" (Computer Forensics e Data Analysis), una novità, che riconduce ad un utile aiuto nel campo delle attività della Polizia Giudiziaria, nel caso in cui non sia stato nominato un consulente tecnico ex art 359 c.p.p. o non sia stato nominato un Ausiliario di Polizia Giudiziaria ex art 348 co 4 c.p.p.[261]

I principi giuridici della digital forensics sono stati definiti nel 2001 con la Convenzione di Budapest sul Cybercrime, recepita, in seguito, in Italia, con la Legge 18 marzo 2008, n. 48 (pubblicata in Gazzetta Ufficiale 4 aprile 2008, n. 80) promulgata dal Presidente della Repubblica. Questa reca significanti modifiche ai Codici penali e di procedura penale, introducendo sanzioni più pesanti per i reati informatici e maggiori tutele per i dati personali[262].

Nasce, dunque, anche la dizione di digital evidence, come qualsiasi dato che può stabilire un collegamento tra un crimine e la sua vittima o tra un crimine e chi lo ha commesso, questa è un'informazione trasmessa in formato binario. Dunque, lo sviluppo e il proliferare della tecnologia sta facendo sempre più parte del mondo giudiziario, fino a creare nuovi mezzi di prova che devono, però, essere sempre congrui con i principi ordinamentali, cercando di porre un equilibrio tra l'esigenza di attendibilità della prova e i principi di difesa e libertà della persona[263].

L'utilizzo di tools predittivi preoccupa principalmente per i possibili esiti discriminatori[264], eppure la loro applicazione si sta espandendo in aree sempre maggiori, soprattutto in ambito investigativo: è il caso della polizia predittiva, la quale si serve di algoritmi diretti a prevedere la consumazione dei reati, quindi a prevenirli, anche pervenendo all'arresto in flagranza di colui che sia colto nell'atto di commettere un reato[265].

---

[260] L. LUPARIA, G. ZICCARDI, Investigazione penale e tecnologia informatica, Giuffrè, Milano, 2007.

[261] https://www.ictsecuritymagazine.com/articoli/le-basi-della-digital-forensics-nella-circolare- 12018-della-guardia-finanza/

[262] https://www.altalex.com/documents/leggi/2009/10/02/criminalita-informatica-ratificata-la- convenzione-di-budapest

[263] L. LUPARIA, G. ZICCARDI, Investigazione penale e tecnologia informatica, Giuffrè, Milano, 2007

[264] Si veda il caso del tool COMPAS, in precedenza analizzato.

[265] http://www.questionegiustizia.it/articolo/intelligenza-artificiale-e-processo-penale-tra-norme- prassi-e-prospettive-_18-10-2019.php

5.4. Valutazione della prova e legame con le neuroscienze

In un sistema basato sulla prova legale[266], l'utilizzo dell'intelligenza artificiale sarebbe stato sicuramente ideale. Innegabile è invece, che in un sistema basato principalmente sulla libera valutazione del giudice, gli strumenti di IA potrebbero riscontrare più difficoltà, nonostante ciò, i dati che vengono riscontrati sono pur sempre oggettivi e questo può farci risultare utile la collaborazione di sistemi di intelligenza artificiale[267].
Gli studi in questo campo si sono avvalsi della collaborazione di rilevanti studi anche nel campo della neurobiologia e delle neuroscienze, in quanto, per poter capire come far funzionare un sistema di IA e permettergli di simulare la mente umana, prima bisogna capire come funzioni la stessa mente umana.
Esistono macchine, che per rapidità di calcolo riescono a superare la stessa mente umana nella risoluzione di problemi, ma la mente umana è molto più complessa di così, oltre i calcoli logici essa si basa anche su stati mentali quali la coscienza, l'intenzionalità, e la soggettività[268].
- Partiamo dalla testimonianza. La testimonianza non è altro che una dichiarazione effettuata da soggetti che non sono parti del processo, su fatti avvenuti in loro presenza o di cui abbiano conoscenza. Tali dichiarazioni servono allo stesso giudice al fine di ottenere informazioni valide riguardo il caso in questione, quindi quest'ultimo sarà interessato a mettere le parti nelle condizioni migliori per poter effettuare delle dichiarazioni soddisfacenti e utili al fine prestabilito.
L'algoritmo, di fatto, può effettuare le medesime attività svolte dal giudice, in quanto può scegliere e scartare le testimonianze che non si rivelano adatte alla risoluzione della causa, sulla base degli stessi input che gli vengono propinati e che insegnano che la memoria umana può essere influenzata da numerose variabili (ad esempio la testimonianza di chi ha assistito agli eventi da lontano non sarà molto affidabile, come anche chi viene minacciato con un'arma si focalizzerà più sull'arma che non sulla persona che la impugnava). Queste variabili, però, poggeranno su elementi che mancano di certezza e precisione matematica: non vi è la certezza che chi assista a un evento da lontano sia poco attendibile, o che chi è minacciato da una pistola non si

[266] Il sistema della prova legale è un'eccezione alla regola, basata sull'obbligo del giudice di conformarsi a quanto stabilito dalla legge, senza poter valutare il contenuto della prova (Art 116 co2 cpc.).
[267] J. NIEVA FENOLL, Inteligencia artificial y proceso judicial, Marcial Pons Ediciones Juridicas y Sociales,Madrid,2018
[268] A. DAMASIO, Cervelli che parlano: il dibattito su mente, coscienza ed intelligenza artificiale, Mondadori Bruno,Milano, 2003.

focalizzi sulla persona, dunque, analizzare questi dati in modo obiettivo potrebbe far ricadere il tutto nell'ambito della prova legale[269]. Bisognerebbe collegare questi dati ed analizzarli correttamente, tentativi sono già stati effettuati con un sistema di intelligenza artificiale chiamato ADVOKATE, un programma che valuta la credibilità dei testimoni[270].

La soluzione nell'utilizzo degli algoritmi nel campo della testimonianza, dunque, è inserire in essi tutte le informazioni ricavate dagli psicologi della testimonianza, i quali, studiando i comportamenti delle persone, sono in grado di intuire anche da minimi gesti quando mentano oppure no.

Inserendo tali elementi nella macchina, essa li immagazzinerà e sarà in grado di prendere decisioni sempre più simili a quelle che avrebbe preso un essere umano, questi elementi, dunque, necessitano di una valutazione umana preventiva.

Inoltre, la macchina può offrire un servizio davvero utile nella formulazione di un interrogatorio: nello svolgere le domande la macchina potrebbe risultare sicuramente molto obiettiva, soprattutto per quanto riguarda la scelta delle domande da porre e nella valutazione delle risposte date, domande che una persona potrebbe, magari per la curiosità o la poca attenzione, portare a sviare dalla causa principale, invece la macchina porrebbe tenendo attenzione alle minime sottigliezze[271].

Dunque, sistemi esperti sembrano poter fungere da intermediari tra gli utenti e le basi di dati, e a tal fine, bisogna definire la natura della conoscenza specialistica richiesta, la scelta delle tecniche più idonee alla formalizzazione di questa conoscenza e, infine, l'elaborazione di specifiche attività di supporto. Il modello d'interrogazione può essere generato per mezzo d'una analisi sintattico-semantica del testo delle domande rivolte al sistema, individuando i concetti fondamentali e le loro possibili varianti; ovvero, in un'ipotesi diversa, può essere condotta un'analisi più profonda del testo delle domande[272].

---

[269] J. NIEVA FENOLL, Inteligencia artificial y proceso judicial, Marcial Pons Ediciones Juridicas y Sociales, Madrid, 2018.

[270] M.C. BROMBY, M.J HALL, The Development and Rapid Evaluation ofthe Knowledge Model of ADVOKATE: anAdvisory System to Assess the Credibility ofEyewitness Testimony, January 2002. ADVOKATE è un'applicazione di supporto decisionale basata sulla conoscenza abilitata per il Web e progettata per l'uso in indagini penali, contenziosi civili o come ausilio didattico per la formazione investigativa. ADVOKATE fornisce una valutazione indicativa della credibilità dei testimoni oculari.

[271] J. NIEVA FENOLL, Inteligencia artificial y proceso judicial, Marcial Pons Ediciones Juridicas y Sociales,Madrid,2018

[272] http://www.ittig.cnr.it/EditoriaServizi/AttivitaEditoriale/InformaticaEDiritto/1991-1-3- EFameli5-26.pdf

Anche la branca delle neuroscienze ha offerto utili aiuti in questo senso: un esempio valido è dato dal fMRI[273], i cui risultati possono essere combinati con l'apprendimento automatico al fine di visualizzare il contenuto percettivo del cervello. Dunque, la IA ricostruisce le immagini e le proietta sullo schermo, i risultati sono sorprendenti: l'interazione tra capacità umana e tecnologica potrebbe arrivare a creare un nuovo livello di comunicazione, ove la IA funge da mediatore[274]. Importante è il contributo del Caltech (California Institute of Technology), che, in collaborazione con il Cedars-Sinai Medical Center e l'Università di Salerno, ha sviluppato un algoritmo in grado di leggere i dati dell'attività cerebrale, in particolare i cambiamenti nel flusso sanguigno in specifiche regioni del cervello, rilevati proprio grazie ad una Risonanza Magnetica Funzionale, dunque, di sviluppare una mappa dell'attività cerebrale che si traduce in una previsione di "intelligenza"[275]. Partendo da questi studi potrebbe risultare non difficile un giorno arrivare addirittura a predire quando una persona possa risultare propensa al reato.

- Per ciò che concerne la prova documentale: l'efficienza della IA nell'analisi documentale è ormai risaputa, grazie alla sua velocità essa infatti può smaltire documenti magari lasciati accumulare da tempo, garantendo così maggiore rapidità a garanzia di un'attività rapida ma efficace.

L'analisi però non deve essere rivolta solo alla rigida valutazione grammaticale, bisogna, infatti, che si tengano in conto anche altri fattori, quali il contesto in cui si trovava l'autore al momento della scrittura, al fine di rilevare eventuali vizi, e le modalità, quindi a quale ambito sia rivolta (per affari, questioni personali, ecc...). Nell'indagine sulla paternità del documento, dunque, sarà facile per la macchina, anche raffrontandolo ad altri documenti, riconoscere che magari la persona che lo ha scritto non poteva essere di basso livello culturale, dato l'utilizzo di forme lessicali complesse, quindi dedurre anche dal tipo di scrittura le abitudini e magari lo stile di vita della persona.

Eppure, l'agire della IA ha dei limiti, in quanto non basta solo l'analisi lessicale, come già accennato, ma le parole hanno bisogno di essere contestualizzate, la qual cosa, appunto, richiede inevitabilmente l'intervento umano.

---

[273] fMRI (Risonanza Magnetica Funzionale) è un test diagnostico con finalità mediche, che si serve della tecnica di imaging a risonanza magnetica per valutare le funzionalità di un organo o di un apparato.

[274] J. NOSTA, L'AI può leggere nel pensiero. E tradurlo in immagini e parole, articolo scritto per Fortune.com, Maggio 2018.

[275] S. PIACENTI, AI: è possibile predire la capacità intellettiva?, 25 Settembre 2018. https://www.ingenium-magazine.it/ai-possibile-predire-capacita-intellettiva/

In ogni caso (ex art 233 cpp.)[276], nulla impedisce a chi voglia contestare la prova algoritmica, di nominare un consulente tecnico al di fuori della perizia.

L'utilizzo di sistemi tecnologici in campo probatorio d'altronde era stato già ampiamente sondato con il poligrafo: la classica macchina della verità, che agiva per mezzo di fasce intorno alla testa e cavetti legati al corpo. Eppure, Julia Hirschberg[277], ha lavorato alla possibilità di rimpiazzare il tradizionale poligrafo con un software in grado di analizzare il discorso di una persona attraverso un metodo approfondito che si serve delle sonorità, degli intercalari, del cambio di tono di una persona e decine di altri piccoli segnali, utili a dedurre il vero e il falso. Il 70% dei bugiardi sono stati scoperti.

Questa innovazione può condurre alla concreta possibilità per la macchina di analizzare anche il "discorso emotivo", il che è da sempre stato ritenuto proprio il limite dei sistemi artificiali.

David F. Larcker[278], ha analizzato le dichiarazioni effettuate dai dirigenti finanziari di Wall Street, scoprendo che essi spesso utilizzano parole quali "chiaramente" o "molto chiaramente" che di fondo suggeriscono un inganno.

Julia Hirschberg e due suoi colleghi hanno ricevuto anche una sovvenzione di ben 1,5 milioni di dollari dall'aeronautica militare, la Air Force, con l'obiettivo di sviluppare algoritmi per analizzare i discorsi di chi parla in inglese, ma anche in arabo e in cinese[279].

I sistemi di IA, applicati al campo probatorio, potrebbero divenire dei veri e proprio "Sherlock Holmes", verificando con maggiore intensità e accuratezza anche i più piccoli segni derivanti dalle emozioni, che spesso tendono a "tradire" le bugie delle persone. Questa novità potrebbe davvero livellare le barriere uomo-macchina.

Infine, si può concludere citando alcuni esempi concreti di applicazione della IA in campo probatorio:

• QuestMap. È uno strumento informatico che si basa su IBIS, un sistema informativo il cui scopo è la identificazione e la risoluzione di problemi, tramite la loro scomposizione in singole questioni, esso, dunque, media discussioni, supporta argomentazioni collaborative e crea mappe informative.

[276] Art 233 cpp co 1: "Quando non è stata disposta perizia, ciascuna parte può nominare, in numero non superiore a due, propri consulenti tecnici..."

[277] Professoressa di computer science alla Columbia University.

[278] Professore a Stanford Graduate School of Business.

[279] https://www.repubblica.it/tecnologia/2011/12/07/news/macchina_verit-26122762/

• Convince me. Esso è uno strumento informatico a supporto dell'argomentazione. Gli argomenti consistono in reti causali di nodi (che possono mostrare prove o ipotesi) e la conclusione che gli utenti traggono da essi. Convince Me predice le valutazioni dell'utente delle ipotesi basate sugli argomenti prodotti, e fornisce un feedback sulla comprensibilità delle inferenze che gli utenti disegnano.

• STEVIE. È uno strumento informatico basato su argomentazioni destinato a supportare le indagini penali. Stevie permette agli analisti di ottenere prove e deduzioni. Il programma è descritto come distillazione tra quelle informazioni coerenti, che sono ricostruzioni ipotetiche di ciò che potrebbe essere accaduto, e quelle definite come una raccolta di rivendicazioni prive di conflitti[280].

5. Giustizia predittiva: Articolo 33 della legge n. 2019/222

La legge n. 2019/222, promulgata il 23 marzo 2019 dal Presidente della Repubblica francese dopo il rinvio al Consiglio costituzionale per verificarne la costituzionalità, prevede una ricostituzione amministrativa del sistema della giustizia, sulla base di situazioni nuove che necessitano di una regolazione che a loro si adegui, dunque, impone ingenti sanzioni penali per le società o chiunque, partendo dalle decisioni e sentenze, raccolga, analizzi e riutilizzi "i dati di identità dei magistrati con lo scopo o l'effetto di valutare, analizzare, confrontare o prevedere le loro pratiche effettive o presunte pratiche professionali". Il divieto vale sia per la giustizia ordinaria sia per quella amministrativa e sorpassarlo potrebbe costare fino a cinque anni di prigione.

Questa è una riforma globale e concreta della che mira a fornire una giustizia più accessibile, più rapida ed efficace al servizio dei cittadini e di coloro che rendono giustizia[281].

L'accoglimento di questa riforma serve da un lato ad operare una migliore scelta degli open data della giustizia, dall'altro serve anche a limitare il servirsi di tali open data per impostare su base predittiva le ipotesi di difesa[282].

L'art. 33 dunque mira a regolare la divulgazione delle decisioni giudiziarie, non sarà, dunque, più possibile valutare e analizzare statistiche quali, ad esempio, quelle che riguardano se i giudici sono più propensi alla prigione o alla libertà, quante volte hanno

[280] E. NISSAN, Digital technologies and artificial intelligence's presentand foreseeable impact on lawyering, judging, policing and law enforcement, Londra, 14 Ottobre 2015.
[281] http://www.justice.gouv.fr/la-garde-des-sceaux-10016/la-reforme-de-la-justice-entre-en- vigueur-32242.html
[282] https://www.privacyitalia.eu/privacy-niente-profilazione-dei-giudici-in-funzione-predittiva-in- francia/11282/

concesso una misura cautelare, ecc.... Tutti questi dati sono preziosi per una misura giudiziale, difatti molte aziende si lamenteranno di questa norma perché si occupano proprio di elaborare intelligenze artificiali che si basano su questi dati e il cui operato potrebbe esser messo alla dura prova alla luce di una tale riforma[283].

[283] Convegno tenutosi presso l'Università degli Studi di Salerno, il quale ha visto come ospite il professore Jordi Nieva-Fenoll, per la presentazione del suo libro "Intelligenza artificiale e processo", 7 Giugno 2019.

IV CAPITOLO
IA E RESPONSABILITA' PENALE

SOMMARIO: 1. Personalità giuridica dell'Intelligenza artificiale – 1.1 Tesi positiva: responsabilità diretta dell'intelligenza artificiale e personalità elettronica – 1.2 Tesi funzionalista – 1.3 Tesi del comportamentismo metodologico – 1.4 Tesi strutturalista-ontologica – 1.5 Paragone con la responsabilità degli enti – 2. Responsabilità penale personale: principi generali applicati all'Intelligenza artificiale – 2.1 Elemento oggettivo: condotta, evento e rapporto di causalità – 2.2 Elemento soggettivo: la colpevolezza. Coscienza e volontà – 2.3 Fine rieducativo della pena – 3. Responsabilità del programmatore, dell'utilizzatore o dell'Intelligenza artificiale? – 3.1 Il problema della responsabilità oggettiva: responsabilità almeno a titolo di colpa – 3.2 Il problema del controllo significativo: la delegazione e la responsabilità da controllo indiretto – 3.3 Il problema delle posizioni di garanzia ex art. 40 cpv. c.p. – 3.4 Il problema dell'individuazione del responsabile: colpa eventuale del programmatore

1. Personalità giuridica dell'Intelligenza artificiale

A questo punto dell'analisi, è il momento di volgere l'attenzione su un tema centrale della trattazione: l'intelligenza artificiale può essere dotata di un'autonoma personalità giuridica? Rispondendo affermativamente a questa domanda si porranno questioni relative alla possibilità di considerare il sistema intelligente come penalmente responsabile dei crimini da questo commessi e si potranno analizzare i vari aspetti della responsabilità penale declinati per l'intelligenza artificiale. Qualora dovessimo rispondere negativamente, invece, risulterà fondamentale comprendere chi possa essere il responsabile di tali illeciti senza rischiare di ricadere in forme di responsabilità oggettiva, di più semplice ed immediata soluzione.

È possibile considerare l'intelligenza artificiale come un agente di diritto?
Sussistono quattro principali significati della parola "agente": una forza o una sostanza che provoca un cambiamento, una forza naturale che produce risultati (come nel caso di agenti chimici o insetti intesi come agenti di fecondazione); l'essere umano che

agisce o ha il potere di agire; colui che viene autorizzato ad agire in rappresentanza di altri; e la persona o la res attraverso la quale l'azione viene compiuta[284].

Dunque, per "agente" si intende tutto ciò, non necessariamente umano, che sia in grado di provocare un cambiamento nell'ambiente circostante e quindi causalmente responsabile per le azioni che commette.

Grazie all'evoluzione tecnologica e all'avvento dei sistemi adattivi complessi, viene aggiunta come fondamentale caratteristica dell'"agente" la sua capacità di elaborare informazioni grazie alla percezione dell'ambiente attraverso i sensori e grazie alle azioni che può svolgere tramite gli attuatori[285]. Dunque, nonostante i robot intelligenti abbiano la natura di res (e non di persona fisica), potranno essere considerati come "agenti" del diritto in senso ampio. Esso diventa "razionale" in quanto capace di selezionare l'azione che massimizza il risultato e "autonomo" perché in grado di compensare le parziali informazioni attraverso l'apprendimento esperienziale.

Tale questione è connessa al concetto di "persona" per il diritto.

"Noi non siamo immobili al centro dell'universo (rivoluzione copernicana); noi non siamo innaturalmente distinti e differenti dal resto del mondo animale (rivoluzione darwiniana) e siamo ben lungi dall'essere cartesianamente interamente trasparenti a noi stessi (rivoluzione freudiana). Noi stiamo ora lentamente accettando l'idea che potremmo non essere così nettamente diversi da altre entità e agenti informazionali e intelligenti, e da artefatti ingegnerizzati (rivoluzione di Turing)"[286].

Hans Kelsen critica le definizioni comuni di "persona fisica" considerata come l'essere umano investito di diritti e doveri e opposta alla "persona giuridica" vista come non-umano[287]. Dire che un essere umano ha un certo dovere o diritto significa solo che un determinato comportamento dell'individuo è contenuto in una norma giuridica. Le norme però non disciplinano l'intera esistenza dell'essere umano ma solo alcune azioni o omissioni particolari. La persona esiste giuridicamente solo in quanto destinatario di diritti e doveri. La "persona fisica" è un concetto generato dall'analisi di norme giuridiche in relazione all'"essere umano" (concetto biologico e fisiologico delle scienze naturali). L'essere umano biologicamente inteso risulta un concetto differente da quello della persona fisica secondo i criteri giuridici. Il primo esiste per

[284] Random House Kernerman Webster's College Dictionary, 2010: https://www.thefreedictionary.com/agent
[285] A. SANTOSUOSSO, op. cit., Paragrafo 7.9 (pag. 193 – 195) "Diritti, storicità, artificialità – L'agente per il diritto".
[286] L. FLORIDI, "Philosophy of computing and information. 5 Questions", (pag. 95) Automatic Press/VIP, 2008.
[287] H. KELSEN,"General Theory of Law and State", (pag. 93 – 95) Harvard University Press, Cambridge, 1945

l'ordinamento esclusivamente nella parte in cui sussistono norme che su di lui vanno ad incidere. Infatti, si può affermare che l'essere umano sia posto a fondamento della persona fisica delimitandola, inoltre, nello spazio. Il rapporto tra i due è legato al fatto che i diritti e doveri, compresi nella nozione di persona, si riferiscono al comportamento di quello specifico essere umano. Secondo Kelsen, la "persona fisica", essendo un costrutto giuridico può essere considerata anch'essa come "persona giuridica" in senso ampio e dunque non si differenzia da quest'ultima: entrambe sono creazioni artificiali del diritto[288].

Bisogna quindi comprendere quale sia, per il diritto, il limite secondo cui un "agente" possa essere considerato tale in senso giuridico. Il limite coincide con la differenza tra cose e persone?

Come abbiamo già esplicato nel primo capitolo, i robot sono macchine disegnate per evocare nelle sembianze e nelle relazioni con l'ambiente, un essere vivente umano o animale. Alcuni di essi possiedono al loro interno un codice morale e sono dotati di un'autonomia che si alimenta grazie alla capacità di autoapprendimento e di compimento di azioni non programmate. Il codice morale dovrebbe permettere alla macchina di individuare la condotta più appropriata e meno nociva tra il novero delle azioni che può compiere[289].

Il tratto qualificante del robot, in assenza di una definizione univoca, è incentrato sulla sua capacità di eseguire autonomamente azioni nell'ambiente circostante senza un controllo dell'uomo. Questo risulta fortemente ambiguo e di difficile inquadramento giuridico[290]. Come già affermato, non sussiste una definizione univoca di intelligenza artificiale. Le categorie giuridiche esistenti non riescono ad inquadrare le forme di intelligenza artificiale più evolute ed autonome e la dottrina ha anche valutato la possibilità di riconoscere tali robot come soggetti di diritto andando a generare una serie di problematiche.

Innanzitutto, bisogna valutare se l'intelligenza artificiale debba essere ricondotta alla categoria delle persone giuridiche oppure delle persone fisiche andando a realizzare un robot con le stesse caratteristiche giuridiche dell'essere umano. È possibile notare come la differenza tra persone fisiche e giuridiche non sia esente da incongruenze.

---

[288] A. SANTOSUOSSO, op. cit., Paragrafo 7.8 (pag. 190 – 192) "Diritti, storicità, artificialità – Creature di Dio e artefatti giuridici"

[289] E. PALMERINI, "Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea", Responsabilità Civile e Previdenza, fasc.6, 2016, pag. 1815B; Paragrafo 2 "Il robot sulla scena giuridica"

[290] E. PALMERINI, op. cit., Paragrafo 3.2. "Che cos'è un robot? Base tecnologica e questioni giuridiche"

Nonostante le loro differenze, entrambe sono considerate come entità titolari di diritti e doveri. Le società sono enti che, mediante contratto con il quale due o più persone conferiscono beni o servizi per l'esercizio in comune di un'attività economica allo scopo di dividerne gli utili, possiedono una personalità giuridica. Esse vengono infatti trattate come individui secondo la legge e la loro responsabilità è indipendente e distinta da quella delle persone che la compongono. Si tratta di una fictio iuris che va a considerare responsabile un ente ovviamente non-umano[291].

Le società non sono in grado di agire autonomamente ma esclusivamente attraverso un rappresentante legale umano e dunque, se l'intelligenza artificiale dovesse rientrare nella categoria di persona giuridica non potrebbe essere considerata come interamente autonoma ma sempre connessa ad un rappresentante legale. In realtà l'intelligenza artificiale e le società possiedono notevoli differenze potendo considerare le prime, a differenza delle seconde, come veri attori dell'azione e dunque, forse, responsabili dei reati commessi. Infatti, a differenza delle società, i robot intelligenti hanno l'abilità di agire concretamente nel mondo esterno essendo dotati di una fisicità propria e separata da quella dell'essere umano che li ha realizzati: gli eventi dipendenti dalle loro azioni sono un logico decorso causale di una loro attività successiva ad una loro decisione assunta grazie ad un meccanismo di auto machine learning completamente indipendente dall'essere umano e da lui imprevedibile (per quanto concerne i sistemi di intelligenza artificiale più evoluti). Le persone giuridiche, invece, non esisterebbero in assenza dei membri (persone fisiche) che le compongono: agiscono solo ed esclusivamente attraverso le azioni dei loro rappresentanti senza poter modificare causalmente l'ambiente circostante in modo autonomo. Alla fine di questo paragrafo, proveremo ad analizzare le analogie tra il Decreto legislativo 231/2001 (Disciplina della responsabilità amministrativa delle persone giuridiche, delle società e delle associazioni anche prive di personalità giuridica) e una possibile responsabilizzazione dell'intelligenza artificiale.

Come sopra detto, i sistemi intelligenti hanno l'abilità di agire in completa autonomia rispetto all'essere umano, di conseguenza sussiste una mancanza di controllo da parte del produttore sulla futura attività realizzata dall'intelligenza artificiale dovuta inoltre all'oscurità del processo di elaborazione seguito dalla macchina nella trasformazione degli input raccolti dall'esterno (o immessi dal programmatore) in output. Per questo,

[291] A. SANTOSUOSSO, op. cit., Paragrafo 7.9 (pag. 193 – 195) "Diritti, storicità, artificialità – L'agente per il diritto"

il fenomeno del black box algorithm, che rende imprevedibile la condotta della macchina e quasi incomprensibili i suoi processi decisionali, risulta fortemente connesso al tema del responsibilty gap dovuto alla mancanza di controllo da parte del produttore.

Un modello black box è un sistema che, similmente ad una scatola nera, è descrivibile essenzialmente nel suo comportamento esterno ovvero solo per come reagisce in uscita (output) a una determinata sollecitazione in ingresso (input), ma il cui funzionamento interno è non visibile o ignoto. Tale definizione nasce dalla considerazione che nell'analisi del sistema ciò che è veramente importante a livello macroscopico ovvero a fini pratici è il comportamento esterno, specie in un contesto di interconnessione di più sistemi, piuttosto che il funzionamento interno il cui risultato è appunto proprio il comportamento esterno.

La nozione di "responsability gap " (lett. "divario di responsabilità") legata all'intelligenza artificiale è stata originariamente introdotta nel dibattito filosofico per indicare la preoccupazione che gli "automi di apprendimento" possano rendere più difficile o impossibile attribuire la responsabilità morale alle persone per eventi spiacevoli. Il gap di responsabilità, si sostiene, non è un problema ma un insieme di almeno quattro problemi interconnessi – gap di colpevolezza, responsabilità morale e pubblica, responsabilità attiva – causati da diverse fonti, alcune tecniche, altre organizzative, legali, etiche e sociale. Fermo restando che lacune di responsabilità possono verificarsi anche con sistemi che non apprendono[292].

Le proposte principali per superare il problema del responsibilty gap risultano essere di tre tipologie.

Innanzitutto, c'è chi affida alla limitazione della responsabilità due compiti: promuovere l'innovazione della ricerca robotica riducendo i rischi economici e garantire immunità ai produttori rispetto agli eventi di danno che non avrebbero potuto evitare neanche usando la necessaria diligenza[293].

L'idea di esonerare i produttori dalla responsabilità per eventi che rimangono al di fuori della loro capacità di controllo, risulta connessa alla difficoltà, per i produttori e programmatori, di anticipare i rischi a cui potrebbero essere esposti. Ma al contempo sussiste chi considera tale esonero come ingiustificato alla luce della mera

---

292 *"Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them"*, di Filippo SANTONI de SIO & Giulio MECACCI, in Philosophy & Technology volume 34, pagg. 1057–1084 (2021)

293 E. PALMERINI, op. cit., Paragrafo 6.1 "Le proposte di schemi alternativi alle comuni regole di responsabilità"

desiderabilità sociale dell'innovazione tecnologica e che dunque i danneggiati dovrebbero pur sempre essere tutelati[294]. Inoltre, questa tipologia di responsabilità (e di limitazione della stessa) potrebbe essere applicata soltanto nel caso in cui la capacità cognitiva e decisionale delle macchine intelligenti venga assimilata ai soggetti che per età o incapacità non possono essere considerati responsabili direttamente per le proprie azioni lesive venendo sostituiti da chi se ne occupa (come avviene per gli animali, dotati di limitata razionalità).

In secondo luogo, c'è chi considera la personalità giuridica del robot utile per renderli direttamente responsabili per i danni arrecati a terzi. La personalità elettronica potrebbe essere applicata sia ai robot dotati di un corpo (meccanico o che emula le sembianze umane o animali), sia ai software intelligenti (non dotati di un corpo fisico ma digitale costituito da algoritmi e dati) che siano considerati autonomi. Con tale previsione bisognerebbe realizzare un registro in cui inserire ogni robot al momento della messa in commercio garantendone un controllo anche sulle possibili assicurazioni connesse ad un fondo patrimoniale alimentato da coloro che sono intervenuti nella sua creazione, in modo tale da permettere alla macchina intelligente di rispondere delle obbligazioni[295]. La creazione di una personalità giuridica per i sistemi di intelligenza artificiale va intesa in mero senso funzionale come meccanismo che consente l'imputazione di effetti direttamente in capo alla macchina. Sicuramente tale approccio risulta necessario per evitare un freno allo sviluppo tecnologico nel caso in cui si considerasse responsabile il solo produttore che inoltre non sempre potrebbe essere titolare di un patrimonio idoneo per risarcire gli eventuali danni commessi dalla macchina.

Allo stesso tempo, tale tesi si espone a due tipologie di critiche: il possibile fraintendimento nella creazione di una personalità giuridica per i robot, rischiando di attribuire soggettività ad un ente meccanico andando a porre similitudini tra esseri umani e macchine in base a capacità cognitive; e la sproporzione del mezzo (dotare i robot di personalità elettronica) rispetto ai fini sopra elencati[296]. Il riconoscimento della personalità elettronica al sistema autonomo potrebbe generare un crescente antropomorfismo nei confronti degli agenti artificiali rischiando di comportare una

[294] E. PALMERINI, op. cit., Paragrafo 6.2 "Alcuni rilievi critici"
[295] E. PALMERINI, op. cit., Paragrafo 6.1 "Le proposte di schemi alternativi alle comuni regole di responsabilità"
[296] E. PALMERINI, op. cit., Paragrafo 6.2 "Alcuni rilievi critici"

deresponsabilizzazione degli operatori. L'autonomia dei robot non viene considerata come sufficiente al fine del riconoscimento della soggettività giuridica[297].

Infine, l'ultimo orientamento spinge verso un inasprimento della responsabilità del proprietario a tutela del danneggiato. Fissando una soglia massima di risarcimento nel caso di danni in capo al proprietario in modo tale da assicurare il rischio[298]. Tale ideologia punta addirittura a considerare la responsabilità come oggettiva e dunque risulta criticabile dato che ne viene investito esclusivamente l'utente finale e non anche il produttore della macchina senza inoltre riuscire a garantire la certezza di soddisfazione dei danneggiati[299].

1.1 Tesi positiva: responsabilità diretta dell'intelligenza artificiale e personalità elettronica

Data la difficoltà di inserire le macchine intelligenti all'interno di una delle due categorie (persona fisica o giuridica), agli inizi degli anni novanta si iniziò a parlare del riconoscimento, all'intelligenza artificiale, della qualifica di "personalità elettronica" nel caso in cui la struttura del programma informatico non permetta di prevedere a priori le future condotte poste in essere dalla macchina. In tale modo, si riconosce alla stessa autonomia giuridica e si libera il produttore dalla responsabilità per le lesioni provocate dal robot[300].

L'intelligenza artificiale verrebbe quindi concepita come un autonomo centro di interessi e di imputazione destinato a rispondere dei danni provocati. Risulta quindi necessaria la creazione di uno specifico registro nel quale iscrivere le macchine con un numero identificativo, affidandole un fondo patrimoniale alimentato da coloro che sono intervenuti nella sua creazione, per rispondere alle obbligazioni e prevedendo forme di assicurazione obbligatoria[301]. In tal modo, si genera una separazione dei patrimoni e dunque, uno schermo per i produttori e investitori evitandogli un'eccessiva esposizione al rischio.

---

[297] B. PANATTONI, "Intelligenza artificiale: le sfide per il diritto penale nel passaggio dall'automazione tecnologica all'autonomia artificiale", Diritto dell'Informazione e dell'Informatica (Il), fasc.2, 1 aprile 2021, pag. 317, Paragrafo 2 "Dall'automazione tecnologica all'autonomia artificiale"

[298] E. PALMERINI, op. cit., Paragrafo 6.1 "Le proposte di schemi alternativi alle comuni regole di responsabilità"

[299] E. PALMERINI, op. cit., Paragrafo 6.2 "Alcuni rilievi critici"

[300] M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, op. cit., Paragrafo 3.4 (pag. 31 - 36) "L'Intelligenza Artificiale più evoluta considerata come una persona"

[301] C. PIERGALLINI, "Intelligenza artificiale: da 'mezzo' ad 'autore' del reato?" Rivista Italiana di Diritto e Procedura Penale, fasc.4, 1 dicembre 2020, pag. 1745; Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'"

Per questo i fautori di tale idea sottolineano la necessità di attribuire personalità giuridica all'intelligenza artificiale al fine di considerarla come soggetto di diritto. Altri obiettano dicendo che tale riconoscimento servirebbe, esclusivamente, a limitare la responsabilità dei produttori o programmatori e che, inoltre, non si modificherebbe il soggetto che di fatto sopporta i costi del danno, dato che il capitale verrebbe comunque conferito da un umano in quanto il robot non possiede capacità di ricevere personalmente un compenso per le sue attività. Non si avrebbe, dunque, una modificazione dell'onere della prova dato che si tratterebbe di un'ipotesi di responsabilità oggettiva in cui l'obbligo di risarcimento del danno verrebbe a gravare sul creatore del robot, che lo rappresenta legalmente, a seguito della prova del danno, del difetto e del nesso causale tra i due. Quindi, si potrebbe giungere al medesimo risultato prevedendo un obbligo di assicurazione per le macchine autonome a beneficio di terzi[302]. Per questo parte della dottrina risulta contraria al riconoscimento della personalità elettronica ai robot ritenendola superflua.

Altra parte della dottrina ritiene che riconoscendo tale personalità si eliminerebbe il problema della complessa individuazione del soggetto responsabile nella catena di produzione delle intelligenze artificiali in quanto composte da prodotti digitali disgregati e poi assemblati in varie fasi. Sarebbe dunque utile riconoscere la personalità giuridica al robot al fine di riunire i profili di responsabilità in capo ad un'unica entità[303]. Nel momento in cui il progresso tecnologico porterà le macchine intelligenti ad autodeterminarsi e a comprendere il disvalore sociale delle loro azioni, il legislatore dovrà valutare l'opportunità di prevedere una forma di responsabilità diretta degli agenti artificiali completamente autonomi che dovranno essere dunque dotati di capacità penale. A tal punto potrà essere mosso agli agenti artificiali un rimprovero per aver commesso un fatto antigiuridico e colpevole[304].

Il maggior esponente della possibilità di considerare una forma di responsabilità penale diretta dei sistemi di intelligenza artificiale è Gabriel Hallevy, professore di diritto penale presso la Facoltà di Giurisprudenza, Ono Academic College, in Israele. Parte della dottrina, però, reputa tale teoria ipotizzabile solo in campo teorico non riconoscendo di fatto la possibilità di considerare le entità artificiali come soggetti di

[302] M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, op. cit., Paragrafo 3.4 (pag. 31 - 36) "L'Intelligenza Artificiale più evoluta considerata come una persona"

[303] M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, op. cit., Paragrafo 3.4.2 (pag. 31 - 36) "L'Intelligenza Artificiale più evoluta considerata come una persona – EPersons: rilievi critici e prospettive"

[304] I. SALVADORI, "Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale", Rivista Italiana di Diritto e Procedura Penale, fasc.1, 1 marzo 2021, pag. 83; Paragrafo 9, "Considerazioni finali"

diritto liberi e capaci di compiere scelte volontarie e dunque non considerabili come penalmente rimproverabili o punibili per le condotte da loro realizzate[305].

La sua tesi positiva consiste nel ritenere che non ci siano ragioni valide per negare la punibilità delle macchine intelligenti in quanto l'elemento oggettivo del reato potrebbe essere ricondotto direttamente al sistema intelligente in grado di compiere un atto penalmente rilevante[306]. Mentre, per quanto concerne l'elemento soggettivo, i sistemi più evoluti di intelligenza artificiale hanno la capacità di rappresentarsi la realtà grazie all'acquisizione e all'analisi dei dati raccolti dal mondo esterno. In tal modo possiedono addirittura la capacità di prevedere e volere un certo risultato come conseguenza della propria azione grazie a processi di decision-making (attività di ragionamento che genera una scelta a seguito di un'informazione ricevuta per raggiungere un obiettivo prefissato) indirizzando l'operato dell'agente intelligente in relazione alla probabilità che si verifichi un determinato evento. Inoltre, è possibile anche configurare l'imprudenza della macchina nel caso in cui la stessa non consideri una probabilità che avrebbe dovuto prevedere in base agli input immessi. Non è rilevante, dunque, la possibilità che il sistema intelligente provi dei sentimenti perché essi sono considerati estranei al concetto di dolo o colpa, non rientrando nella tipicità della maggior parte delle fattispecie penali[307].

A tale tesi viene obiettato principalmente il contrasto con l'articolo 27 comma 1 della Costituzione secondo cui la responsabilità penale è personale. Sulla base di tale principio, per infliggere una pena a seguito del reato è necessario che l'autore abbia la possibilità di agire diversamente, ciò non accade nei sistemi intelligenti perché nonostante la loro evoluzione, essi sono programmati per agire (anche se le neuroscienze di recente stanno iniziando a mettere in dubbio anche la sussistenza di un effettivo libero arbitrio nell'uomo, destinato ad agire senza una reale percezione della volontà di azione). Inoltre, si lederebbe anche l'articolo 27 comma 3 della Costituzione in quanto la pena non potrebbe assolvere al suo scopo principale, l'unico previsto dalla Carta costituzionale, vale a dire quello rieducativo, dato che

[305] B. PANATTONI, op. cit., Paragrafo 5.1 "Forme di responsabilità diretta a carico dell'agente artificiale"
[306] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?"
[307] A. CAPPELLINI, "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale", op. cit., Paragrafo 4 (pag. 10 - 13) "La responsabilità diretta dell'IA: la tesi positiva di Gabriel Hallevy"

l'intelligenza artificiale non può provare timore, rimorso, rimpianto e dunque non può operare nei suoi confronti l'effetto dissuasivo prima e risocializzante poi della pena[308]. Inoltre, per quanto concerne l'elemento soggettivo viene criticato il fatto che l'abilità della macchina ad elaborare i dati captati dall'esterno, la sua capacità di rappresentarsi la realtà, di prevedere un risultato come conseguenza di una sua azione e il suo volerlo perché programmate per ottenerlo non risulta sufficiente per scardinare i principi fondamentali dell'ordinamento penale. Il dolo che sembra così essere dimostrato, in realtà non sussiste, si tratta di una mera apparenza in quanto le macchine intelligenti non sono in grado di autodeterminarsi al pari dell'uomo. Ad oggi i robot, non hanno una piena capacità di scelta, non possiedono il libero arbitrio e di conseguenza non possono essere punite come colpevoli[309].

1.2 Tesi funzionalista

Il substrato della teoria positiva di Hallevy, che ha lo scopo di affermare giuridicamente la possibilità di riconoscere la personalità elettronica ai sistemi di intelligenza artificiale, è connesso alla teoria funzionalista degli anni Sessanta e Settanta i cui principali esponenti sono Hilary Putnam e Jerry A. Fodor (filosofi e ricercatori statunitensi nel campo delle scienze cognitive e dei meccanismi mentali). Tale teoria si ripromette di rispondere alla domanda sulla sussistenza o meno di un libero arbitrio artificiale. Applicando tale teoria, nonostante non sia il suo obiettivo finale, è possibile considerare le macchine come soggetti di diritto e dunque destinatari di norme giuridiche[310].

Lo scopo di tale teoria, scollegato da possibili approdi giuridici, consiste in uno studio del funzionamento della mente umana per poi applicarlo ai sistemi di intelligenza artificiale. Si tratta di un'analogia tra mente e computer in senso non ontologico-strutturale ma in senso funzionale: ricondurre la mente umana al modello simbolico-computazionale dei sistemi intelligenti.

La ricerca scientifica, per migliorare le tecniche di machine learning, si è basata sull'apprendimento neurologico dei bambini. Infatti, i robot, come gli infanti,

[308] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?"

[309] A. CAPPELLINI, "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale", op. cit., Paragrafo 5 (pag. 13 - 15) "Prima critica: persistente assenza di colpevolezza"

[310] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'"

scoprono e apprendono il mondo grazie alla percezione dell'ambiente esterno tramite il proprio corpo e i propri sensori.

Si tende quindi a supporre una completa analogia tra umano e computer in senso meramente funzionale.

In tale ambito non sussiste alcuna differenza tra i due, l'unico discrimine riguarda il supporto in cui viene immagazzinata la conoscenza: il cervello per l'essere umano e le componenti meccanico-digitali per il robot. L'intelligenza umana è data dall'attività neuronale e dunque risulta possibile ottenere un sistema neurale artificiale che imiti tale attività generando una mente artificiale intelligente[311].

Gli stati mentali del cervello umano sono qualificati in base alla loro funzione e non alla loro sede materiale quindi sono slegati dal sistema organico e dunque il sistema nervoso umano risulta riproducibile grazie alle reti neurali: realizzando artificialmente l'attività neuronale espressione dell'intelligenza, si genera di conseguenza un duplicato dell'intelligenza stessa[312].

Infatti, le funzioni cognitive di un essere umano sono indistinguibili da quelle della macchina che potrebbe risultare addirittura più intelligente grazie alla logica computazionale-conseguenziale: l'essere umano risulta affetto da una razionalità limitata rispetto alla macchina[313].

Come sappiamo l'intelligenza artificiale è la disciplina secondo cui si può emulare ogni aspetto dell'intelligenza umana. Essa imita e riproduce, tramite componenti elettroniche, l'attività mentale umana. Gli artefatti artificiali intelligenti possiedono strutture mutuate da quelle umane. L'idea di Alan Turing consisteva nel considerare le macchine capaci di esprimere un pensiero e di produrre processi intellettuali identici a quelli umani. Il cervello umano viene così paragonato all'hardware e la mente al software. Non interessa, dunque, la struttura degli stati mentali ma la funzione esercitata, a prescindere dalla sede fisica in cui si trovano. Alla base della teoria funzionalista sussiste tale principio: un computer è paragonabile ad un essere umano

[311] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 3 (pag. 6 - 7) "Gli agenti intelligenti agiscono autonomamente, e sono perciò agenti liberi? La tesi funzionalista"

[312] M. B. MAGRO, "Biorobotica, robotica e diritto penale", in D. Provolo, S. Riondato, F. Yenisey (a cura di), Genetics, Robotics, Law, Punishment, Padova University Press, 2014, pp. 510 s.; Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali"

[313] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 3 (pag. 6 - 7) "Gli agenti intelligenti agiscono autonomamente, e sono perciò agenti liberi? La tesi funzionalista"

se le sue prestazioni non possono essere distinte da quelle svolte dall'uomo (the imitation game).

Questo, però, non genera anche l'attribuzione di una coscienza in capo ai robot, in quanto gli stessi non possiedono stati psicologici rilevabili che possono comportare una consapevolezza delle azioni.

Secondo tesi opposte, non sussiste un'autenticità nel pensiero meccanico e dunque ci troviamo difronte ad una diversità ontologica tra intelligenza artificiale e umana. Le macchine simulano i processi umani, ma non li realizzano come propri: copiano. Risulta complesso attribuire ai computer una nozione di intelligenza intesa come intenzionalità. Essa risulta fondamentale per l'uomo e ciò lo distingue dalla macchina. Alla tesi funzionalista si oppone lo strutturalismo che considera la struttura (o la sede) e non la funzione delle attività mentali al centro del dibattito: l'intenzionalità si trova nel cervello umano. E dunque non è possibile paragonare un computer ad un essere umano in quanto non possiede la sua stessa abilità di pensiero[314]. Trasmigrando tale pensiero nel mondo giuridico, secondo tale tesi, non potremmo di conseguenza affermare (come vedremo nel secondo paragrafo di questo capitolo) la sussistenza di una coscienza in capo alle macchine tale da permettergli di essere considerate come soggetti di diritto paragonabili alle persone fisiche.

Nonostante ciò, è indubbio che i robot siano oggi soggetti ad un processo di umanizzazione. Infatti, i sistemi basati sull'intelligenza artificiale sono capaci di evolversi grazie al loro apprendimento e sono dotati di una propria libertà di movimento e cambiamento. Tale libertà è però differente rispetto a quella dell'essere umano il cui funzionamento cognitivo risulta tutt'ora in parte sconosciuto.

Il concetto di responsabilità della macchina è connesso ma non consegue al grado di libertà di cui gode l'intelligenza artificiale, la quale potrà essere considerata responsabile anche se non sussiste una piena sovrapposizione tra libertà umana e artificiale. Nel corso della Storia[315] non è mancato chi considerasse il libero arbitrio come una mera illusione dell'essere umano in quanto in realtà esso non ha una piena libertà di azione. Seguendo tali teorie è possibile riscontrare che anche gli umani, come le macchine, sono determinati a priori da un filo che li guida per tutta la vita, e dunque, secondo tale teoria, libertà non verrebbe considerata come un fenomeno naturale ma

[314] M. B. MAGRO, "Biorobotica, robotica e diritto penale", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali"

[315] G. GALUPPI, "Libero arbitrio, imputabilità, pericolosità sociale e trattamento penitenziario", Dir. famiglia, fasc.1, 2007, pag. 328; Paragrafo 2 "Cenni sulla negazione del libero arbitrio nei secoli passati"

come mera attribuzione normativa al reo al fine esclusivo di risultare funzionale per scopi di organizzazione sociale. In tal senso, riportando il discorso al tema dell'intelligenza artificiale, le azioni commesse da una persona fisica e una elettronica risultano funzionalmente similari[316].

Il concetto di libero arbitrio si atteggia, infatti, come una costruzione interna al sistema sociale a prescindere dalle caratteristiche biologiche del soggetto. La persona è un artificio socialmente indotto e proiettabile verso qualsiasi entità capace di deludere le aspettative sociali di comportamento. Si è "persona" non per questioni naturali ma per regole sociali che le affidano diritti e doveri.

Dunque, in base alla teoria funzionale, partendo dal presupposto che anche i sistemi di intelligenza artificiale possono deludere tali aspettative, gli stessi potranno essere assoggettati alle norme di diritto penale: in seguito ad un riconoscimento sociale della macchina intelligente, non sussisterebbero limiti alla possibilità di sanzionarla[317].

1.3 Tesi del comportamentismo metodologico

Lo sviluppo di robot sempre più antropomorfici in grado di relazionarsi con l'essere umano genera sistemi di intelligenza artificiale dotati di emozioni artificiali modulate su quelle umane grazie ad un processo di computazione emotiva.

Su tali basi, la tesi del comportamentismo metodologico o dell'equivalenza performativa, tende a considerare i robot come soggetti di diritto in quanto dotati di uno status morale significativo dovuto all'emulazione dell'essere umano che ne è in possesso.

Le critiche a tale tesi seguono il detto secondo cui "l'abito non fa il monaco", infatti, il fatto che la macchina simuli l'essere umano, non significa che essa lo sia e dunque ciò non basta per riconoscerle soggettività giuridica. È importante valutare se l'apparato interno all'entità artificiale sia in grado di comprendere emotivamente e di essere consapevole di sé stesso come pensante. Tutto ciò deve inoltre essere empiricamente dimostrabile grazie ad un comportamento esteriorizzabile: gli stati mentali interni come la sensibilità, la consapevolezza e la coscienza fenomenica non possono essere conosciuti direttamente ma solo tramite la presenza di comportamenti

[316] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 3 (pag. 6 - 7) "Gli agenti intelligenti agiscono autonomamente, e sono perciò agenti liberi? La tesi funzionalista"

[317] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'"

esterni coerenti con tali stati. Infatti, analizzare l'algoritmo del sistema intelligente, sicuramente permetterebbe una conoscenza diretta dell'azione interna della macchina (per quanto la complessità e l'oscurità dello stesso non sempre lo consenta). Ma, a quel punto, si avrebbe una comprensione non di un vero e proprio stato mentale interno del robot ma della programmazione che ha permesso al sistema di simulare uno stato mentale umano all'interno di una macchina intelligente.

Secondo le moderne neuroscienze, il comportamento esteriore è fondamentale per fornire una prova della sussistenza di tali stati mentali.

Il comportamentismo metodologico accoglie i limiti umani dell'incapacità di comprendere gli stati mentali metafisici senza un riscontro empirico dato da test comportamentali che ne esteriorizzano l'esistenza. La capacità dei robot intelligenti di interagire anche a livello emotivo non significa che gli stessi provino emozioni coscienti ma solo simulate. La macchina si comporta come se capisse la realtà senza però essere in grado di riprodurre un'attività puramente umana. I robot simulano le emozioni senza provarle[318].

1.4 Tesi strutturalista – ontologica

La tesi strutturalista – ontologica si pone, al contrario, nell'ottica di una netta distinzione tra l'intelligenza artificiale e naturale considerando la prima come non autentica.

Il machine learning risulta basato esclusivamente su calcoli formali e consequenziali, ciò esprime sicuramente razionalità ma ciò non è sinonimo di intelligenza.

Si pone al centro del dibattito la struttura da cui dipendono le decisioni: materia biologica per l'essere umano e materia inorganica per il robot. La differenza si basa essenzialmente in questo, a prescindere da una possibile capacità della macchina di replicare il comportamento umano.

Secondo le scienze psicologiche, gli esseri umani, a differenza dei sistemi di intelligenza artificiale, non utilizzano modelli logico-formali-computazionali nelle decisioni che prendono: le loro azioni dipendono dalla percezione ed elaborazione emotiva dell'input esterno. Non si ha dunque, nell'essere umano una completa rappresentazione cosciente. Questo è dovuto alle limitate capacità cognitive umane nell'elaborare le informazioni che gli vengono fornite. L'essere umano non ha bisogno

[318] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 6 (pag. 10 - 13) "La questione dello statuto morale delle macchine umanizzate: il comportamentismo metodologico e la computazione emotiva"

di una completa conoscenza della realtà: esso si basa principalmente su procedure di tipo intuitivo, inconscio ed emotivo.

Sono infatti le emozioni a superare la lacuna subita dall'uomo a causa dei sui limiti cognitivi mentali. Grazie al sistema emozionale riesce a gestire situazioni complesse in base a scelte anche passionali-impulsive che spesso vincono sulle motivazioni razionali-coscienti: opposto a ciò che avviene in un sistema di intelligenza artificiale le cui decisioni sono minuziosamente calcolate, razionali e pianificate.

Le emozioni concorrono dunque con le altre informazioni di cui dispone la mente per prendere decisioni e agire in determinati modi esternando il comportamento.

La sfera emotiva è indispensabile per una razionalità umana funzionale alla conoscenza e alla decisione, al punto tale che se i sistemi affettivi ed emotivi risultano danneggiati, lo saranno anche i sistemi deliberativi. Non si può dunque parlare di intelligenza umana senza un essere umano. I robot non possiedono la dimensione emotiva e dunque il sistema cognitivo artificiale risulta fortemente diverso rispetto a quello di un essere vivente dotato di capacità intuitive. Le macchine sono dotate di mera capacità razionale e computazionale senza poter dunque ricomprendere un pensiero umano anche creativo e apparentemente illogico.

Il cervello umano risulta dunque libero da vincoli computazionali ed è in grado di correggere azioni automatiche nel caso in cui in concreto non risultino ottimali come in astratto, riuscendo ad uscire da schemi ripetitivi grazie alla sua spontaneità. L'essere umano possiede una volontà cosciente di azione e un libero arbitrio metafisico e al contempo empiricamente dimostrabile.

La mente umana risulta dunque privilegiata rispetto a quella della macchina: l'umano è libero e non replicabile pienamente da un sistema di intelligenza artificiale.

Le entità intelligenti possono essere considerate meramente come razionali e non come intelligenti, dato che hanno la capacità di muoversi solo se è possibile automatizzare il processo decisionale[319].

1.5 Paragone con la responsabilità degli enti

Come abbiamo accennato all'inizio di questo paragrafo, le società sono enti artificiali che possiedono una personalità giuridica e per questo, parte della dottrina ha ipotizzato la possibilità di realizzare una personalità elettronica dei sistemi di intelligenza

[319] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 7 (pag. 13 – 16) "La tesi strutturalista- ontologica. Le proprietà dell'intelligenza umana: pensiero logico e pensiero creativo intuitivo"

artificiale basata sulla stessa struttura della personalità giuridica degli enti: una fictio iuris utile per considerarli come soggetti di diritto titolari di diritti e doveri.

Analizzando la personalità giuridica delle società, è possibile notare come nell'essere un soggetto di diritto sia completamente ininfluente il possedere o meno una dignità. Infatti, le società possiedono sia la capacità giuridica che la capacità di agire pur essendo prive di dignità: esse non sono vive da un punto di vista naturalistico. Viene quindi a realizzarsi una finzione necessaria per raggiungere gli scopi sociali[320].

Che cosa significa essere reali? L'ente non esiste in un mondo materiale, in assenza dei suoi membri, a differenza di un robot che possiede la sua fisicità a prescindere dal soggetto che l'ha creato. Mentre, nel mondo giuridico, la società esiste in quanto persona giuridica a differenza dell'artefatto intelligente che non è dotato di personalità. Infatti, secondo tale teoria, è possibile effettuare un parallelismo tra le azioni illecite commesse dai robot intelligenti e il sistema della responsabilità penale degli enti[321].

I sostenitori di tale dottrina reputano opportuno estendere l'applicabilità di tale modello di responsabilità anche ad altre entità non umane. Come i sistemi di intelligenza artificiale, anche gli enti non hanno un corpo fisico umano e un'anima; nonostante ciò, sono comunque considerati soggetti di diritto capaci di porre in essere reati per il tramite delle persone fisiche che li compongono. I robot possiedono un corpo fisico che conosce l'ambiente tramite i sensori possedendo inoltre, a differenza degli enti giuridici, un elevato grado di autonomia dall'essere umano.

Per tali ragioni si ipotizza una responsabilità dell'agente artificiale che, similmente alla responsabilità amministrativa degli enti, renderebbe responsabile il sistema intelligente per le sue azioni e per quelle dell'utilizzatore o programmatore.

In realtà tale ideologia ha subito molte critiche dovute al fatto che, per quanto sia vero che esistano delle somiglianze tra l'ente e il robot, sussistono altrettante differenze, tra cui, la più importante: gli agenti intelligenti, a differenza delle società, non sono costituiti da persone fisiche che agiscono in loro nome[322]. Infatti, i sistemi basati sull'intelligenza artificiale, non coincidono con l'insieme di persone fisiche che si aggrega in una nuova formazione sociale capace di incidere economicamente nella

[320] S. QUINTARELLI, op. cit., Capitolo 6.2 (pag. 110 – 113) "Personalità giuridica. La convivenza tra uomini e software"

[321] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'"

[322] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 4 (pag. 8 - 9) "La punizione penale degli agenti non umani come la responsabilità degli enti giuridici"

società e a cui vengono riconosciuti diritti e doveri. Essi sono artefatti distinti e separati dall'umano che li ha realizzati e che li utilizza[323].

Mentre la società funge da schermo agli esseri umani che la compongono e la muovono tramite i loro compiti di rappresentanza, direzione, amministrazione e controllo, i sistemi intelligenti, possiedono una propria fisicità distinta da quella del loro ideatore potendo agire autonomamente. Infatti, le norme giuridiche riguardanti le società, in realtà, si riferiscono alle persone fisiche che la compongono[324], ed è loro che il decreto legislativo 231/2001[325] vuole colpire andando a considerare una responsabilità penale degli enti. [326]

Per questo, punire una macchina, per quanto intelligente essa sia, non avrebbe effetti dissuasivi nei confronti degli esseri umani che la programmano o la utilizzano dato che gli stessi sono del tutto estranei dal processo decisionale del robot[327]. Inoltre, le sanzioni elaborate per punire le società hanno una funzione deterrente in quanto incidono sul loro profitto, questione poco ipotizzabile per le sanzioni relative ai sistemi intelligenti, perché anche le sanzioni pecuniarie o di spegnimento della macchina andrebbero pur sempre a ricadere sul proprietario del robot nonostante lo stesso non possa influenzare in tempo reale il comportamento del sistema intelligente al fine di non commettere illeciti.

Continuando ad analizzare il decreto legislativo 231/2001[328], la responsabilità amministrativa da reato dell'ente è accertata sulla base di una fattispecie complessa, di cui il reato è solo un segmento. Il reato in questione è pur sempre quello commesso dalla persona fisica, un soggetto interno all'ente (apicale o sottoposto), a vantaggio o nell'interesse dell'ente[329]. Ad essere imputato è un fatto proprio dell'ente e la

[323] B. PANATTONI, op. cit., Paragrafo 2 "Dall'automazione tecnologica all'autonomia artificiale"

[324] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?"

[325] Decreto legislativo 8 giugno 2001, n. 231, "Disciplina della responsabilità amministrativa delle persone giuridiche, delle società e delle associazioni anche prive di personalità giuridica", a norma dell'articolo 11 della legge 29 settembre 2000, n. 300.

[326] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'"

[327] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 4 (pag. 8 - 9) "La punizione penale degli agenti non umani come la responsabilità degli enti giuridici"

[328] Decreto legislativo 8 giugno 2001, n. 231, "Disciplina della responsabilità amministrativa delle persone giuridiche, delle società e delle associazioni anche prive di personalità giuridica", a norma dell'articolo 11 della legge 29 settembre 2000, n. 300

[329] G. LATTANZI, P. SEVERINO, A. GULLO, "Responsabilità da reato degli enti", Volume I "Diritto Sostanziale", G. Giappichelli Editore, Torino 2020; Parte Seconda "La disciplina della responsabilità da reato delle persone giuridiche", M. PELISSERO, E. SCAROINA, V. NAPOLETANI, Capitolo 1 (pag. 71 – 171) "Principi generali"

colpevolezza rientra in una colpa di organizzazione propria della società[330] (l'ente è responsabile solo se non si è organizzato per evitare il reato)[331]; ma è sempre la persona fisica ad essere responsabile in ultima istanza. Quindi il parallelismo con la responsabilità penale degli enti non risulta idoneo a considerare i sistemi di intelligenza artificiale come soggetti di diritto[332].

2. Responsabilità penale personale: principi generali applicati all'intelligenza artificiale

Procedendo di questo passo nell'evoluzione tecnologica, a breve potremmo trovarci difronte a situazioni in cui un'entità artificiale intelligente ponga in essere un fatto antigiuridico e colpevole che, se commesso da un essere umano, sarebbe considerato penalmente rilevante[333].

Tuttavia, si è davanti ad un notevole scetticismo su una possibile disciplina normativa ad hoc per i reati commessi dall'intelligenza artificiale. La stessa critica era stata avanzata nei confronti dei computer crimes all'epoca della loro introduzione: considerati utili come una "law of the horse"[334]. Ciò perché il diritto penale risulta connesso esclusivamente all'essere umano e non alle res. In tal modo, però, si va ad assimilare la robotica ad altre forme di tecnologia ritenendo sufficienti le norme di diritto penale oggi in vigore su di esse. Secondo tale visione, una normativa specifica per i robot crime viene dunque considerata superflua per l'ordinamento: un nuovo fenomeno scientifico o tecnologico non sempre necessita di nuove norme, essendo le attuali dotate di notevole elasticità da potersi adattare all'innovazione.

In tale ottica, si inserisce l'interazione dell'intelligenza artificiale al Diritto Penale.

---

[330] A. CAPPELLINI, "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale", op. cit., Paragrafo 7 (pag. 16 - 18) "Terza critica: fallacia del parallelo con la corporate liability"

[331] G. LATTANZI, P. SEVERINO, A. GULLO, op.cit., Parte Seconda "La disciplina della responsabilità da reato delle persone giuridiche", M. PELISSERO, E. SCAROINA, V. NAPOLETANI, Capitolo 1 (pag. 71 – 171) "Principi generali"

[332] B. PANATTONI, op. cit., Paragrafo 5.1 "Forme di responsabilità diretta a carico dell'agente artificiale"

[333] I. SALVADORI, op. cit., Paragrafo 4, "Gli agenti artificiali come autori di un reato"

[334] F.H. EASTERBROOK, "Cyberspace and the Law of the Horse", University of Chicago Legal Forum, 1996, 2017:
https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=2147&context=journal_articles
Con il termine "law of the horse" si andava a paragonare internet ai cavalli i quali possono essere oggetto di compravendita, di obbligazioni risarcitorie in caso di commissione di lesioni e di obblighi di diligenza da parte di veterinari e fantini, ma questo non risulta sufficiente per realizzare un'autonoma branca del diritto incentrata esclusivamente sui cavalli.

L'IA ha la capacità di incidere su fattori essenziali del diritto penale: nesso di causalità tra gli eventi, rapporti mezzo-fine e agente-strumento nella definizione dei comportamenti penalmente rilevanti.

Il "mezzo" risulta essere un elemento, privo di autonomia e discrezione, che contribuisce a causare un evento: è l'agente, nelle cui mani viene posto lo strumento, colui che realizza le azioni che condurranno al fatto illecito. Fino ad oggi, le tecnologie sono sempre state considerate come mezzi, reputando responsabile esclusivamente la persona fisica che utilizzandole ha realizzato l'evento. Con l'evoluzione scientifica, tale assunto viene posto in crisi dato che l'intelligenza artificiale risulta avere un grado di autonomia che le permette di prendere decisioni rilevanti per gli esseri umani[335]. Infatti, secondo la Risoluzione del Parlamento Europeo del 16 febbraio 2017 concernente le norme di diritto sulla robotica, "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 (...) 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; (...) 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"[336].

Si assiste a tale mutazione, da strumento a soggetto, tramite due modalità principali: diretta e indiretta. Nel primo caso sono gli agenti umani a chiedere ai sistemi automatizzati di prendere decisioni per loro conto; nel secondo caso, invece, le conoscenze alla base delle decisioni prese dagli individui sono fornite da strumenti tecnologici spesso automatizzati (il web consente a ciascuno un accesso immediato a un numero elevato di fonti di informazione)[337].

Ad oggi, ancora si tende ad escludere che i sistemi di intelligenza artificiale possano essere considerati come penalmente responsabili per i reati commessi: il diritto penale risulta fortemente antropocentrico. Anche se taluni sostengono l'esatto opposto, considerando possibile un rimprovero doloso o colposo alla macchina, potendo, la

[335] A. SIMONCINI, op. cit., Paragrafo 3 (pag. 67 - 69) "L'impatto della rivoluzione cibernetica su diritto costituzionale: intelligenza artificiale, autonomia e libertà"

[336] RISOLUZIONE DEL PARLAMENTO EUROPEO DEL 16 FEBBRAIO 2017 recante "Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica" (2015/2103(INL)), op cit.

[337] A. SIMONCINI, op. cit., Paragrafo 3.1 (pag. 69 – 71) "Il soggetto "catturato" dallo strumento"

stessa, realizzare tutti gli elementi oggettivi e soggettivi del reato: in tal modo si viene a garantire, al robot, una personalità elettronica che si riflette anche nel diritto penale. L'ultima tesi risulta fortemente innovativa e sussistono tutti i presupposti per poter pensare ad un futuro in cui la macchina possa avere piene capacità di agire in modo consapevole, ma ad oggi, gli agenti artificiali non possiedono una piena libertà di autodeterminazione tanto che neanche i sistemi intelligenti più evoluti hanno un'autocoscienza che gli permette di comprendere le norme penali e il disvalore delle loro azioni[338].

Si è già parlato dei dubbi in merito alla compatibilità tra un sistema basato sull'intelligenza artificiale come responsabile per un reato commesso ed il principio della personalità di cui all'art. 27 della Costituzione. Il principio di cui parla l'articolo necessita di essere riferito ad un essere umano a cui può essere psicologicamente attribuito l'illecito in modo tale da poterlo rieducare attraverso la pena (art. 27 comma 3 della Costituzione). Questo non può riguardare i robot, privi di una coscienza e dunque non potrebbe comprendere l'illiceità della sua azione e l'irrogazione di una sanzione penale risulterebbe superflua non riuscendo a garantire le funzioni necessarie della pena (prevenzione, speciale o generale, e rieducazione) le quali si basano su caratteristiche tipiche degli stati mentali umani.

Tesi avanguardiste (come quella di Hallevy) tendono a riconoscere una responsabilità diretta dell'intelligenza artificiale fornendo un ragionamento simile a quello realizzato in relazione allo sgretolamento del brocardo societas delinquere non potest.

Secondo tale approccio, i robot possiedono la capacità di porre in essere azioni penalmente rilevanti grazie ad una loro fisicità che gli permette di muoversi nello spazio e dunque di commettere illeciti. Viene quindi garantita la possibilità dell'azione intesa come elemento del fatto tipico. Per quanto concerne l'elemento soggettivo, risulta possibile ravvisare, in capo al sistema intelligente, la capacità di cognizione e di volizione. Grazie alla sua abilità di acquisire, rappresentare e rielaborare dati raccolti dall'ambiente esterno apprendendo da essi tramite le tecniche di machine learning è possibile ipotizzare la sussistenza della cognizione. Per quanto concerne la volizione, invece, la stessa sussisterebbe dal momento in cui l'intelligenza artificiale ha sviluppato la capacità di valutare la probabilità di verificazione di un evento e agire di conseguenza. Infine, in relazione alle funzioni della pena, risulta, per tale tesi, la possibilità di ritrovare gli scopi della riabilitazione e della prevenzione anche in

[338] I. SALVADORI, op. cit., Paragrafo 4, "Gli agenti artificiali come autori di un reato"

sanzioni relazionate con l'intelligenza artificiale. Tutto ciò grazie alla possibilità di imporre una sanzione detentiva, cioè una sanzione che limiti la libertà di agire dell'agente artificiale intelligente e grazie alla possibilità di rieducare il sistema andando a correggere direttamente il programma di funzionamento della macchina[339].

A questo punto, si possono analizzare le componenti del reato in relazione ai fatti commessi dall'intelligenza artificiale in base alla concezione bipartita di Antolisei costituita dall'elemento oggettivo (condotta ed evento) e soggettivo (dolo o colpa), in base alla visione tripartita secondo cui si necessita una valutazione in relazione al fatto, all'antigiuridicità e alla colpevolezza, oppure in base alla concezione quadripartita che va ad aggiungere alla moderna teoria tripartita, anche la punibilità.

Il fatto ricomprende tutti gli elementi oggettivi che individuano e caratterizzano ogni singolo reato come specifica forma di offesa a uno o più beni giuridici. Esso comprende: la condotta (attiva o omissiva) e i sui presupposti (situazioni che devono preesistere alla condotta), l'evento inteso come gli accadimenti causati dalla condotta ma separati da essa, il rapporto di causalità, l'oggetto materiale e l'offesa.

L'antigiuridicità riguarda il rapporto di contraddizione tra il fatto e l'ordinamento.

La colpevolezza invece contiene in sé i requisiti dai quali dipende la possibilità di muovere all'agente un rimprovero per aver commesso il fatto antigiuridico. Tali requisiti, che devono essere riferiti al caso concreto, comprendono il dolo, la colpa o il dolo misto a colpa, l'assenza di scusanti, la conoscenza della legge penale e la capacità di intendere e di volere.

Infine, la punibilità consiste nell'insieme di eventuali condizioni, ulteriori ed esterne rispetto al fatto antigiuridico e colpevole che fondano o escludono l'opportunità di punirlo[340].

Analizziamo quindi gli elementi del reato alla luce della possibile commissione dello stesso da parte dell'intelligenza artificiale.

2.1 Elemento oggettivo: condotta, evento, rapporto di causalità

Seguendo la concezione quadripartita (come anche la tripartita) del reato, l'elemento oggettivo dello stesso, come accennato precedentemente, comprende tre requisiti che

[339] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'"

[340] G. MARINUCCI, E. DOLCINI, G.L. GATTA, "Manuale di diritto penale. Parte generale", Ottava edizione, Giuffrè Francis Lefebvre S.p.a. Milano – 2019; Sezione III "Il Reato", Capitolo V (pag. 207 – 225) "Analisi e sistematica del reato"

devono sussistere per considerare quella determinata situazione come fatto illecito penale.

Il primo è la condotta, consistente nell'azione (o omissione) manifestata esteriormente e (fino ad ora) considerata esclusivamente umana. Per azione si intende un movimento del corpo del soggetto percepibile all'esterno ed essa risulta essere il contrario dell'omissione (un non facere) penalmente rilevante limitatamente al caso in cui sussista una violazione di un obbligo giuridico di agire.

Il secondo è l'evento, inteso come conseguenza della condotta e legato ad essa tramite il rapporto di causalità. L'evento è dunque separato dall'azione ma dipendente da essa risultandone l'effetto. Esso consiste in una modificazione della realtà fisica, psichica o economico-giuridica e tendenzialmente considerato dipendente da un comportamento umano.

Il rapporto di causalità risulta fondamentale per attribuire un evento all'individuo. L'evento deve verificarsi come conseguenza della sua azione o omissione (art. 40 c.p.). Sussistono diverse teorie connesse alla causalità: la teoria condizionalista, la teoria della causalità adeguata, della causalità umana, dell'imputazione oggettiva dell'evento (correttivi della teoria condizionalista).

La teoria condizionalista (anche detta teoria dell'equivalenza causale) si basa sul principio della "conditio sine qua non": l'azione A è causa dell'evento B, se può dirsi che senza A, tenuto conto di tutte le circostanze del caso concreto, l'evento B non si sarebbe verificato. Si considera causa ogni condizione dell'evento senza il quale l'effetto non si sarebbe verificato: basta che il soggetto abbia posto in essere un antecedente necessario per il verificarsi del risultato. I rischi sono legati ad un'eccessiva estensione del concetto di causa finendo per ricadere in un regressus ad infinitum; quindi bisogna valutare attentamente il caso concreto e la volontà dell'azione lesiva. Bisognerà dunque seguire il procedimento di eliminazione mentale: escludere un elemento e vedere se l'evento si sarebbe verificato comunque, se ciò non avviene, allora quella è la causa. Per riempire di significato il procedimento di eliminazione mentale è necessario seguire le leggi scientifiche (universali o statistiche con carattere probabilistico anche a seguito dei principi espressi dalla Sentenza 30328/2002)[341]: dunque, grazie a tale correttivo, è possibile affermare che è causa dell'evento ogni azione che, tenendo conto delle circostanze del caso concreto, non

[341] Sentenza della Cassazione Penale, Sezioni Unite 30328/2002 (Franzese): https://www.giurisprudenzapenale.com/wp- content/uploads/2014/01/Cass-Pen-Sez-Un-Franzese-2002.pdf

può essere eliminata mentalmente sulla base di leggi scientifiche senza che l'evento venga meno. Nel caso in cui sussistano cause sopravvenute sufficienti, non è ravvisabile la causalità con questo evento.

Quindi: il concorso di fattori causali preesistenti, simultanei e sopravvenuti, non esclude il rapporto di causalità quando l'azione è necessaria per l'evento; il rapporto di causalità non è escluso neanche se il fattore causale ulteriore consiste in un'azione illecita di un terzo; e il rapporto di causalità è invece escluso se la causa è autonoma rispetto all'azione. I primi due corollari esposti sono ritrovabili anche nel primo e terzo comma dell'articolo 41 c.p. mentre sussistono controversie in relazione all'articolo 41 cpv. c.p. secondo cui "le cause sopravvenute escludono il rapporto di causalità quando sono state da sole sufficienti a determinare l'evento" anche se, in realtà, potrebbe rientrare nel terzo corollario della teoria condizionalista, infatti, la causa sopravvenuta così descritta si sarebbe inserita tra l'azione e l'evento andando a far si che l'azione rappresenti un mero antecedente temporale e non una conditio sine qua non. Così interpretata, tale teoria rispecchia pienamente il dettato dell'articolo 41 c.p.

La teoria della causalità adeguata si basa sullo schema secondo cui: l'azione A è causa dell'evento B quando, senza l'azione A, l'evento B non si sarebbe verificato e inoltre, l'evento B rappresenta una conseguenza prevedibile dell'azione A. Dunque, si necessita di un individuo che provoca un evento con un'azione adeguata e idonea a provocare quell'evento in base al principio "id quod prelunque acidit". Viene quindi fatto un rinvio alla statistica e alle regole dell'esperienza in base ad una valutazione ex ante secondo cui non sono causati dall'uomo gli effetti straordinari e atipici dell'azione umana. Tale teoria ha il rischio di restringere eccessivamente la responsabilità penale generando anche un'incertezza applicativa dovuta al fatto di contenere in sé concetti (come l'adeguatezza) che risultano fortemente ambigui.

La teoria della causalità umana ha come esponente Antolisei e segue una struttura basata sul fatto che: l'azione A è causa dell'evento B quando, senza l'azione A, l'evento B non si sarebbe verificato e inoltre il verificarsi dell'evento B non è dovuto al concorso di fattori eccezionali. Tale teoria è connessa alla capacità dell'essere umano di valutare gli effetti delle proprie azioni: una sfera di signoria dell'uomo che può dominare i suoi comportamenti. Quindi, come per la teoria della causalità adeguata, l'uomo deve aver posto in essere almeno una condotta senza che siano, inoltre, intervenuti fattori straordinari tali da interrompere il nesso causale con l'azione dell'essere umano. Ma, a differenza della teoria appena menzionata, il rapporto di

causalità viene escluso non ogni qualvolta sussistano fattori anomali, ma solo nel caso in cui tali fattori causali siano estremamente rari. Dunque, si va ad ampliare il campo del penalmente rilevante anche agli sviluppi dell'azione che l'essere umano poteva dominare (escludendo, come appena detto, solo i fattori causali di minima realizzabilità).

La teoria dell'imputazione oggettiva dell'evento va, invece, a restringere il campo della teoria condizionalistica nell'ipotesi di un decorso causale atipico andando ad inserire un requisito ulteriore rispetto alla causalità. L'evento generatosi dall'azione potrebbe essere imputato al soggetto a due condizioni: che l'agente, con la sua condotta antigiuridica, abbia creato o non impedito il rischio di verificazione di un evento e che l'evento sia la concretizzazione del rischio che la regola cautelare violata mirava ad evitare.

A ciò bisogna aggiungere il fatto che l'azione e l'evento debbano incidere sulla persona o res nominata dalla norma come oggetto del reato. L'offesa deve riguardare un bene giuridicamente tutelato come individuale (riferito alle singole persone fisiche), collettivo (facente capo alla generalità dei consociati o allo Stato ed enti pubblici), strumentale (la cui integrità è strumento per la sopravvivenza di beni superiori) o finale (protetto dai beni strumentali)[342].

Fin qui, abbiamo descritto gli elementi oggettivi tipici del reato in relazione all'essere umano, in quanto, fino ad ora, gli stessi si sono sempre e solo riferiti ad una persona fisica titolare di soggettività giuridica e penale. Ma, come abbiamo precedentemente affermato, i sistemi basati sull'intelligenza artificiale sono anch'essi capaci di porre in essere azioni. È possibile affermarlo grazie al fatto che anche i robot sono dotati di una fisicità simile a quella umana e dunque hanno l'abilità di muoversi nello spazio circostante (si pensi ad un braccio robotico). L'azione è intesa come il movimento (o l'assenza di movimento) del corpo, ma, secondo Beling (giurista tedesco di fine '800), questa deve essere attribuita ad un soggetto umano il quale abbia la signoria del proprio corpo. Dunque, non basta la mera fisicità e capacità di movimento per considerare l'azione del robot come penalmente rilevante, si necessita pur sempre della volontarietà dell'azione.

Da un punto di vista meccanico, il sistema intelligente agisce nel mondo fisico: bisogna però valutare la volontarietà. Il problema sussiste nel fatto che la macchina intelligente

[342] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VI "Il Fatto", Sottocapitolo A (pag. 225 – 257) "Il fatto nei reati commissivi"

non agisce ma "è agita", per quanto sia vero che essa rielabori i dati della realtà esterna tramite tecniche di machine learning, le decisioni a cui giunge e i movimenti che realizza sono esclusiva espressione dell'algoritmo che la governa: il sistema artificiale risulta sprovvisto di una piena libertà di autodeterminazione (differente dall'essere umano). Il decision-making è il risultato obbligato del c.d. black box algorithm che lo governa.

Non sussistendo una piena autodeterminazione, considerare la capacità di azione della macchina come penalmente rilevante risulta estremamente complesso dato che, ad oggi, non sussistono sistemi di intelligenza artificiale che siano in grado di emulare pienamente i meccanismi dell'intelligenza umana. La macchina, a differenza dell'essere umano, non può agire differentemente rispetto a quanto dettato dall'algoritmo: l'essere umano risulta tendenzialmente libero di esprimersi e di muoversi senza condizionamenti limitativi del suo comportamento posti a priori[343].

È necessario, quindi, focalizzarci attentamente sull'elemento soggettivo del reato declinato per i sistemi di intelligenza artificiale.

2.2 Elemento soggettivo: la colpevolezza. Coscienza e volontà

Dobbiamo ora analizzare la colpevolezza come elemento del reato anche alla luce delle implicazioni dovute all'avvento dell'intelligenza artificiale.

Tale requisito sembra comportare gravi ostacoli ad un pieno riconoscimento della responsabilità dei sistemi intelligenti. Infatti, come analizzeremo di seguito, la colpevolezza necessita di un coinvolgimento soggettivo del reo al fine di muovergli un rimprovero penalmente rilevante. La soggettività psicologica è sempre stata riferita all'essere umano, come anche la capacità di intendere e di volere e dunque, tradizionalmente, solo un uomo può essere considerato colpevole. Di recente, però, parte della dottrina ha iniziato a riconoscere tale elemento anche in capo all'intelligenza artificiale date le sue abilità di emulazione del comportamento e ragionamento umano grazie alle sviluppate tecniche di machine learning. I robot stanno diventando sempre più autonomi e capaci di imparare dalle loro esperienze, non limitandosi ai binari della programmazione inizialmente impartitagli dato che, spesso, le decisioni assunte dai sistemi più evoluti, non sono neanche prevedibili dai programmatori.

[343] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'".

Su queste basi si fonda la prospettiva di una possibile responsabilità penale in capo alle macchine intelligenti[344].

La colpevolezza è l'elemento soggettivo-psicologico del reato: per integrare un reato, oltre al fatto materiale, serve il concorso della volontà; infatti il brocardo latino recita "nulla poena sine culpa".

Tale requisito è talmente importante per il nostro ordinamento da essere sancito anche a livello costituzionale all'articolo 27 comma 1: "la responsabilità penale è personale". Il moderno principio di colpevolezza, nato con la Sentenza 364/1988[345], si basa sul concetto secondo cui la commissione del reato deve essere personalmente rimproverata all'autore: il fatto deve quindi essere personalmente colpevole. In base a tale concezione, l'articolo 27 comma 1 della Costituzione viene così ad essere analizzato: divieto di responsabilità penale per fatto altrui (a differenza del diritto civile); esclusione della responsabilità oggettiva che discende dalla semplice realizzazione di un evento anche in assenza dell'elemento soggettivo-psicologico; e responsabilità penale personale intesa come colpevole, in tal modo viene posta la colpevolezza come principio garantistico di legalità e di rango costituzionale.

I requisiti della colpevolezza sono quattro: l'elemento psicologico in senso stretto (il dolo, la colpa e il dolo misto a colpa); l'assenza di scusanti che eliminano l'elemento soggettivo; la conoscenza o conoscibilità della legge penale violata; e l'imputabilità concepita come la capacità di intendere e di volere.

La colpevolezza ha quindi un carattere psicologico che rimanda al nesso psichico tra fatto e soggetto agente e uno normativo legato al rapporto di contraddizione tra la volontà del soggetto e la norma giuridica[346]. Prima di parlare dell'elemento psicologico (dolo, colpa e dolo misto a colpa) è importante soffermarsi sul concetto di coscienza e volontà dell'azione. L'articolo 42 primo comma c.p. stabilisce il principio secondo cui nessuno può essere punito per un fatto previsto dalla legge come reato se non l'ha commesso con coscienza e volontà. Secondo la tradizionale tesi di Antolisei, per affermare la sussistenza di tale elemento, risulta necessario valutare a priori la suitas (attribuibilità) del fatto al suo autore. Esistono infatti atti automatici (istintivi, riflessi o abituali) che la volontà umana non può interamente controllare: essi sfuggono, in

[344] F. BASILE, op. cit., Paragrafo 6.3.3 (pag. 30 – 31) "Una colpevolezza "disumana"?"

[345] Sentenza della Corte Costituzionale n. 364, del 23 marzo 1988: https://www.cortecostituzionale.it/actionSchedaPronuncia.do?anno=1988&numero=364

[346] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VIII "La Colpevolezza", Paragrafo 1 (pag. 349 – 352) "La colpevolezza: nozione, fondamento e rilevanza costituzionale"

determinate condizioni, alla gestione dell'essere umano. Infatti, secondo Antolisei, bisogna risalire alla suitas e richiamare le capacità di controllo umano. In base a tale concezione, dunque, esistono atti automatici che possono essere controllati grazie ad un nisus cosciente (sforzo dell'energia interiore) potendo rispettare il concetto di coscienza e volontà di azione. La dottrina moderna non parla più di suitas, ma richiama le funzioni di controllo dell'uomo sulle sue azioni ed omissioni. Nonostante la diversità di concetti, la sostanza rimane simile alla dottrina tradizionale: si fa riferimento alle scusanti (in assenza delle quali si è colpevoli) ossia circostanze interne e anomale che scusano la violazione di regole rendendo inesigibile un comportamento diverso da quello tenuto. Esse sono dovute a reazioni di spavento e terrore che paralizzano le funzioni di controllo della coscienza e volontà oppure a atti umani che sfuggono totalmente al controllo dell'essere umano (come ad esempio un movimento involontario del braccio, dovuto crisi epilettica, che genera una lesione ad un altro soggetto)[347].

È importante ricordare come la Dichiarazione Universale dei diritti dell'Uomo del 1948 stabilì per la prima volta una piena simmetria tra l'entità (che possiede l'insieme dei diritti e delle libertà) e l'essere umano connettendo ad esso i requisiti di coscienza e volontà. All'articolo 1 della Dichiarazione si legge: "Tutti gli esseri umani sono nati liberi e uguali in dignità e diritti. Loro sono dotati di ragione e coscienza" e all'articolo 2: "A ogni individuo spettano tutti i diritti e le liberà enunciati nella presente Dichiarazione, senza distinzione alcuna per ragioni di razza, colore, sesso, lingua, religione, opinione politica o di altro genere, origine nazionale o sociale, ricchezza, nascita o altra condizione."

L'affermazione secondo cui tutti gli esseri umani sono dotati di ragione e coscienza sembra supporre che le stesse facciano parte dell'essenza di tutti gli individui senza le quali perderebbero il carattere dell'umanità. Ma, affermando questo, significherebbe dire che un essere umano privo delle suddette qualità non potrebbe essere considerato tale e dunque non avrebbe gli stessi diritti e dignità degli altri[348].

La soggettività giuridica è sempre stata attribuita agli esseri umani ma non è detto che la stessa non possa essere estesa anche alle entità artificiali (come avvenuto per gli animali ai quali sono stati riconosciuti alcuni diritti in relazione alla loro dignità di

[347] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VIII "La Colpevolezza", Sotto-paragrafo A (pag. 352 – 423) "Dolo, colpa e dolo misto a colpa"

[348] A. SANTOSUOSSO, op. cit., Paragrafo 7.11 (pag. 198 – 202) "Diritti, storicità, artificialità – Creature artificiali e proprietà ontologiche"

esseri senzienti dotati di emozioni e desideri). Dunque l'intelligenza non è necessaria per il riconoscimento dei diritti (al massimo la sua assenza potrà rilevare come inesigibilità dei doveri)[349].

Bisogna però sottolineare che lo stato della conoscenza all'epoca della Dichiarazione Universale era limitato e l'obiettivo del testo era quello di porre il confine degli animali non umani, che costituivano il limite oltre il quale l'umanità cessava di esistere. La rivendicazione dei diritti per gli animali avverrà nei decenni successivi e non era immaginabile durante la stesura della Dichiarazione. Infatti, grazie alle scoperte neuro-scientifiche a proposito dell'attività mentale degli animali non umani è stato possibile negare che solo gli esseri umani posseggono il sostrato neurologico che genera la coscienza dato che anche gli animali non umani hanno le capacità di mostrare comportamenti intenzionali. Infine, con le varie scoperte scientifiche è stato possibile affermare che non esiste un solo tipo di coscienza i cui gradi possono essere misurati secondo un'unica scala: esistono varie tipologie di coscienza che si articolano in modo differente tra gli animali umani e non[350].

Relazionando tali principi all'intelligenza artificiale è possibile notare come la stessa sia in grado di agire e di muoversi nello spazio ma la sua fisicità è governata da un algoritmo che a priori le impone di muoversi all'interno di binari delimitati (anche nel caso di sistemi intelligenti molto evoluti e dotati di autoapprendimento). Difetta, quindi, di quella signoria che gli esseri umani hanno sul proprio corpo e dunque della suitas posta come condizione di pre-tipicità del fatto: ad oggi, non sono state ancora realizzate macchine intelligenti dotate di una piena capacità di colpevolezza[351].

Ai sistemi intelligenti mancano le proprietà della coscienza cognitiva e fenomenica.

La coscienza cognitiva permette l'accesso dell'essere umano ai propri stati mentali e al proprio comportamento in relazione alle informazioni che gli giungono dall'esterno. Ciò avviene grazie ai processi di formazione dei ragionamenti, delle credenze e riflessioni limitate all'aspetto della rappresentazione. Serve quindi ad organizzare e pianificare l'azione, a controllare gli impulsi e a costruire i modelli per le interazioni sociali. Essa solitamente non è consapevole a causa della complessa architettura mentale dell'essere umano che permette la rappresentazione di stati mentali impliciti

[349] S. QUINTARELLI, op. cit., Capitolo 6.2 (pag. 110 – 113) "Personalità giuridica. La convivenza tra uomini e software"

[350] A. SANTOSUOSSO, op. cit., Paragrafo 7.11 (pag. 198 – 202) "Diritti, storicità, artificialità – Creature artificiali e proprietà ontologiche"

[351] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'"

e non consapevoli che però si riflettono sul sistema cognitivo: si parla quindi di inconscio cognitivo. La coscienza cognitiva non risponde a logiche computazionali ma ad una dimensione emotiva ed esperienziale potendosi verificare fuori dalla consapevolezza avendo ricadute importanti sulle nostre attività consapevoli (il c.d. unconscious will). La coscienza fenomenica esprime le rappresentazioni mentali delle esperienze e sensazioni vissute dal soggetto in prima persona. Essa sfugge ad una logica oggettiva dato che le sensazioni soggettive non sono accessibili ad un osservatore terzo esistendo solo in prima persona e potendo essere parzialmente comprese da un terzo solo attraverso un racconto ma senza una possibile osservazione diretta.

I robot godono della capacità di accedere alle informazioni (e quindi della coscienza cognitiva) grazie alla loro abilità di conoscere e memorizzare grandi quantità di dati. Ma questa attività non può essere qualificata come sistema cognitivo soggettivo dato che l'accesso avviene in modo automatico e meccanico senza una forza emotiva-sensoriale. Dunque, sussiste in capo ai sistemi intelligenti un'evoluta capacità di accesso ma scollegata dalle emozioni non avendo un'architettura mentale idonea. A differenza delle macchine, le azioni degli esseri umani sono sempre connesse a stati mentali da lui provati, anche a livello inconscio.

In relazione alla coscienza fenomenica, bisogna affermare che la stessa sussiste solo in capo a chi ha vissuto dato che le caratteristiche interne dell'essere in quanto tale non possono essere analizzate oggettivamente ma esclusivamente nella dimensione soggettiva. Ai robot manca tale dimensione: la coscienza dei significati dell'esperienza[352]. Si può affermare che i sistemi basati sull'intelligenza artificiale non sono coscienti dell'essere coscienti, possedendo una capacità di organizzazione funzionale senza avere la consapevolezza del proprio agire[353].

Questo viene confermato anche dal test della stanza cinese ideato da Searle, filosofo statunitense, per confutare la tesi di Turing – di cu si è già parlato e che qui si riporta – : ipotizzando di porre all'interno di una stanza un soggetto madrelingua inglese che non conosce il cinese e all'esterno un soggetto madrelingua cinese che non conosce l'inglese; al soggetto posto all'interno della stanza verranno fornite istruzioni (in

[352] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 8 (pag. 16 - 19) "La coscienza artificiale e i limiti della computazione emotiva degli agenti artificiali umanizzati"

[353] M. B. MAGRO, "Biorobotica, robotica e diritto penale", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali"

inglese) su come utilizzare gli ideogrammi cinesi; grazie alle istruzioni il soggetto riuscirà a comunicare all'esterno in cinese pur non conoscendone la lingua. Searle paragona il soggetto inglese al computer che non conosce realmente il mondo esterno ma riesce ad elaborare gli input che gli vengono forniti grazie al programma in lui immesso (paragonato alle istruzioni fornite al soggetto inglese) generando l'output desiderato (comunicare in cinese con il soggetto all'esterno), dunque, non è importante che il computer abbia coscienza di ciò che sta compiendo, lui non sta realmente imparando ma solo eseguendo meccanicamente un algoritmo di elaborazione dei dati. In tal modo, Searle dimostra come l'esecuzione di un programma su un dato input non sia sufficiente per l'effettiva intenzionalità di un'azione. Un computer ha l'abilità di manipolare simboli ma ciò non significa comprendere: si comporta come se capisse senza, però, possedere l'attività celebrale umana. La simula. Gli stati mentali umani, però, non possono essere ridotti a processi computazionali.

Inoltre al robot, per quanto evoluto, mancano le conoscenze date dal senso comune che gli uomini possiedono senza aver affrontato studi particolari. Il senso comune permette all'uomo di risolvere in modo elastico i problemi che gli vengono posti e ciò si oppone alla forte rigidità del ragionamento algoritmico della macchina intelligente. Questo viene definito intelligenza emotiva; ad oggi sconosciuta dai robot. Esistono ricerche finalizzate allo scopo di fornire, ai dispositivi intelligenti, emozioni artificiali per renderli simili all'essere umano: passare dalla progettazione di macchine che simulano le emozioni, a sistemi che le provano. Le emozioni sono una caratteristica fondamentale degli esseri umani e sono fortemente connesse alla razionalità e alle funzioni fisiologiche; forniscono un'identità all'individuo. Nelle macchine, tali stati mentali sono solamente artificiali; tramite la computazione emotiva, i robot studiano gli stati emotivi umani attraverso i dati biometrici connessi a comportamenti manifesti dei soggetti, in questo modo il sistema sceglierà la migliore modalità di interazione con l'utente in base al suo stato emotivo[354].

Quindi, il problema del riconoscimento di una responsabilità penale dell'intelligenza artificiale ed ancor prima del riconoscimento di una colpevolezza in capo ai sistemi intelligenti risulta connesso all'individuazione degli stati mentali nei robot dato che la colpevolezza ha un fondamento anche di tipo psichico. Per garantire il rispetto dell'articolo 27 della Costituzione è necessario che il reo (anche artificialmente inteso)

[354] M. B. MAGRO, "Biorobotica, robotica e diritto penale", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali"

sia in grado di percepire e comprendere la sua condotta grazie anche agli stati mentali tipici dell'architettura psichica umana in modo tale da essere capace di una risposta emotiva agli eventi che accadono[355].

In base a questo approccio un sistema intelligente potrebbe ottenere soggettività giuridica e capacità di agire superata una certa soglia di coscienza emotiva[356].

Ad oggi (ma l'evoluzione è fortemente rapida) è, dunque, possibile negare che gli stati mentali possano essere ridotti a processi meccanici di tipo logico-computazionale: il robot ha una capacità performativa superiore a quella umana ma agisce in modo completamente diverso da esso in quanto pecca di emotività[357].

Passando all'analisi dei requisiti della colpevolezza, come esplicato poche righe sopra, l'elemento psicologico penalmente rilevante può essere suddiviso in tre categorie: dolo, colpa e dolo misto a colpa.

Il dolo è la forma tipica della volontà colpevole: una piena ribellione alla legge e dunque comporta la forma più grave di responsabilità penale.

L'articolo 43 comma 1 c.p. definisce il delitto doloso (o secondo l'intenzione) come l'evento dannoso o pericoloso considerato tale dalla legge e che è il risultato dell'azione od omissione preveduto e voluto dall'agente come conseguenza della propria azione od omissione. Le caratteristiche sono due: la previsione e la volontà dell'evento. I momenti costitutivi del dolo sono quindi anch'essi due: il momento rappresentativo (o conoscitivo) e il momento volitivo. Il primo riguarda la fase in cui il soggetto agente si rappresenta il fatto realizzandosi una visione anticipata dello stesso; dunque, il soggetto ha chiaro il fatto antigiuridico da porre in essere conoscendo effettivamente (non una mera conoscenza potenziale) tutti gli elementi (descrittivi-naturali e normativi) rilevanti del caso concreto valutato al momento di inizio dell'esecuzione della fattispecie tipica. Il secondo, invece, sussiste nella reale volontà del soggetto agente di commettere quell'evento e dunque la realizzazione del fatto antigiuridico. Questo deve sussistere al momento della condotta (non basta che sia successivo o antecedente). Esistono tre tipologie di dolo che si ricavano dal momento volitivo. Il dolo intenzionale, esistente nel momento in cui il soggetto agisce proprio

[355] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 5 (pag. 9 - 10): "Il problema della colpevolezza degli agenti artificiali intelligenti"

[356] S. QUINTARELLI, op. cit., Capitolo 6.2 (pag. 110 – 113) "Personalità giuridica. La convivenza tra uomini e software"

[357] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 8 (pag. 16 - 19) "La coscienza artificiale e i limiti della computazione emotiva degli agenti artificiali umanizzati"

con lo scopo di realizzare quel fatto e quell'evento. Non è necessario che il fatto sia il fine ultimo della condotta (può essere anche intermedio) e basta la minima possibilità di successo. Il dolo diretto, invece, si ha quando il soggetto agente non persegue la realizzazione del fatto ma si rappresenta come certo o probabile (al limite della certezza) il verificarsi dell'evento. Il dolo eventuale (o indiretto), infine, risulta essere il più controverso e problematico; è il dolo di minore intensità, non molto distante dal grado più elevato della colpa (colpa cosciente) distinguendosi da essa in base al fatto che nel dolo eventuale si ha l'accettazione del rischio di verificazione dell'evento (pur di non rinunciare all'azione e ai vantaggi collegati, il soggetto agente accetta che l'illecito possa verificarsi) seguendo la seconda formula di Frank "avvenga questo o quest'altro, io agisco comunque".

L'accertamento del dolo risulta problematico in base al fatto che non si può entrare nella psiche del soggetto e quindi, sono importanti i comportamenti esteriori analizzati secondo le massime di comune esperienza applicate al caso concreto relative alle modalità della condotta (mezzi adoperati, durata della condotta e condotta successiva) e alla personalità dell'agente[358].

In relazione all'intelligenza artificiale, per quanto sia possibile affermare che i robot hanno la capacità di rappresentarsi la realtà prevedendo e volendo un risultato come conseguenza di una loro azione, ciò avviene solo in ragione del fatto che le stesse sono programmate per raggiungere un obiettivo specifico e grazie alle tecniche di machine learning riescono a "ragionare" come se fossero esseri umani. Ciò non basta: la volontà di realizzazione dell'evento è solo apparente e dunque il dolo non sussiste a causa dell'incapacità delle macchine di autodeterminarsi. Non è un vero dolo, è un'apparenza di dolo: socialmente sembrerebbe un fatto doloso perché espressione di una direzionalità capace di adattarsi alla realtà mentre punta al suo obiettivo, ma ciò è solo l'involucro esteriore. La tipicità delle azioni della macchina, per quanto possa essere spinta da una logica dolosa in astratto, non riguarda la colpevolezza che sussiste solo qualora ci sia libertà di autodeterminazione in capo ad una macchina che possiede la capacità di scegliere liberamente di compiere un fatto antigiuridico. Questo manca nelle macchine intelligenti perché le stesse non sono ancora in grado di porsi dei singoli obiettivi egoistici volendoli raggiungere anche a costo di ledere i beni giuridici

[358] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VIII "La Colpevolezza", Sottoparagrafo A (pag. 352 – 423) "Dolo, colpa e dolo misto a colpa".

altrui; i robot non sono dotati di capacità di scelta pura: mancando di colpevolezza non possono neanche essere considerati rimproverabili e dunque soggette a pena[359].

La colpa invece, si differenzia dal dolo perché non possiede un contenuto psicologico. Secondo l'articolo 43 comma 3 c.p., il delitto è colposo o contro l'intenzione quando l'evento anche se preveduto non è voluto dall'agente e si verifica a causa di negligenza o imprudenza o imperizia ovvero per inosservanza di leggi, regolamenti, ordini o discipline. I requisiti sono quindi due: la non volontà dell'evento (e dunque un'assenza di dolo) e una violazione delle regole cautelari (non scritte per la colpa generica, scritte per la colpa specifica) che se osservate avrebbero impedito l'evento.

La colpa si divide in cosciente e incosciente. La prima è anche detta colpa con previsione (effettiva e non potenziale) dell'evento e risulta di difficile distinzione con il dolo essendo il grado di colpa più elevato; il soggetto si è rappresentato la possibilità di verificazione dell'evento ma confida sinceramente nella non verificazione dell'esito infausto. La seconda è di grado inferiore e consiste in una colpa senza previsione dell'evento; è legata ad una leggerezza dell'agente (una sottovalutazione del pericolo o una sopravalutazione della propria capacità di evitarlo) che non gli permette di rendersi conto che la sua condotta potrebbe provocare eventi dannosi.

Per l'accertamento della colpa bisogna analizzare, nel caso della generica, i parametri della prevedibilità e dell'evitabilità dell'evento; nel caso della colpa specifica invece, bisogna valutare la violazione della norma scritta e la realizzazione di un evento che la norma mirava ad evitare. Per valutare il comportamento tenuto in un contesto pericoloso, ci si rifà alle azioni che avrebbe tenuto (nello stesso contesto e con le stesse conoscenze e condizioni) un uomo ideale (l'agente-modello) considerato come "homo eiusdem professionis et condictionis"[360]. Secondo alcuni, in relazione all'intelligenza artificiale, è possibile accertare la sussistenza della negligenza, dell'imperizia e dell'imprudenza andando a paragonare l'azione della macchina intelligente a quella di un agente–modello (artificiale e intelligente) che sia munita delle stesse esperienze e capacità di machine learning della intelligenza artificiale concreta[361].

Il dolo misto a colpa, invece, è connesso al concetto di responsabilità oggettiva. Riguarda la situazione in cui un elemento o l'intero fatto viene addossato all'agente

[359] A. CAPPELLINI, "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale", op. cit., Paragrafo 5 (pag. 13 - 15) "Prima critica: persistente assenza di colpevolezza"

[360] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VIII "La Colpevolezza", Sotto-paragrafo A (pag. 352 – 423) "Dolo, colpa e dolo misto a colpa"

[361] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'"

senza che sia necessario accertare il dolo o la colpa. Si assiste dunque ad una responsabilità penale a prescindere dall'elemento psicologico, sulla base del mero rapporto causale. Questo risulta in contrasto con il moderno principio di colpevolezza. L'articolo 42 comma 3 c.p. tratta la responsabilità oggettiva affermando che la legge determina i casi nei quali l'evento è posto altrimenti (a prescindere dall'elemento soggettivo) a carico dell'agente, come conseguenza della sua azione od omissione. Per effettuare un'interpretazione conforme a Costituzione, è necessario seguire il principio espresso dalla Sentenza 364/1988[362] (esposto all'inizio di questo sotto-paragrafo)[363].

Per quanto concerne l'assenza di scusanti, queste consistono in un catalogo tassativo di circostanze anomale che hanno influito in modo irreversibile sulla volontà e capacità psicofisica del reo. Sono connesse al principio dell'inesigibilità: non si poteva pretendere un comportamento diverso. Esse vanno, quindi, ad eliminare l'elemento psicologico[364].

Il requisito della conoscenza o conoscibilità della legge penale è posto alla base del principio di colpevolezza in quanto si necessita di accertare se l'agente sapesse o potesse sapere che quel fatto (da lui commesso) fosse previsto dalla legge come reato. Su di esso possono influire gli errori, ossia, stati dell'intelletto per cui una circostanza del mondo esterno non è più conosciuta come dovrebbe essere realmente. L'errore consiste in una creazione di un convincimento sbagliato sullo stato di fatto delle cose. Esso si differenzia dall'ignoranza (che consiste in una mancanza di conoscenza) e dal dubbio (nel quale sussiste un conflitto di opinioni o giudizi e che dunque non sfocia in un vero e proprio convincimento).

Gli errori legati alla formazione della volontà sono di due tipologie: l'errore di diritto e l'errore di fatto.

Il primo riguarda la norma penale ed è connesso al principio di obbligatorietà della legge penale. L'obbligo di rispettare le norme comprende in sé l'obbligo di conoscere la legge penale: "ignorantia legit non excusat". Al momento della commissione del fatto il soggetto deve sapere che si tratta di un reato. Il rimprovero che viene mosso verso il soggetto deve essere connesso al presupposto della conoscibilità della norma al fine di infliggere una sanzione che possa effettivamente rieducare il reo. Se sussiste

[362] Sentenza della Corte Costituzionale n. 364, del 23 marzo 1988: https://www.cortecostituzionale.it/actionSchedaPronuncia.do?anno=1988&numero=364

[363] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VIII "La Colpevolezza", Sotto-paragrafo A (pag. 352 – 423) "Dolo, colpa e dolo misto a colpa"

[364] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VIII "La Colpevolezza", Sotto-paragrafo B (pag. 423 – 429) "Assenza di scusanti"

un'oggettiva impossibilità di conoscere la legge penale (valevole per tutti i consociati a causa di un'assoluta oscurità del testo legislativo) allora, l'ignoranza è accettata e scusabile.

L'errore di fatto, invece, consiste nell'assenza del momento rappresentativo (e quindi del dolo) andando ad escludere la punibilità dell'agente. È un errore che cade su uno degli elementi richiesti per l'esistenza del reato a causa di un'errata percezione sensoriale dell'evento. Se l'errore è dovuto a colpa, il soggetto sarà punito per colpa solo se il delitto è previsto dalla legge anche nella forma colposa[365].

Per quanto riguarda l'intelligenza artificiale bisogna tenere presente la sussistenza dei bias intesi come errori di valutazione che possono incidere negativamente sugli output finali e che derivano da sviste nella generazione del dato e nelle annotazioni realizzate dagli esseri umani, oppure, da errori nella raccolta autonoma dei dati effettuata dal sistema stesso[366]. Essi si dividono in data bias (relativi a problematiche concernenti i dati di input) e automation bias (riguardanti le modalità di acquisizione automatica delle informazioni e la loro analisi, a causa, anche, di discriminazioni congenite della società) che generano errori di valutazione comportando azioni omissive o commissive[367]. Infatti, i sistemi intelligenti, non possedendo dei binari fissi e muovendosi grazie a calcoli probabilistici possono generare bias nell'algoritmo che oltretutto non risultano neanche immediatamente riscontrabili a causa dell'oscurità dello stesso[368]. Tali bias potrebbero astrattamente essere paragonati agli errori degli esseri umani in quanto consistono in un'errata percezione della realtà che genera un'azione lesiva. Quest'azione è senz'altro realizzata dal sistema intelligente ma sussiste un vizio nella formazione della decisione di realizzare quell'illecito.

Inoltre, ci si potrebbe porre il quesito sulla possibilità per la macchina di conoscere la legge penale e se dunque una sua azione in contraddizione con la stessa sia dovuta ad una non conoscenza incolpevole o se invece sia una violazione di una legge penale in realtà conosciuta dal sistema. Infatti, i bias possono anche essere connessi ad un'errata interpretazione della norma come nel caso di infrazioni compiute delle self-drive cars

---

[365] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VIII "La Colpevolezza", Sotto-paragrafo C (pag. 429 – 433) "Conoscenza o conoscibilità della legge penale violata"

[366] M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, op. cit., Paragrafo 2 (pag. 2 – 6) "Il dato"

[367] M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, op. cit., Paragrafo 6 (pag. 46 – 49) "Conclusioni. La "Competenza Artificiale": monitoraggio, formazione e informazione dei TSRM quali strumenti principali per fronteggiare il fenomeno"

[368] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".

che dovrebbero necessariamente muoversi nel rispetto del codice della strada. Ciò si riconnette al tema della conoscenza e della comprensione effettiva che il sistema intelligente può avere del mondo esterno e delle sue regole: l'intelligenza artificiale non conosce ma imita un essere umano che conosce. Il suo grado di comprensione e conoscibilità dipende dal programma che viene inserito al suo interno. Con l'evoluzione dei robot dotati di alte capacità cognitive, è stato ipotizzato lo sviluppo di macchine munite di un codice etico da seguire durante il processo decisionale. I dubbi sono stati posti sulla tipologia dei criteri da considerare al fine di condizionare le scelte del robot. Si potrebbe, quindi, ipotizzare la possibilità di inserire all'interno dello stesso le norme di diritto penale al fine di imporre la conoscenza della legge anche alla macchina, che non potrebbe più "essere scusata" per una non comprensione delle norme giuridiche. Ad oggi però, non è (ancora) stata codificata una modalità efficace che permetta di inserire i principi morali nell'addestramento dell'algoritmo al loro rispetto[369].

Infine, l'ultimo requisito della colpevolezza consiste nella capacità di intendere e di volere (art. 85 c.p.) secondo cui nessuno può essere punito per un reato, se, al momento della commissione non era imputabile.

A differenza delle precedenti concezioni (la più importante quella di Antolisei) in base alle quali l'imputabilità non era un presupposto della colpevolezza ma un mero stato del soggetto che riguardava la figura del reo e non il concetto del reato (potendo dunque considerare colpevole anche un soggetto non capace di intendere e volere), la colpevolezza risulta, oggi, in stretto rapporto con l'imputabilità. Dato che la colpevolezza implica il rimprovero personale del fatto commesso il quale, a sua volta necessita della maturità psicologica del reo, se non sussiste la piena capacità di intendere e di volere non si può parlare di imputabilità[370].

Per imputabile si intende chiunque abbia la capacità di intendere e di volere; in mancanza di tale elemento non si può essere puniti. La capacità di intendere consiste nella consapevolezza del valore sociale dei propri atti e delle loro conseguenze. La capacità di volere consiste, invece, nell'abilità di controllare un impulso motorio potendolo inibire o assecondare a seguito di un processo razionale. Esse sono legate

[369] M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, op. cit., Paragrafo 3.4 (pag. 31 - 36) "L'Intelligenza Artificiale più evoluta considerata come una persona"

[370] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VIII "La Colpevolezza", Paragrafo 1 (pag. 349 – 352) "La colpevolezza: nozione, fondamento e rilevanza costituzionale"

alla piena maturità psichica e alla piena sanità mentale del soggetto[371]. L'individuo deve essere consapevole del significato delle proprie azioni e delle loro conseguenze, possedendo inoltre la capacità di autodeterminarsi liberamente. La libertà di autodeterminazione, connessa alla capacità di volere, riguarda l'abilità dell'essere umano di determinarsi secondo la propria ragione. Essa risulta legata al libero arbitrio, il quale infatti consiste nel potere di decidere degli scopi del proprio agire, in pieno possesso delle proprie capacità mentali e senza essere condizionati dal volere altrui. Dunque, non si può essere imputati per la commissione di un reato se non si è in possesso di una la libertà emotiva e mentale di scelta.

Nel corso della storia, l'uomo si è sempre posto domande in relazione alla sua capacità di intendere e di volere e soprattutto, in merito alla sua effettiva libertà di scelta. Nel mondo antico, era chiara l'idea di forze sovrannaturali (gli Dei o il Fato) capaci di incidere sul comportamento degli esseri umani in modo positivo o negativo senza che il destinatario avesse possibilità di ribellarsi. Nascono così i rituali superstiziosi e religiosi volti a limitare gli effetti negativi di tali forze sugli individui. Anche nel Medioevo è possibile ritrovare l'affermazione dei limiti al libero arbitrio tramite S. Agostino che attribuiva il destino individuale al volere divino. Lo stesso è possibile ritrovarlo nelle religioni Buddiste e Musulmane, le quali rifiutano il pieno libero arbitrio. Dio è la causa di tutto. Anche Schopenhauer esclude la libertà e la volontà: tutto il mondo compreso il comportamento umano è determinato e basato sul principio di causalità. Il libero arbitrio null'è se non una semplice illusione[372].

Per libero arbitrio si intende il fatto che l'essere umano può essere influenzato ma rimane libero nelle sue scelte di comportamento. È un concetto pratico e psicologico. L'imputabilità, invece, è un concetto puramente giuridico. Essa riguarda la possibilità di essere considerati responsabili per comportamenti tenuti in contrasto alla legge. L'imputabilità, di per sé non necessiterebbe del libero arbitrio, ma viene ad esso connessa al fine di evitare rischi di responsabilità oggettiva e di eccessiva repressione punitiva senza alcun fine preventivo o rieducativo. Libero arbitrio e imputabilità sono dunque due presupposti fondamentali e interconnessi per applicare in concreto le sanzioni penali con fine rieducativo.

---

[371] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VIII "La Colpevolezza", Sotto-paragrafo D (pag. 433 – 453) "Capacità di intendere e di volere"

[372] G. GALUPPI, "Libero arbitrio, imputabilità, pericolosità sociale e trattamento penitenziario", Dir. famiglia, fasc.1, 2007, pag. 328; Paragrafo 2 "Cenni sulla negazione del libero arbitrio nei secoli passati"

Il concetto di libero arbitrio nasce con la scuola classica nel periodo illuminista secondo cui la necessità di punire il reo si annulla nel momento in cui si può dimostrare la sua infermità mentale. Nel XIX secolo, con la scuola positiva, si ipotizzò l'inutilità della sanzione penale a causa dell'inesistenza di un vero libero arbitrio e dunque i criminali dovevano semplicemente essere posti nella condizione di non nuocere: al centro viene messa la tutela della vittima e non la rieducabilità del reo. Attualmente, in Italia, è possibile osservare una concezione posta al centro tra la scuola classica e positiva: la primaria azione della pena è l'educazione basandosi sul principio secondo cui ogni criminale è rieducabile. La problematica riguarda il fatto che in concreto, nelle carceri, risulta complesso assistere ad un effettivo ed efficiente modello rieducativo rischiando quindi di rendere la pena meramente punitiva e contenitiva[373].

Secondo M. Minsky, che come abbiamo accennato nel primo capitolo è considerabile come il padre dell'intelligenza artificiale, la volontà è esclusivamente illusoria[374] e dunque anche le macchine intelligenti possono essere paragonate all'uomo in tema di imputabilità.

È possibile riconoscere come l'avvento dell'intelligenza artificiale abbia messo in crisi il normale modello di imputazione della responsabilità connesso al libero arbitrio e alla capacità di intendere e di volere. Le macchine stanno sviluppando la capacità di muoversi in modo autonomo anche in assenza di un controllo umano, dando luogo ad una parvenza di autonomia. I robot possono agire nell'ambiente circostante senza necessità di una guida in modo indipendente da istruzioni iscritte precedentemente. Hanno quindi la capacità di reagire in maniera efficiente e non necessariamente preordinata agli input che ricevono grazie alla

capacità di apprendimento che rende gran parte delle loro decisioni imprevedibili ex ante[375].

I robot come sappiamo sono interattivi, reattivi con l'ambiente, dotati di azione autonoma, imprevedibile e non determinata, flessibile e influenzabile; possiedono quindi autonomia, interattività e adattabilità essendo in grado di migliorare le loro abilità autonomamente grazie ai meccanismi di machine learning[376]. Ma, al contempo,

[373] G. GALUPPI, op. cit., Paragrafo 3 "La prospettiva delle varie "scuole" giuridiche sull'imputabilità e sul libero arbitrio"

[374] G. GALUPPI, op. cit., Paragrafo 4 "Libero arbitrio e neuropsicologia moderna"

[375] E. PALMERINI, op. cit., Paragrafo 6 "I problemi della sicurezza e della responsabilità nel mercato emergente della robotica"

[376] M. B. MAGRO, "Biorobotica, robotica e diritto penale", op. cit., Paragrafo 1 (pag. 1 – 7) "Biorobotica, interfacce cervello- macchina e potenziamento umano: filosofia precauzionale e euristica di avversione al rischio"

la macchina intelligente risulta (ancora) priva di una coscienza ed intenzionalità delle proprie azioni e dunque, risulta priva di una capacità di determinarsi diversamente. Il sistema intelligente non è libero ma determinato: la mancanza di liberà di intendere e volere si riflette su una mancanza di colpevolezza. Il fatto che i robot sappiano muoversi da soli nello spazio, sulla base di decisioni assunte senza l'intervento dell'uomo,

è in realtà una mera parvenza di autonomia essendo prive di stati mentali psichici[377].

2.3 Fine rieducativo della pena

Accanto, al quesito "machina delinquere potest?", bisognerà porsi anche il connesso quesito: "(quomodo) machina puniri potest?"[378]

Le problematiche (simili a quelle sollevate in relazione al diritto penale degli enti) concernono principalmente l'assenza di sentimenti in capo all'intelligenza artificiale, la possibile assenza di un corpo fisico e di un conto bancario a cui applicare rispettivamente le sanzioni detentive o pecuniarie.

L'articolo 27 della Costituzione pone al terzo comma il principio della rieducazione del condannato. La pena deve essere proporzionale alla gravità del fatto commesso e non deve consistere in trattamenti disumani o degradanti. Sono quindi vietate pene contrarie al senso di umanità, pene corporali e la pena di morte. Si ha dunque un nesso indissolubile tra la pena e il senso di umanità. Il detenuto può essere privato della liberà ma non della dignità (Sentenza Torreggiani 2013)[379]. La finalità del carcere non è la neutralizzazione ma la rieducazione.

Nel 1930 la visione della pena era principalmente carcerocentrica. Lo scopo era eliminare e isolare il reo considerato meritevole della privazione della libertà personale. Con la Costituzione si sancisce il fine rieducativo della pena abbandonando il primato della pena detentiva carceraria intesa come segregazione.

Lo Stato deve garantire un trattamento risocializzante ad ogni detenuto: serve anche la volontà dello stesso ad essere rieducato. Infatti la pena "tende" alla rieducazione

[377] A. CAPPELLINI, "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale", op. cit., Paragrafo 2 (pag. 4 – 5) "Le origini del machina delinquere non potest"

[378] F. BASILE, op. cit., Paragrafo 6.3.4 (pag. 31 – 32) "Quali pene per i sistemi di IA?"

[379] Sentenza della Corte Europea dei Diritti dell'Uomo, II Sezione, Causa Torreggiani e Altri c. Italia, Strasburgo 8/1/2013; Ricorsi nn. 43517/09, 46882/09, 55400/09, 57875/09, 61535/09, 35315/10 e 37818/10: https://www.camera.it/application/xmanager/projects/leg17/attachments/sentenza/testo_ingleses/000/000/541/Torreggiani.pdf

perché non è detto che l'obiettivo venga raggiunto (e spesso, purtroppo, il sistema carcerario non riesce a garantirlo)[380].

Come possiamo notare, tale assetto è incentrato su un soggetto umano che ha commesso un reato e che dunque può essere rieducato. Per quanto concerne il caso in cui sia l'intelligenza artificiale a realizzare l'illecito e volessimo considerare possibile un'applicazione della pena ad essa, risulta necessario dunque effettuare un adeguamento della pena normalmente inflitta all'essere umano per concordarla con tale innovazione.

Secondo la teoria di Hallevy, il significato della punizione rimane lo stesso per gli esseri umani e per le entità di intelligenza artificiale.

Qualora volessimo applicare la pena di morte (incostituzionale in Italia e in altri ordinamenti) al robot (consistente nello spegnimento definitivo del sistema o nella sua distruzione) lo scopo sarebbe lo stesso nel caso di applicazione all'essere umano: l'assoluta impossibilità di reiterazione dell'illecito da parte del reo oramai privato della vita. Per l'intelligenza artificiale, la vita è la sua esistenza come entità con o senza un corpo fisico. Una volta cancellato il software di controllo dell'intelligenza artificiale, la stessa non potrà commettere ulteriori reati. Invece, l'incarcerazione, per gli esseri umani ha come significato la privazione della libertà personale. La libertà di un sistema intelligente include la sua capacità di agire come entità nello spazio; in ambito pratico, per la macchina intelligente, consisterebbe nella sua inutilizzabilità per un determinato periodo. Per quanto riguarda il servizio alla comunità, questo può essere utilizzato come sostituto per pene detentive brevi ed ha lo scopo rieducativo e risocializzante per l'essere umano. Per l'intelligenza artificiale, questo può essere garantito imponendo il lavoro obbligatorio nella collettività. Infine, in relazione alla pena pecuniaria, questa ha lo scopo di colpire il patrimonio di un soggetto; tale sanzione risulta complessa per i sistemi di intelligenza artificiale, non avendo gli stessi una disponibilità economica che gli permette di adempiere alla stessa (potrebbe però essere trasformata in servizio alla comunità imponendo un lavoro socialmente utile all'intelligenza artificiale)[381].

[380] P. BALDUCCI E A. MACRILLÒ, "Esecuzione penale e ordinamento penitenziario", Giuffrè Francis Lefebvre S.p.a. Milano – 2020; Parte III "Il diritto penitenziario", L. VIOLI, Capitolo 1 (pag. 675 – 765) "Trattamento penitenziario"

[381] G. HALLEVY, "The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control", Akron Intellectual Property Journal: Vol. 4: Iss. 2, Article 1; Paragrafo 4 (pag. 194 – 199) "General punishment adjustment considerations": https://ideaexchange.uakron.edu/akronintellectualproperty/vol4/iss2/1

L'attribuzione di soggettività giuridica ai sistemi di intelligenza artificiale permetterebbe di applicare tali pene alla macchina e di risolvere la lacuna nel caso di reati commessi dai robot[382].

Analizzando le critiche mosse alla teoria di Hallevy, è possibile affermare il fatto che, nei confronti di un agente artificiale intelligente, la pena non potrebbe rispettare le sue tipiche funzioni[383]: la retributiva, la disincentivante (general-preventiva), e la rieducativa-dissuasiva (special-preventiva). La prima consiste nell'infliggere la sanzione esclusivamente per punire e "farla pagare" al soggetto per il reato che ha commesso; la seconda opera come minaccia che frena il criminale dal commettere l'illecito; e la terza ha come obiettivo la risocializzazione del reo in modo tale da non fargli più commettere crimini in futuro.

La funzione retributiva e la rieducativa potrebbero, in astratto, essere applicate anche nei confronti dei sistemi intelligenti tramite la soppressione della macchina o tramite la sottoposizione della stessa ad un training rieducativo[384]. In realtà, la funzione retributiva non potrebbe essere applicata dato che, come affermato nei sotto-paragrafi precedenti, le macchine intelligenti non sono (ancora) suscettibili di un rimprovero colpevole[385]. E la prevenzione speciale connessa alla rieducazione e risocializzazione (e non alla neutralizzazione) risulta poco effettiva nei loro confronti in quanto esse (non possedendo il libero arbitrio)[386] non sono capaci di apprendere dalla sanzione irrogata come conseguenza di una propria azione sbagliata; salvo nel caso in cui si vada ad incidere sul meccanismo di machine learning ottimizzando progressivamente il suo comportamento[387] (ma questo significherebbe agire direttamente su di essa, senza la necessità di uno strumento indiretto di pressione come la sanzione penale)[388].

Sicuramente complessa (se non impossibile) risulta essere la finalità di prevenzione generale in relazione al robot dato che una macchina non è in grado di provare emozioni come la paura e dunque non potrebbe essere disincentivata dalla minaccia

[382] S. QUINTARELLI, op. cit., Capitolo 6.2 (pag. 110 – 113) "Personalità giuridica. La convivenza tra uomini e software"

[383] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?"

[384] F. BASILE, op. cit., Paragrafo 6.3.4 (pag. 31 – 32) "Quali pene per i sistemi di IA?"

[385] A. CAPPELLINI, "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale", op. cit., Paragrafo 6 (pag. 15 - 16) "Seconda critica: perdita di senso delle funzioni della pena"

[386] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'"

[387] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?"

[388] A. CAPPELLINI, "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale", op. cit., Paragrafo 6 (pag. 15 - 16) "Seconda critica: perdita di senso delle funzioni della pena"

della pena alla commissione del reato[389]. Inoltre, la pena peccherebbe di quella attitudine comunicativa della sanzione inflitta ad uno, nei confronti degli altri: la pena rimarrebbe così un fatto isolato e relativo solo al robot che la subisce, risultando estranea agli altri[390].

Alla luce di quanto esposto in questo secondo paragrafo è possibile concludere affermando un'impossibilità di riconoscere l'intelligenza artificiale come direttamente responsabile dei fatti illeciti commessi, a causa della mancanza dell'elemento soggettivo in capo alla stessa. Con ciò, però, non si vuole escludere che in futuro, a seguito dell'evoluzione tecnologica, una soggettività penale possa essere riconosciuta anche in capo alle macchine intelligenti. La scienza corre e il diritto penale la segue: nel momento in cui sarà possibile (e se sarà possibile) considerare esistenti gli stati mentali tipici degli esseri viventi anche in capo ad agenti artificiali, allora il legislatore dovrà intervenire ipotizzando nuove categorie giuridiche e nuove situazioni per evitare lacune di tutela.

3. Responsabilità del programmatore, dell'utilizzatore o dell'intelligenza artificiale?

Affermando l'assenza di una responsabilità penale diretta in capo all'intelligenza artificiale si pongono problematiche in relazione alle situazioni che, se non analizzate attentamente, resterebbero prive di tutela. La macchina infatti può porre in essere un'azione lesiva considerabile come reato, dunque è necessario individuare il soggetto persona fisica indirettamente responsabile e penalmente sanzionabile[391].

Come abbiamo precedentemente sottolineato dilemma si pone, nel momento in cui il robot non sia un mero strumento nelle mani del reo (in tali situazioni è facile considerare il soggetto come il reale autore del reato), ma quando il sistema abbia la capacità di muoversi autonomamente. Bisogna quindi andare ad individuare la persona fisica, nella catena di produzione e utilizzazione del sistema intelligente che, in concreto, sia il destinatario dell'azione penale[392].

[389] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?"

[390] A. CAPPELLINI, "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale", op. cit., Paragrafo 6 (pag. 15 - 16) "Seconda critica: perdita di senso delle funzioni della pena"

[391] A. CAPPELLINI, "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale", op. cit., Paragrafo 1 (pag. 1 - 4) "Macchine e modelli imputativi"

[392] A. LONGO E G. SCORZA, op. cit., Capitolo 7 (pag. 198 – 240) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela"

Escludere che la condotta possa essere realizzata coscientemente dalla macchina, impedisce di reputarla come il reo e risulta dunque necessario individuare un altro soggetto come tale, al fine di evitare vuoti di tutela.

Per considerare l'essere umano come responsabile, si necessita sia della prova del nesso di causalità tra la sua azione di produzione del sistema intelligente e l'evento che lo stesso ha posto in essere che la sussistenza dell'elemento soggettivo in capo all'agente. Su tale tema sono stati posti dubbi in dottrina soprattutto nei casi in cui l'intelligenza artificiale sia fortemente evoluta e dotata di sistemi di machine learning che rendono la sua condotta simile a quella umana: in questi casi, neanche chi ha progettato il robot è in grado di ricostruire una correlazione tra l'attività di programmazione della macchina svolta dall'essere umano e il danno da essa provocato[393]. L'unico danno-evento penalmente rilevante è quello di diretta e immediata conseguenza dell'azione. È importante la prevedibilità dell'evento dannoso, requisito che, nel caso di algoritmi complessi risulta quasi impossibile dato che il creatore dell'algoritmo non sempre può prevedere ex ante le azioni che la macchina andrà a compiere a causa della vastità degli input che il robot può acquisire e dai quali può autonomamente imparare ad agire (per questo si parla di black box algorithm). Tra l'azione umana di programmazione del sistema e l'azione lesiva commessa dal robot, risulterà complesso individuare sia la colpevolezza (in quanto sarà difficile individuare una volontarietà o prevedibilità dell'azione da parte dell'essere umano) che la suitas (intesa come riconducibilità dell'azione al soggetto in base alla sua sfera di signoria e controllo)[394].

Sappiamo che il nostro ordinamento non permette di far rispondere un soggetto per la sola responsabilità oggettiva in assenza di un elemento psicologico in quanto ciò si porrebbe in contrasto con l'articolo 27 comma 1 della Costituzione. Infatti, come vedremo nel corso della trattazione, una volta individuati i soggetti rimproverabili per l'azione lesiva dei robot, la loro responsabilità dovrebbe essere proporzionata al livello di istruzioni impartite al robot e al suo grado di autonomia. In base a tale ragionamento e a quanto espresso nel secondo paragrafo di questo capitolo, la responsabilità va imputata all'umano e non al robot[395].

---

[393] M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, op. cit., Paragrafo 3 (pag. 14 - 36) "Qualificazione giuridica dell'Intelligenza Artificiale"

[394] M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, op. cit., Paragrafo 4 (pag.36 - 40) "Le difficoltà nella ricostruzione del nesso causale: "in dubio pro machina"

[395] A. LONGO E G. SCORZA, op. cit., Capitolo 7 (pag. 198 – 240) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela"

Innanzitutto, allo scopo di ricondurre il paradigma di responsabilità al modello imposto dalla Costituzione, cioè quello della responsabilità personale, potrebbero individuarsi i profili di responsabilità dei soggetti che a vario titolo ed in varie fasi della "vita" della macchina intelligente intervengono su di essa. Si pensi al programmatore (che scrive gli algoritmi e fornisce alla macchina il mezzo fondamentale con cui essa interagirà con il mondo fisico), al produttore (che tramite i suoi fattori produttivi consente la messa in commercio o comunque l'utilizzo da parte di altri della macchina) e da ultimo, si pensi all'utilizzatore finale della stessa (che se ne serve per scopi per cui essa è programmata o meno).

Interrogarsi sull'adattabilità degli attuali modelli di responsabilità o sulla necessità di istituirne di nuovi (come colpa di programmazione o di automazione che coinvolgono l'impresa produttrice della macchina sul modello della product liability) risulta inevitabile in casi in cui i reati sono commessi da sistemi di intelligenza artificiale in cui la condotta è opera dell'azione condivisa tra umano e robot[396]. È quindi utile che il legislatore si interroghi su come eliminare i vuoti di tutela oggi (e soprattutto in futuro) riscontrabili a seguito dell'impiego di sistemi intelligenti nella società.

È indubbio che la possibilità di considerare la responsabilità diretta della macchina affascina e incuriosisce (in primis con le teorie di Hallevy); ma ad oggi, è rimasta relegata a mero caso di scuola senza mai essere applicata in concreto. Di conseguenza, l'unica tipologia di responsabilità ravvisabile nella realtà è la responsabilità vicaria dell'essere umano.

Considerando i robot solo come mezzi e non agenti non umani del reato, si pone il problema della responsabilità esclusivamente umana del programmatore o dell'utilizzatore. Nei paragrafi e nel capitolo precedente è stata sottolineata più volte la capacità del robot intelligente di agire in modo imprevedibile anche per il programmatore. L'imprevedibilità dovuta al black box algorithm pone non poche problematiche in relazione all'imputazione della responsabilità penale all'umano dato che, la macchina, imparando ad agire autonomamente tramite le tecniche di machine learning non permette di ipotizzare antecedentemente le azioni che compirà in futuro.

[396] V. MANES, op. cit., Paragrafo 1.1 (pag. 2 – 5) "L'imputazione della responsabilità penale per azioni (autonome) delle macchine"

La responsabilità dell'essere umano deve essere considerata come diretta o indiretta? A titolo di dolo (perché la macchina agisce su volontà dell'uomo) o di colpa (dovuta ad un difetto di programmazione o utilizzo del robot)?[397]

Hallevy prospetta tre modelli (applicabili separatamente o in combinazione tra essi) di responsabilità penale in relazione ai reati commessi dell'intelligenza artificiale: la responsabilità per reato tramite un altro, la responsabilità per conseguenze naturali e la responsabilità diretta[398].

Il primo modello, anche detto "the perpetration-via-another liability model"[399], considera l'intelligenza artificiale come agente innocente. Le sue caratteristiche di autoapprendimento, di autonomia e di movimento al di fuori del controllo umano le permettono di porre in essere un reato ma essa rimane pur sempre una macchina e di conseguenza, nonostante possa essere ritenuta come materialmente capace di realizzare un fatto di reato, la stessa viene considerata come un soggetto limitatamente capace di intendere e di volere e dunque non imputabile.

Quindi si considera responsabile non la macchina, ma il soggetto che la utilizza o che la progetta per commettere l'illecito e la responsabilità della persona fisica è determinata sulla base della condotta della macchina. I soggetti da considerare sono il programmatore del software (se lo progetta per commettere reati) e l'utente finale (che utilizza il sistema e il suo software a suo vantaggio). Se il robot commette un crimine risponde il programmatore, se lo progetta per farlo, e l'utilizzatore, se lo usa a tale scopo (come nel caso di un animale che aggredisce su ordine del suo padrone). In entrambi i casi il reato è commesso dall'entità artificiale intelligente senza che l'essere umano abbia compiuto alcuna azione ma ne risponde comunque sulla base di un utilizzo strumentale della macchina innocente.

L'elemento oggettivo è realizzato dal sistema su ordine dell'utente/programmatore che possiede il pieno elemento soggettivo a differenza del robot. L'intelligenza artificiale è solo il mezzo attraverso cui si realizza il reato, al pari di una pistola o un animale impiegati nella commissione dell'illecito. Ma, a differenza della pistola o dell'animale, il robot può eseguire un ordine anche molto complesso.

---

[397] M. B. MAGRO, "Biorobotica, robotica e diritto penale", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali".

[398] G. HALLEVY, op. cit., Paragrafo 1 (pag. 171 – 175) "Introduction: The Legal Problem".

[399] Hallevy, Prof. Gabriel, The Basic Models of Criminal Liability of AI Systems and Outer Circles (June 11, 2019). https://ssrn.com/abstract=3402527 or http://dx.doi.org/10.2139/ssrn.3402527

Tale tipologia di responsabilità, secondo Hallevy, può essere considerata solo nel caso di utilizzo di un sistema intelligente non altamente sofisticato che dunque non possiede un black box algorithm di machine learning. La perpetration-via-another liability non risulta idonea nel caso in cui l'intelligenza artificiale decida di commettere un reato (senza che sia stata programmata o utilizzata per realizzarlo) in base ad azioni di autoapprendimento. In tali ipotesi, la macchina si comporta come un agente semi-innocente[400].

Una simile posizione si presta a possibili opposizioni e correzioni che riportano la perpetration-via-another liability ad una "semplice" responsabilità a carico del soggetto agente (persona fisica) a titolo di dolo. Infatti, l'autore del reato decide autonomamente di creare o utilizzare il sistema intelligente per commettere quello specifico illecito. La macchina, nell'ipotesi prospettata da Hallevy, difatti, non si comporta come un agente semi-innocente ma come un mero strumento nelle mani del reo. Precisamente, al soggetto agente non viene imputata una condotta altrui commessa dal robot ma un fatto da lui commesso utilizzando uno strumento che nel caso di specie è un sistema intelligente. Quindi, non si è di fronte ad una c.d. perpetrator-by-another liability ma ad una responsabilità diretta e dolosa dell'utilizzatore o programmatore che decide di utilizzare il robot come strumento del reato[401].

Il secondo modello, anche detto "the natural-probable-consequence liability model", riguarda i reati commessi dall'intelligenza artificiale che potevano essere previsti. I programmatori o utenti del sistema non hanno interesse a commettere l'illecito e non ne erano coscienti fino a che la macchina non lo ha realizzato. Loro non hanno progettato o partecipato ad alcuna parte del reato ma le azioni della macchina, non volute dal programmatore o dall'utilizzatore, sono state comunque realizzate secondo il programma impartitogli (ad esempio nel caso in cui il sistema consideri un soggetto come minaccia per la sua missione e decida di eliminarlo).

Il programmatore o l'utente non erano a conoscenza, non avevano pianificato e non avevano l'intenzione di realizzare il reato commesso dal sistema intelligente. Dunque, questo secondo modello si basa sulla capacità dei programmatori o degli utenti di prevedere la potenziale commissione dei reati. Quindi saranno considerati responsabili esclusivamente se sussisteva la probabilità prevedibile di realizzazione dell'illecito: è necessaria almeno la colpa valutata in base al comportamento che un agente-modello

[400] G. HALLEVY, op. cit., Paragrafo 3 (pag. 177 – 194) "Three models of the criminal liability of artificial intelligence entities"
[401] I. SALVADORI, op. cit., Paragrafo 5.1, "La responsabilità a titolo doloso"

avrebbe tenuto. Si rimprovera al soggetto il fatto-reato che lo stesso avrebbe dovuto prevedere al fine di evitarlo.

Nell'applicazione di tale modello di responsabilità è possibile distinguere due casi: quando il soggetto è stato negligente durante la programmazione o l'utilizzo dell'intelligenza artificiale senza però avere intenzione di commettere il reato, e quanto il soggetto ha progettato o utilizzato l'intelligenza artificiale al fine (consapevole e voluto) di commettere un reato ma il sistema intelligente ha deviato il piano commettendo un reato differente da quello previsto. Nella prima situazione si tratta di una pura negligenza; nel secondo di aberratio delicti. I soggetti saranno dunque ritenuti responsabili almeno a titolo di colpa[402].

La creazione di tale teoria da parte di Hallevy, non risulta però innovativa in quanto è possibile considerare queste ipotesi come semplice responsabilità per colpa della persona fisica valutata in base alla prevedibilità, da parte del creatore o utente, delle azioni commesse dall'agente intelligente. Gli illeciti realizzati a causa di errori di programmazione possono essere imputati al programmatore solo a titolo di colpa. E chi utilizza il sistema risponde per colpa delle azioni lesive che avrebbe potuto impedire. Tale tipologia di responsabilità risulta più complessa nel caso in cui, a seguito della notevole autonomia del robot, l'essere umano non abbia le capacità di correggere o prevedere le azioni che lo stesso andrà a realizzare. Nei casi in cui il software non abbia le funzioni per permettere all'individuo di intervenire ed evitare i reati, allora, forse la responsabilità andrà addossata non all'utilizzatore ma al produttore che non ha implementato la macchina con tale sistema in base ad una colpa di programmazione.

Nel caso in cui l'evento si sia realizzato per l'errore di più soggetti nella catena di produzione e utilizzazione, si potrà ipotizzare un concorso di cause indipendenti (art. 41.3 c.p.) o una cooperazione colposa (art. 113 c.p.)[403]. La differenza tra i due casi sta nel fatto che nel primo si avrà tale concorso di più condotte che generano un unico evento in quanto poste da più persone ma l'una all'insaputa dell'altra. Mentre, per la cooperazione colposa, si necessita di una consapevolezza della convergenza della propria condotta con quella altrui andando a generare una colpa di concorso. La cooperazione colposa ex art. 113 c.p. riguarda, infatti, il caso in cui ciascuna delle parti versa in una responsabilità colposa per il medesimo fatto-reato. Tutti i compartecipi

[402] G. HALLEVY, op. cit., Paragrafo 3 (pag. 177 – 194) "Three models of the criminal liability of artificial intelligence entities"

[403] I. SALVADORI, op. cit., Paragrafo 5.2, "La responsabilità colposa"

sono in colpa e tutti potrebbero prevedere l'altrui comportamento: dunque, rispondono tutti per le pene del delitto stesso. Si ha quindi la necessità di una pluralità di persone, della realizzazione del reato, del contributo causale tra le condotte e della colpa nella condotta di partecipazione[404].

Il terzo modello di responsabilità, anche detto "the direct liability model", paragona l'intelligenza artificiale ai rei persone fisiche. Tale tipologia di responsabilità si va ad aggiungere alla responsabilità del programmatore o dell'utente umano, non la sostituisce ma al contempo esse non sono tra loro dipendenti. Ci si riferisce in questo caso all'ipotesi in base alla quale la macchina abbia deciso autonomamente, con coscienza e volontà, di commettere un illecito di cui dovrà necessariamente essere considerata responsabile. Essa possiede abilità di machine learning che le permettono di apprendere dall'esperienza e di effettuare ragionamenti incontrollabili dall'essere umano che ne resterebbe all'oscuro. In questo modo si paragona il pensiero intelligente umano all'autonomia razionale della macchina.

Come affermato precedentemente, tale teoria afferma la sussistenza della coscienza e volontà anche in capo ai sistemi di intelligenza artificiale[405]: ipotizza dunque, in capo al robot, un'anima pensante e sensibile paragonabile a quella umana.

Abbiamo più volte negato, all'interno di questo capitolo, la possibilità, ad oggi, di considerare le macchine come titolari di una personalità: esse detengono la razionalità e i valori morali immessi al loro interno da un programmatore, possono distinguere il giusto dallo sbagliato in merito a quanto impartito dall'algoritmo, ma la loro coscienza rimane limitata a ciò. I sistemi intelligenti sono determinati: seguono dei binari prestabiliti senza possibilità di muoversi al di fuori di essi; possono imparare e apprendere dall'esperienza ma solo in quanto programmate a farlo.

Per considerare un soggetto direttamente e penalmente responsabile è necessario che lo stesso possa essere considerato come imputabile e ad oggi, come espresso nel paragrafo precedente, i sistemi intelligenti non possono essere considerati come tali e dunque, è necessario escludere (almeno per ora) la possibilità di considerare la macchina come direttamente responsabile in quanto impossibile considerarla come titolare di una personalità elettronica giuridicamente rilevante.

---

[404] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione IV "Le forme di manifestazione del reato", Capitolo X "Tentativo e concorso di persone nel reato", Sotto-capitolo B "Il concorso di persone nel reato", Paragrafo 22 (pag. 546 – 549) "La cooperazione nel delitto colposo"

[405] G. HALLEVY, op. cit., Paragrafo 3 (pag. 177 – 194) "Three models of the criminal liability of artificial intelligence entities"

La responsabilità penale deve sempre essere riportata in capo ad un essere umano destinatario di diritti e doveri e di conseguenza sottoposto al diritto penale.
Quindi possiamo in realtà considerare due tipologie di responsabilità in capo alla persona: a titolo doloso o colposo.

3.1 Il problema della responsabilità oggettiva: responsabilità almeno a titolo di colpa

Come abbiamo affermato precedentemente, la responsabilità oggettiva, nel nostro ordinamento, in base alla Costituzione non è accettata. Sussistono però, nel codice penale alcuni casi residui come i reati aggravati da evento, il delitto preterintenzionale, l'aberratio ictus, l'aberratio delicti, il concorso anomalo (art. 116 c.p.) e il mutamento del titolo del reato per taluno dei concorrenti (art. 117 c.p.). Nonostante ciò, grazie ad un'interpretazione costituzionalmente orientata si deve sempre analizzare l'elemento soggettivo del reo almeno a titolo di colpa.
Analizzando i casi più vicini alle ipotesi che possono verificarsi in relazione all'intelligenza artificiale, per quanto concerne il delitto preterintenzionale sancito all'articolo 43 comma 2 c.p., il delitto è considerato oltre l'intenzione quando dall'azione o omissione deriva un evento dannoso o pericoloso più grave di quello voluto dall'agente. Il termine "oltre l'intenzione" sta a indicare la realizzazione di un fatto che va al di là dell'evento voluto risultando dunque più grave. Per giungere ad un'interpretazione costituzionalmente orientata è necessaria almeno la colpa intesa come prevedibilità dell'evento. È un dolo misto a colpa perché sussiste un dolo nell'intenzionalità dell'evento voluto e una colpa per l'evento in concreto realizzatosi (uno sviluppo prevedibile ed evitabile del fatto commesso secondo la diligenza dell'uomo ragionevole).
L'aberratio ictus (art. 82 c.p.) riguarda un'offesa ad una persona diversa da quella alla quale l'offesa era diretta dovuta ad un errore nell'uso dei mezzi di esecuzione (o altra causa). In tal caso il colpevole risponde come se avesse commesso il reato di danno alla persona che voleva offendere. Tale situazione è diversa da un errore di persona (art. 60 c.p.) perché è connessa ad un errore nella fase esecutiva del reato. La persona in concreto offesa non è oggetto del dolo perché non era a lei diretta l'azione. Sussiste quindi un dolo nel comportamento ma una colpa nel destinatario della condotta offensiva.
L'aberratio delicti (art. 83 c.p.) consiste in un evento diverso da quello voluto dall'agente a causa di un errore nell'uso dei mezzi di esecuzione del resto (o altra

causa). Il colpevole risponde a titolo di colpa per l'evento non voluto se previsto dalla legge come delitto colposo. La colpa deve essere effettivamente valutata per essere conforme all'articolo 27 comma 1 della Costituzione[406].

Il modello di imputazione della responsabilità indiretta dell'uomo entra in crisi a causa di sistemi intelligenti capaci di autoapprendere dall'esperienza e di conseguenza porre in essere decisioni e azioni autonomamente grazie ai meccanismi di machine learning. Il loro comportamento non è prevedibile e dunque risulta complesso individuare un soggetto persona fisica da considerare responsabile per tali azioni. Se l'azione imprevedibile della macchina era voluta dal soggetto, non sussistono problematiche dato che, anche nel caso di cambiamento del decorso causale voluto da quello realizzato, sussiste comunque un dolo in capo all'agente che l'ha voluto seguendo le regole del delitto preterintenzionale, aberratio delicti e aberratio ictus. Le problematiche riguardano il fatto che il funzionamento dei sistemi di intelligenza artificiale rimane spesso opaco senza riuscire ad avere chiarezza su come il robot sia giunto a tale risultato[407].

Ciò che caratterizza i sistemi di intelligenza artificiale è, infatti, un'imprevedibilità soggettiva, che non permette al programmatore, produttore ed utilizzatore di prevedere il risultato raggiungibile dal sistema intelligente, e una oggettiva legata all'opacità dell'algoritmo che nasce di per sé come imprevedibile. Per questo si pongono dubbi su chi debba rispondere nel caso di commissione di un reato da parte del sistema intelligente[408].

Le forme di responsabilità per danni causati dall'agente artificiale devono pur sempre ricadere sugli esseri umani in conformità con il principio di colpevolezza. Considerando responsabile il programmatore per non aver costruito il software al fine di evitare ogni possibilità di illecito, si estenderebbero eccessivamente i concetti della prevedibilità ed evitabilità dell'evento rischiando di uscire da un rimprovero colposo e rientrando in una responsabilità oggettiva. È sempre necessaria almeno la sussistenza della colpa da valutare sulla scorta della diligenza necessaria per evitare l'evento e il comportamento prevedibile del robot.

---

406 G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VIII "La Colpevolezza", Sotto-paragrafo A (pag. 352 – 423) "Dolo, colpa e dolo misto a colpa"

407 R. BORSARI, op. cit., Paragrafo 4 (pag. 264 - 266) "Le entità intelligenti di ultima generazione e la crisi del modello di imputazione della responsabilità indiretta dell'uomo"

408 B. PANATTONI, op. cit., Paragrafo 5 "Le sfide poste dall'«emergence»: il responsibility gap".

L'autonomia dei sistemi di intelligenza artificiale mette in crisi il nesso tra la colpa del programmatore o utilizzatore e l'evento che la macchina ha realizzato: risulta complesso dimostrare la prevedibilità e l'evitabilità del fatto commesso. La situazione antigiuridica che i robot potrebbero generare è costituita di due momenti: la fase esecutiva, in cui il sistema crea il modello ed esegue la funzione che gli è stata assegnata (esso risulta spesso imprevedibile) e l'evento lesivo.

Nasce quindi il problema del responsibility gap (di cui abbiamo fatto cenno precedentemente) in base al quale gli algoritmi di machine learning giungono a risultati di cui nessuno può essere considerato responsabile e dunque, generando situazioni eticamente inaccettabili. Per evitare che si realizzi tale lacuna, potrebbe risultare utile la Proposta di Regolamento del Parlamento Europeo[409] e del Consiglio sull'intelligenza artificiale, del 21 aprile 2021, che impone obblighi di controllo e di prudenza (con monitoraggi anche successivi all'immissione nel mercato) nel caso di sistemi intelligenti ad alto rischio: in tal modo potrebbe essere ipotizzato, per il futuro, l'utilizzo del medesimo schema per prevenire la commissione i reati dovuti a comportamenti autonomi della macchina.

Per evitare tale tipologia di responsabilità oggettiva o di situazioni non tutelabili a causa dell'elevata autonomia dei robot sussistono due approcci: un primo legato al principio di prudenza che si basa sulla limitazione dell'autonomia dei sistemi di intelligenza artificiale e un secondo mosso da una ricerca di soluzione al responsibility gap tramite il diritto penale limitando la responsabilità colposa dell'operatore grazie ad una progressiva accettazione sociale del rischio a fronte dei benefici che la stessa produce, oppure, tramite un adeguamento dei modelli di responsabilità oggi vigenti[410].

Nei prossimi paragrafi ci soffermeremo su questo secondo approccio.

3.2 Il problema del controllo significativo: la delegazione e la responsabilità da controllo indiretto

L'evoluzione tecnologica nella realizzazione dei sistemi di intelligenza artificiale, come affermato più volte, è giunta fino alla creazione di entità intelligenti capaci di muoversi nello spazio autonomamente e di prendere decisioni imparando

[409] COMMISSIONE EUROPEA, "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", COM2021/206 final, Bruxelles 21/04/2021: https://eur-lex.europa.eu/legal-content/it/txt/?uri=celex:52021pc0206

[410] B. PANATTONI, op. cit., Paragrafo 5.2 "Responsabilità colposa degli operatori per i comportamenti emergenti dei sistemi di IA"

dall'ambiente esterno immettendo dentro di sé input raccolti senza alcun intervento umano. Infatti, viene così ad enfatizzarsi il problema dell'imprevedibilità del funzionamento della macchina e soprattutto dell'output che genererà.

Nella catena di produzione di tali sistemi, le competenze iper-specialistiche di ingegneri, informatici, programmatori, produttori e tanti altri si fondono insieme senza che però gli uni possano avere un controllo sull'operato degli altri e di conseguenza, nessuno è in grado di prevedere ex ante il comportamento della macchina intelligente dotata di meccanismi di machine learning il cui algoritmo viene per questo definito black box[411].

L'essere umano ha una limitata capacità di prevedere il comportamento dei sistemi intelligenti in quanto si tratta di macchine dotate di capacità di autoapprendimento[412]. L'utilizzatore dunque, non può conoscere in anticipo il loro comportamento ma tale imprevedibilità è però prevedibile dal programmatore e dunque, da lui limitabile[413]. A sua volta, il programmatore può prevedere solo le situazioni ipotizzate da lui in base alle quali la macchina viene creata per agire e non anche quelle che sfuggono alla sua cognizione[414]. Per farlo rispondere del danno arrecato dal sistema è necessaria la dimostrazione che il programmatore avrebbe potuto prevedere ed evitare la verificazione dell'evento: un simile modello di responsabilità entra in crisi nel momento in cui neanche chi crea l'intelligenza artificiale è in grado di predire i suoi comportamenti ed effetti. I dubbi riguardano l'idea secondo cui per considerare fondata la responsabilità deve sussistere una violazione delle misure precauzionali da adottare nella valutazione del rischio della macchina che la legge esige da un agente modello in base all'attività che lo stesso svolge. Meno l'algoritmo che guida il sistema intelligente sarà oscuro, più sarà semplice dimostrare la responsabilità in capo al programmatore o utilizzatore[415].

È quindi il requisito dell'autonomia (più dell'adattabilità e interattività) ad essere caratteristico dell'intelligenza artificiale. Esso consiste nella capacità del sistema di svolgere la funzione per cui è stato programmato cambiando il proprio stato senza l'intervento dell'essere umano. I sistemi intelligenti, modificano il processo

[411] E. PALMERINI, op. cit., Paragrafo 6 "I problemi della sicurezza e della responsabilità nel mercato emergente della robotica"
[412] C. PIERGALLINI, op. cit., Paragrafo 4 "Machine learning e diritto penale"
[413] M. B. MAGRO, "Biorobotica, robotica e diritto penale", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali"
[414] C. PIERGALLINI, op. cit., Paragrafo 4 "Machine learning e diritto penale"
[415] M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, op. cit., Paragrafo 4 (pag.36 - 40) "Le difficoltà nella ricostruzione del nesso causale: "in dubio pro machina""

decisionale e generando soluzioni inaspettate al problema delegatogli, sfuggono, quindi, al controllo umano. Dunque, sussistono problematiche in relazione alla ripartizione della responsabilità[416].

Se quindi consideriamo il fatto che gli agenti artificiali sono in grado di autodeterminarsi, non risulterà possibile una predeterminazione ex ante dei loro comportamenti; per questo, molti li paragonano ad esseri umani penalmente responsabili. Qualora dovesse risultare possibile una comprensione a priori delle modalità di azione della macchina, la stessa verrà considerata come un mero strumento nelle mani dell'essere umano.

Sulla base dell'autonomia è possibile distinguere quattro livelli di intelligenza artificiale: al primo livello, i sistemi operano in modo automatico e sono sottoposti al controllo dell'essere umano (come le self- driving cars oggi disponibili sul mercato); al secondo, invece, gli stessi agiscono in base ad algoritmi deterministici risolvendo problematiche dettate dai programmatori; al terzo livello troviamo i sistemi intelligenti semi-autonomi dotati del meccanismo di machine learning che non necessita di seguire regole predeterminate dallo sviluppatore, infatti, è possibile effettuare solo un controllo sulle fasi di apprendimento correggendo e classificando i dati di input che vengono raccolti; infine, l'ultimo livello di autonomia è posseduto da sistemi capaci di interagire con altri agenti adeguando il proprio comportamento in base all'ambiente dove operano.

Per quanto riguarda le azioni commesse dai sistemi automatizzati di primo e secondo livello, la responsabilità penale sarà sempre in capo all'essere umano. Per i sistemi autonomi di terzo e quarto livello, le difficoltà per l'essere umano di controllare la macchina generano il c.d. responsibility gap che comporta un vuoto di tutela a causa della loro elevata autonomia. Infatti, gli agenti artificiali intelligenti completamente autonomi, grazie al meccanismo di auto machine learning, sono dotati di capacità di apprendimento e di problem solving con comportamenti non prevedibili ex ante da parte dell'essere umano[417].

Il rischio consiste, dunque, nella non comprensione e previsione dell'azione dell'intelligenza artificiale da parte dell'utente o programmatore, ma ciò viene bilanciato dai cospicui vantaggi che la stessa fornisce alla collettività e ai singoli. Infatti, grazie ai sistemi più evoluti di intelligenza artificiale è stato possibile delegare

[416] B. PANATTONI, op. cit., Paragrafo 2 "Dall'automazione tecnologica all'autonomia artificiale"
[417] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali"

compiti direttamente alle macchine, le quali possono agire senza l'intervento dell'essere umano e risultando anche più efficaci. La delegazione allo strumento diventa molto incisiva grazie a tale autonomia funzionale del robot. Il problema della delegazione è che, di fatto, affidando ad una macchina lo svolgimento di un compito, essa viene gravata anche di tutti i valori etici e morali connessi alle attività che deve svolgere: le tecnologie sono influenzate dalla società in cui vengono sviluppate e dunque seguono i suoi codici morali.

Connesso a tale tema, troviamo il problema dei bias: i pregiudizi eticamente rilevanti posseduti dal sistema intelligente perché inseriti da chi lo ha creato oppure basati su generalizzazioni fatte dalla macchia su set di dati incompleti raccolti in modo inadeguato. Si possono più o meno controllare le informazioni che vengono fornite all'algoritmo ma è quasi nulla la cognizione che si possiede in relazione alle modalità con cui l'algoritmo ha elaborato gli input e condotto all'output.

Tornando ai rischi della delegazione ai sistemi di intelligenza artificiale, questa genera una notevole deresponsabilizzazione dell'agire dell'essere umano, che tenderà a disinteressarsi dell'azione svolta dalla macchina in attesa del semplice risultato[418]. Infatti, con le forme di delegazione che l'intelligenza artificiale permette, l'essere umano viene posto ai margini del processo decisionale. È però necessario individuare un controllo umano facendo ricadere la responsabilità sul soggetto anche se non è su di lui che incombe la responsabilità causale del fatto. Maggiore è il grado di controllo che l'essere umano può esercitare, maggiore sarà la responsabilità penale personale dell'utilizzatore o programmatore[419].

La progressiva perdita di controllo da parte dell'operatore può essere ricondotta a diversi fattori: il capacity problem, a causa delle capacità computazionali del sistema, l'essere umano non ha la possibilità di controllare l'operato dello stesso in tempo reale; l'attentional problem, le possibili interruzioni dell'attenzione del controllore sulla macchina; l'attitudinal problem, la deresponsabilizzazione del soggetto difronte all'eccessiva fiducia che ripone nei confronti delle abilità del sistema intelligente; e i meccanismi di machine learning che generano un black box algorithm e dunque un'oscurità del processo che trasforma gli input in output. Ciò genera problematiche

[418] S. QUINTARELLI, op. cit., Capitolo 4 (pag. 72 – 92) "L'Etica dell'Intelligenza artificiale"
[419] S. QUINTARELLI, op. cit., Capitolo 6 (pag. 106 – 117) "Le leggi in gioco con l'intelligenza artificiale".

nell'ambito della riconnessione delle azioni della macchina alla suitas dell'essere umano da ritenere responsabile[420].

Com'è possibile notare si ha un'impossibilità di un controllo completo sull'esecuzione dell'algoritmo e dunque si necessita almeno di un controllo umano significativo[421]. Questo ha lo scopo di realizzare modalità che consentano all'utente di esercitare un controllo sugli aspetti rilevanti dell'azione autonoma della macchina soprattutto nei compiti ad essa delegati che possiedono un notevole valore etico e morale intrinseco. L'idea di base consiste nel non ritenere idoneo lasciare la gestione di tali situazioni esclusivamente al sistema intelligente. Così facendo, l'utente verrà messo nella condizione di valutare le implicazioni etiche delle azioni potendo scegliere in base alla sua coscienza nel rispetto della legge[422].

I soggetti coinvolti nel processo di vita del sistema intelligente sono: il produttore, il progettista del modello matematico e dei parametri di apprendimento, l'addestratore, coloro che raccolgono e selezionano i dataset, i venditori, gli acquirenti del sistema, gli utilizzatori e i fruitori finali. La Proposta di Regolamento dell'Unione Europea[423] riconduce tali soggetti in nuove nozioni: il provider, il soggetto che sviluppa il sistema; il distributor, colui che mette a disposizione il sistema all'interno del mercato europeo; e lo user, chi utilizza il sistema intelligente. Questa pluralità di operatori coinvolti comporta ovviamente problematiche in relazione all'individuare il soggetto penalmente responsabile. È quindi importante delineare le operazioni e la posizione ricoperta da ogni soggetto, così da individuare gli obblighi giuridici che gravano su ognuno. È chiaro che le fasi di progettazione e produzione sono le più delicate e pertanto i soggetti in esse coinvolti catalizzeranno le responsabilità della maggioranza dei casi. Inoltre, in tal modo si evita di responsabilizzare eccessivamente gli utilizzatori nel caso di difetti dovuti ad errori nella fase di produzione oppure nel caso in cui il fatto lesivo dipenda da un comportamento autonomo della macchina[424]. È infatti necessario che tutte le parti coinvolte si assumano la responsabilità di ciò che rientra nella loro sfera di signoria così da garantire la cognizione da parte dell'utente del

[420] B. PANATTONI, op. cit., Paragrafo 3.3 "Modelli di responsabilità e diversi livelli di autonomia: il "problema del controllo""

[421] G.UBERTIS, op. cit., Paragrafo 8 (pag. 83 – 84) "Necessità di un controllo significativo"

[422] S. QUINTARELLI, op. cit., Capitolo 4 (pag. 72 – 92) "L'Etica dell'Intelligenza artificiale"

[423] COMMISSIONE EUROPEA, "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", COM2021/206 final, Bruxelles 21/04/2021: https://eur-lex.europa.eu/legal-content/it/txt/?uri=celex:52021pc0206

[424] B. PANATTONI, op. cit., Paragrafo 3.3 "Modelli di responsabilità e diversi livelli di autonomia: il "problema del controllo""

proprio ruolo e della propria capacità di azione. Bisogna rendere chiari e semplici i processi di distribuzione della responsabilità in modo tale da sensibilizzare gli esseri umani delle azioni commesse dalla macchina sotto il loro controllo senza che gli stessi si sentano deresponsabilizzati a causa della delegazione del compito alla macchina. L'essere umano continua così ad esercitare il proprio giudizio morale senza delegarlo ad un sistema che non può farsene carico[425].

Vengono quindi in considerazione la causalità e la colpa.

I danni devono sempre essere ricondotti all'essere umano anche se l'azione non sembra connessa alle decisioni programmate. Si tratta del più volte richiamato fenomeno della black box secondo cui è possibile conoscere gli input immessi nel sistema ma non si possono programmare ex ante gli output a cui giungerà la macchina. Si pone quindi il problema della trasparenza, l'algoritmo non può spiegare il suo ragionamento. Tale gap potrebbe essere risolto installando scatole nere per comprendere ex post il processo decisionale posto dietro all'azione lesiva realizzata dalla macchina intelligente (ma così facendo si avrebbe un controllo meramente successivo non idoneo a prevenire il danno). Inoltre, seguendo l'articolo 41 comma 2 c.p., secondo cui "le cause sopravvenute escludono il rapporto di causalità quando sono state da sole sufficienti a determinare l'evento", se il danno è stato causato da un'azione autonoma del robot ed estranea al progetto iniziale della macchina, questa dovrebbe essere un fattore sopravvenuto ed interruttivo del nesso causale tra il programmatore del sistema e l'evento lesivo realizzatosi. L'intelligenza artificiale si è quindi intromessa tra la condotta del programmatore e l'evento penalmente rilevante innescando un rischio diverso da quello previsto dal soggetto.

In base alle teorie della causalità esplicate nel paragrafo precedente, l'evento causale sopravvenuto (la decisione autonoma presa dall'intelligenza artificiale) modifica il rischio iniziale ascrivibile alle azioni programmate ex ante, in uno nuovo dipendente dalla decisione della macchina. E dunque, l'eccezionalità dell'evento finale esulerebbe dalla sfera di controllabilità dell'essere umano non potendo quindi essere considerato responsabile il programmatore per l'azione del sistema intelligente: l'azione di programmazione non è di per sé in grado di produrre la lesione dato che il danno non era in quella fase ipotizzabile in quanto il comportamento della macchina programmata non è predeterminabile ex ante. In realtà però, chi programma tali sistemi dotati di auto machine learning sa di realizzare un sistema non completamente prevedibile e

[425] S. QUINTARELLI, op. cit., Capitolo 4 (pag. 72 – 92) "L'Etica dell'Intelligenza artificiale".

controllabile dall'essere umano. Non è quindi corretto negare la responsabilità dell'essere umano in base alla carenza del nesso di causalità perché il soggetto genera un rischio consapevole di non poter governare interamente le azioni della macchina.

La dinamica causale risulta fortemente complessa e ancor più rilevante, forse, è la problematica relativa all'accertamento della colpa a causa dell'imprevedibilità delle azioni del sistema intelligente. L'imprevedibilità, seppur prevedibile da chi realizza il sistema, nega la sussistenza di un elemento soggettivo. Si rischierebbe altrimenti di ricadere in responsabilità oggettiva perché verrebbe meno il concetto del rischio tipico se andassimo a considerare ogni evento realizzato dalla macchina come rimproverabile in base ad un'astratta prevedibilità iniziale anche se di per sé, in concreto, imprevedibile[426]. Una volta programmato, il sistema intelligente non fa più affidamento sul programmatore ed interagisce in maniera autonoma con il mondo. L'essere umano non può prevedere o controllare il comportamento del sistema che decide autonomamente.

Inoltre, sono numerosi i soggetti coinvolti nella realizzazione del fatto lesivo. Si potrà quindi avere una responsabilità distribuita in base al contributo causale di ciascuno e alle interazioni con il robot intelligente. Tutto ciò pone problematiche in relazione al sistema della responsabilità penale dell'uomo a titolo di colpa: l'essere umano che monitora il sistema non può prevedere il comportamento della macchina e non potrà neanche realizzare un controllo effettivo e costante sul suo operato[427].

L'essere umano non può essere considerato responsabile per i fatti illeciti commessi dal sistema intelligente, se, non è possibile da parte sua, esercitare un controllo significativo sulla macchina. Allo stesso tempo, se sussiste un danno connesso ad attività umane e non ad eventi naturali (come ad esempio un terremoto), qualcuno deve assumersene la responsabilità. Per evitare vuoti di tutela bisogna seguire due strade: individuare una forma di controllo umano significativo sul funzionamento autonomo dei sistemi di intelligenza artificiale, agendo a livello ingegneristico rendendo disponibile un qualche grado di controllo significativo sulla macchina; oppure, superare il limite del controllo per individuare la responsabilità dell'essere umano, intervenendo a livello concettuale, analizzando la responsabilità anche da altri elementi che non siano il mero controllo diretto degli eventi.

---

[426] C. PIERGALLINI, op. cit., Paragrafo 4 "Machine learning e diritto penale".

[427] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 2 (pag. 3 - 6) "Il problema dell'agire imprevedibile degli agenti artificiali: the black box alghoritms"

Rimangono comunque problematiche riguardanti l'effettiva possibilità di realizzare modelli nei quali il soggetto possa mantenere tale controllo e in cosa consista la significatività di tale controllo umano[428].

Per questo si parla, inoltre, di responsabilità da controllo indiretto del soggetto. Essa si basa sull'assunto che nessuno può essere in grado di controllare in modo diretto il sistema dell'intelligenza artificiale. Viene quindi realizzato un bilanciamento di interessi tra il soggetto da ritenere responsabile (che non potrebbe considerarsi tale in assenza dei poteri di impedimento dell'evento) e le vittime (che non vogliono vedersi sminuite le loro tutele).

Quando la macchina svolge la sua funzione in modo eccellente, sarà il team di programmatori ad assumersi la responsabilità del successo (come nel caso in cui "AlphaGo" sconfisse il campione mondiale di Go) e dunque non sussisteranno problemi di aggiudicazione della responsabilità anche se i programmatori non potranno essere considerati direttamente responsabili della strategia eseguita dalla macchina. Sulla base di tale assunto si può quindi affermare che la responsabilità non debba necessariamente essere connessa ad un controllo diretto, ma anche ad uno indiretto. I programmatori dovranno realizzare i sistemi di intelligenza artificiale (soprattutto per quelli "ad alto rischio" secondo la Proposta di Regolamento europeo del 21 aprile 2021)[429] che garantiscano una sufficiente trasparenza permettendo agli utenti di utilizzarli in modo appropriato al fine di permettere una supervisione ed un effettivo controllo umano[430].

Infatti, la capacità dei sistemi di adattarsi nel tempo all'ambiente e la sua capacità di apprendere e di auto-modificare il proprio programma grava i produttori di una maggiore responsabilità per l'anticipazione di eventuali problemi e l'introduzione di misure di salvaguardia. Le difficoltà riguardano l'impossibilità di individuare preliminarmente i difetti del prodotto, l'indecifrabile ragione della reazione della macchina all'ambiente che ha causato l'evento dannoso[431].

La figura del controllore umano opera come tutela per impedire che il malfunzionamento della macchina arrechi danni evitabili, altrimenti sarà lui a

[428] S. QUINTARELLI, op. cit., Capitolo 4 (pag. 72 – 92) "L'Etica dell'Intelligenza artificiale"

[429] COMMISSIONE EUROPEA, "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", COM2021/206 final, Bruxelles 21/04/2021: https://eur-lex.europa.eu/legal-content/it/txt/?uri=celex:52021pc0206

[430] B. PANATTONI, op. cit., Paragrafo 3.3 "Modelli di responsabilità e diversi livelli di autonomia: il "problema del controllo""

[431] E. PALMERINI, op. cit., Paragrafo 6.3 "Scenari presenti e futuri"

rispondere di tali eventi. Nel caso di danni da difetti di produzione del sistema che sfuggono al controllo umano, si potrà facilmente ricadere nell'illecito civile, mentre risulterà complesso valutare una responsabilità penale in capo al soggetto[432]. L'essere umano indirettamente responsabile per i fatti commessi dall'intelligenza artificiale avrà la possibilità di prova liberatoria andando a dimostrare di aver adottato tutte le cautele necessarie per evitare il danno.

Per ammettere il modello della responsabilità indiretta, è necessaria l'esistenza di un meccanismo che vada ad imporre ed assicurare un controllo umano significativo in modo tale che il soggetto possa adottare tutte le precauzioni utili ad impedire la lesione[433].

3.3 Il problema delle posizioni di garanzia ex art. 40 cpv. c.p.

In base all'articolo 40 cpv. c.p., secondo cui "non impedire un evento, che si ha l'obbligo giuridico di impedire, equivale a cagionarlo", si incrimina il mancato compimento di un'azione doverosa imposta per impedire un evento. Si tratta di un'omissione causalmente legata all'evento. Il titolare dell'obbligo giuridico viene considerato come un garante dell'integrità del bene giuridico tutelato. L'obbligo giuridico che grava su di lui si fonda su norme extra-penali contenute in fonti anche secondarie, in contratti, in contatto sociale qualificato etc. (taluno considera anche la precedente attività pericolosa). In base ad una concezione formale di tale articolo, si fa riferimento solo agli obblighi derivanti dalla legge o dal contratto. Il garante viene considerato come titolare di una posizione di garanzia e quindi ha obblighi di protezione (in base ai quali deve proteggere i beni giuridici da una gamma di pericoli) e obblighi di controllo (che mirano a neutralizzare le fonti di pericolo, umano o naturale, specifiche a tutela del bene minacciato)[434].

Sussiste inoltre il concorso mediante omissione, in base al quale è possibile avere un concorso di persone in forma omissiva seguendo due requisiti: la sussistenza di una posizione di garanzia secondo cui in capo al soggetto deve sussistere l'obbligo giuridico di impedire la commissione del reato da parte di altri e la necessità dell'omissione per la commissione del reato da parte del concorrente. Bisogna quindi

[432] C. PIERGALLINI, op. cit., Paragrafo 3 "Danno da prodotto e intelligenza artificiale sotto il controllo umano".
[433] S. QUINTARELLI, op. cit., Capitolo 6 (pag. 106 – 117) "Le leggi in gioco con l'intelligenza artificiale"
[434] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione III "Il Reato", Capitolo VI "Il Fatto", Sotto-capitolo B (pag. 257 – 271) "Le peculiarità del fatto nei reati omissivi".

accertare se l'azione doverosa che si è omesso di compiere avrebbe impedito la realizzazione del fatto concreto da parte dell'autore[435].

Gli eventi dannosi connessi all'uso dei sistemi intelligenti possono essere conseguenza di errori umani realizzatisi nelle attività di programmazione, sviluppo, collaudo e produzione; infatti, come accennato precedentemente, in tali fasi è possibile notare una molteplicità di soggetti che lavorano in team o meno per la creazione del prodotto intelligente. Spesso, tali attività vengono realizzate nell'ambito di organizzazioni imprenditoriali (società, aziende e imprese) specializzate nella realizzazione di robot con un personale altamente qualificato. Per queste ragioni, come esplicato precedentemente, risulta complesso individuare i soggetti responsabili di ogni contributo causale rischiando di ricadere in una responsabilità da posizione in base alla competenza del soggetto e al suo ruolo nella realizzazione del sistema intelligente.

La ripartizione dei poteri all'interno delle imprese è connessa di conseguenza ad una serie di posizioni di garanzia in modo tale da far rispondere il garante per gli errori relativi all'adempimento delle sue funzioni secondo la struttura dell'articolo 40 cpv. c.p.: sui garanti grava l'obbligo di controllo sulle fonti di pericolo che possono generare danni successivi[436].

L'individuazione dei soggetti responsabili si è evoluta nel corso della storia: durante l'illuminismo si considerava responsabile solo la persona fisica che aveva commesso il reato ledendo il bene giuridico della vittima; con la rivoluzione industriale nasce l'idea dell'impresa organizzata (ex art. 2082 c.c.) in cui sussiste una divisione del lavoro grazie ad un'organizzazione basata sulla fabbrica generando una spersonalizzazione (la competenza ad agire è data ad uffici e non a persone fisiche). L'impresa è compatta e unica vista dall'esterno e dunque sono riscontrabili difficoltà nell'individuazione dei soggetti responsabili come persone fisiche. Inizialmente per facilitare l'individuazione del responsabile si faceva riferimento esclusivamente al vertice dell'impresa in quanto rappresentante della stessa nei confronti dei terzi: tale soluzione risulta insoddisfacente dato che non sempre chi si trova al vertice è capace di intervenire sulle fonti di rischio e dunque si è poi optato per considerare come responsabile il soggetto che ha svolto le mansioni più vicine al reato. Oggi, invece, si guarda al soggetto che detiene il potere di controllo e di garanzia.

---

[435] G. MARINUCCI, E. DOLCINI, G.L. GATTA, op. cit., Sezione IV "Le forme di manifestazione del reato", Capitolo X "Tentativo e concorso di persone nel reato", Sotto-paragrafo B (pag. 523 – 557) "Il concorso di persone nel reato"

[436] I. SALVADORI, op. cit., Paragrafo 6, "Danno da agente artificiale e responsabilità plurisoggettiva".

Le posizioni di garanzia sono quindi connesse al fatto che il soggetto titolare del bene giuridico non sempre è in grado di proteggere il bene autonomamente anche a causa dell'impossibilità di conoscere e controllare i rischi; dunque, il garante risponde per la cattiva gestione delle fonti di rischio perché è lui il soggetto vincolato per legge al controllo dei pericoli nei confronti dei beni giuridici di soggetti indeterminati.

La posizione di garanzia si suddivide a sua volta in una posizione di protezione (proteggere soggetti determinati rispetto a rischi indeterminati) e di controllo (limitare rischi specifici per soggetti indeterminati).

Alla base della posizione di controllo è posta la responsabilità da mancato impedimento. Il garante risponde solo se possiede i poteri (originari o derivati) necessari per esercitare un effettivo controllo sulla fonte del rischio e se il suo comportamento sia in concreto esigibile. Il garante deve almeno conoscere l'evento da evitare per potergli muovere un rimprovero per la sua inerzia.

Essendo la struttura dell'impresa complessa, non si può affidare tutto al vertice e dunque si assiste ad un decentramento dei poteri considerando come datore di lavoro anche il responsabile dell'unità produttiva che ha adeguati poteri decisionali e di spesa. In tal modo, dopo aver accertato la causa del reato sul piano naturalistico, l'area di rischio e i soggetti che avevano il controllo su tale area, si potrà sanzionare il garante per non aver impedito l'evento (anche se nel caso dei reati commessi dall'intelligenza artificiale, il decorso causale risulta fortemente complesso). È infatti necessario analizzare l'incarico di fatto svolto dal soggetto a prescindere da una formale investitura: secondo l'articolo 2639 c.c. si deve effettuare un'estensione ai soggetti che esercitano di fatto i poteri tipici della figura presa in considerazione in modo continuativo (requisito quantitativo-temporale) e significativo (requisito qualitativo) ed inoltre si equiparano i soggetti che svolgono funzioni uguali ma con qualifiche differenti. Tale articolo si applica limitatamente ai reati societari ma sarebbe stato opportuno inserirlo all'interno del codice penale così da poter estendere la disciplina ad altre situazioni analoghe: infatti, la giurisprudenza lo ritiene applicabile anche ad altri settori (tra cui il fallimentare).

La Sentenza Thyssenkrupp[437] esprime il principio secondo cui il giudice deve valutare in concreto quali fossero gli strumenti a disposizione dei soggetti per prevenire il reato:

[437] Sentenza della Cassazione Penale, Sezioni Unite, 18 settembre 2014 n. 38343: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjXkPux-5X1AhUN_aQKHdC4B1AQFnoECAIQAQ&url=https%3A%2F%2Fwww.vegaengineering.com%2Fnews%2Fallegati%2F2194 %2F1%2FSentenza-n.38343-Thyssenkrupp.pdf&usg=AOvVaw3UahHk9Z6ouVslRdrxQOHD

l'indagine va svolta in concreto e non basta un semplice obbligo di fare il possibile, altrimenti si avrebbe un'indeterminatezza della condotta penale. Bisogna inoltre effettuare un accertamento controfattuale valutando in via ipotetica se la condotta attiva di controllo e iniziativa avrebbe potuto impedire il verificarsi dell'evento[438].

Le posizioni di garanzia gravano sugli apicali e sui soggetti subordinati che posseggono un sapere tecnico fondamentale per la realizzazione dei sistemi intelligenti. Essi dovranno adottare le norme cautelari e di prevenzione per evitare i rischi legati alla produzione delle macchine intelligenti. Dunque, il mancato rispetto di tali obblighi potrà sfociare in una responsabilità omissiva per i fatti lesivi realizzati dal robot se sono espressione del pericolo che il soggetto aveva il compito di prevedere ed evitare.

Infatti, secondo la Direttiva relativa alla responsabilità per danno da prodotto difettoso (Direttiva 85/374/CEE)[439], di cui tratteremo successivamente, i produttori hanno l'obbligo di mettere in commercio dei prodotti che siano sicuri a seguito di costanti manutenzioni e collaudi. Quindi sussiste in capo a loro un obbligo di controllo per la sicurezza di tali prodotti. Inoltre, ipotizzando l'applicazione della Proposta di Regolamento dell'Unione Europea[440] anche ad altri settori dell'ordinamento, essa disciplina le regole cautelari da adottare in base al rischio intrinseco del sistema intelligente. Di conseguenza, in base a tale ragionamento interpretativo, se non dovessero essere rispettate le regole cautelari si potrà rientrare in una responsabilità penale; se invece, il danno si dovesse realizzare comunque, nonostante l'adozione delle precauzioni necessarie, non potremmo considerare il produttore o programmatore come responsabile, rientrando dunque in una area di rischio consentito. Quindi, bisognerà valutare caso per caso la presenza sia di un dovere di controllo che di un potere effettivo di impedire l'evento nel rispetto delle regole cautelari intrinseche all'attività pericolosa che si sta svolgendo.

I due elementi a cui si deve far riferimento nella formulazione dei regimi di imputazione della responsabilità penale nell'ambito dell'intelligenza artificiale sono:

---

[438] A. ALESSANDRI E S. SEMINARA, "Diritto penale commerciale", Volume 1 "I principi generali", G. Giappichelli Editore, Torino 2018, Capitolo 2 (pag. 43 – 87) "La responsabilità penale della persona fisica".

[439] Direttiva 85/374/CEE del Consiglio del 25 luglio 1985 relativa al "Ravvicinamento delle disposizioni legislative, regolamentari ed amministrative degli Stati Membri in materia di responsabilità per danno da prodotti difettosi": https://eur-lex.europa.eu/legal- content/IT/TXT/PDF/?uri=CELEX:31985L0374&from=IT

[440] COMMISSIONE EUROPEA, "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", COM2021/206 final, Bruxelles 21/04/2021: https://eur-lex.europa.eu/legal-content/it/txt/?uri=celex:52021pc0206

la multi-soggettività connessa alla realizzazione dei sistemi intelligenti e il rispetto del principio di colpevolezza. A causa del fatto che nessun soggetto ha la capacità di gestire autonomamente l'intero percorso decisionale della macchina e neanche le fasi della sua realizzazione, si potrà far riferimento solo ad una responsabilità colposa. Dunque non bisognerà configurare una responsabilità esclusivamente in capo ad un unico soggetto, ma una responsabilità diffusa tra i vari operatori che sono intervenuti nel processo di realizzazione del sistema basandosi sul modello della responsabilità penale degli enti[441].

Come affermato nel primo paragrafo di questo capitolo, invece, non è ad oggi possibile effettuare una semplice analogia tra la responsabilità dei sistemi intelligenti con la responsabilità degli enti in base d.lgs. 231/2001[442] dato che, nonostante si sia indotti ad effettuare tale parallelismo in quanto, al pari degli enti, i robot sono dotati di intrinseca capacità criminale che permette una loro responsabilità penale, in realtà tale paragone risulta fallace dato che i sistemi intelligenti, a differenza degli enti, non sono composti da esseri umani e non possono essere considerati come responsabili[443]. Allo stesso tempo, il modello della responsabilità degli enti fornisce un ottimo spunto metodologico ragionando ad esempio sull'opportunità di ricorrere a contaminazioni tra branche del diritto differenti (penale, amministrativo e civile) in modo tale da riuscire a ricreare un assetto normativo coerente e idoneo a tutelare la collettività[444].

Per evitare un'eccessiva responsabilizzazione e al contempo garantire tutele ai danneggiati, servirà trovare un punto di equilibrio in cui accettare un margine di rischio consentito, considerando punibile tutto ciò che va oltre. In tal modo sarà possibile considerare lecita l'attività di produzione e progettazione del sistema intelligente eseguita nel rispetto delle regole cautelari. Nel caso in cui si dovessero riscontrare lesioni ai beni giuridici di terzi nonostante tali accorgimenti, al soggetto che ha realizzato il robot non potranno essere mossi rimproveri colposi in quanto i danni commessi rientrano nel rischio consentito[445].

---

[441] B. PANATTONI, op. cit., Paragrafo 6 "Responsabilità diffuse eLegal Protection by Design"

[442] Decreto legislativo 8 giugno 2001, n. 231, "Disciplina della responsabilità amministrativa delle persone giuridiche, delle società e delle associazioni anche prive di personalità giuridica", a norma dell'articolo 11 della legge 29 settembre 2000, n. 300.

[443] P. SEVERINO, "Intelligenza artificiale e diritto penale", Paragrafo 2 (pag. 533 – 536) "La responsabilità penale degli agenti intelligenti" in RUFFOLO U., "Intelligenza artificiale – Il diritto, i diritti, l'etica", Collana "Tech and Law" di Giuffrè Francis Lefebvre Edizione, Milano 2020; Parte IV "Diritto processuale, penale, costituzionale, amministrativo, tributario, eurounitario", Sezione V "A.I. e diritto penale" Capitolo 5 (pag. 531 – 547).

[444] B. PANATTONI, op. cit., Paragrafo 6 "Responsabilità diffuse eLegal Protection by Design"

[445] I. SALVADORI, op. cit., Paragrafo 6, "Danno da agente artificiale e responsabilità plurisoggettiva"

La concezione etica del rischio graduata in base alla tollerabilità sociale dovrà indirizzare le future scelte politiche al fine di governare tali situazioni[446].

3.4 Il problema dell'individuazione del responsabile: colpa eventuale del programmatore

Le macchine dimostrano abilità di autoapprendimento dall'esperienza assumendo decisioni imprevedibili. Ma allo stesso tempo non è ancora possibile considerarle come soggetti di diritto ma solo come agenti artificiali di una persona fisica o giuridica.

Tali difficoltà interpretative del processo decisionale dei sistemi intelligenti generano dibattiti in tema di progettazione responsabile di tali tecnologie e di responsabilità per colpa dell'utilizzatore nei casi di azioni incontrollabili da parte del soggetto umano. È importante limitare la responsabilità dei programmatori per evitare di ricadere in forme di responsabilità oggettiva e per evitare di porre un freno allo sviluppo tecnologico di tali sistemi intelligenti. Tutto si concentra sul tema della prevedibilità e sulla possibilità di punire i programmatori per non aver previsto i potenziali usi dannosi alternativi che possono essere realizzati con il sistema intelligente da loro ideato. La mancanza di una prevedibilità e spiegabilità dell'agire della macchina non solleva il programmatore da una possibile responsabilità per "colpa eventuale" dato che la stessa non richiede una prevedibilità dell'evento hic et nunc limitandosi ad una prevedibilità astratta del rischio anche concretamente imprevedibile che oltrepassa il limite del rischio consentito. Non si può negare che il programmatore sia astrattamente a conoscenza dei possibili rischi connessi all'utilizzo del sistema intelligente.

Ma così facendo si eliderebbe gran parte del contenuto della colpa andando a minimizzare la dimensione soggettiva della colpevolezza riducendola ad una mera inottemperanza cautelare[447].

Una delle ipotesi per ovviare a tali problematiche riguarda, per il settore penale, l'applicazione della responsabilità penale da prodotto sancita dalla Direttiva 85/374/CEE[448]. In base ad essa, il produttore risulta responsabile del danno causato dal suo prodotto ma sarà il danneggiato a dover provare il danno, il difetto e il nesso

[446] C. PIERGALLINI, op. cit., Paragrafo 2 "Danno da prodotto e Intelligenza Artificiale"

[447] M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 9 (pag. 19 - 22) "La responsabilità a titolo di colpa (eventuale) dello sviluppatore da procreazione di agenti artificiali. Il principio della produzione robotica responsabile e benefica".

[448] Direttiva 85/374/CEE del Consiglio del 25 luglio 1985 relativa al "Ravvicinamento delle disposizioni legislative, regolamentari ed amministrative degli Stati Membri in materia di responsabilità per danno da prodotti difettosi", op. cit.

causale (il che sembrerebbe evocare una responsabilità di tipo oggettivo, che però, se connessa al diritto penale sarebbe incostituzionale e dunque andrebbe interpretata in senso costituzionalmente orientato considerando almeno l'elemento della colpa). Per prodotto difettoso, ai sensi della direttiva, si intende un prodotto che non offre la sicurezza che si può legittimamente attendere in base alla sua presentazione e all'uso a cui è destinato.

Di conseguenza, con l'avvento dell'intelligenza artificiale, è stato avviato un processo di esame di tale disciplina al fine di valutare se la stessa possa garantire un adeguato livello di protezione per i danneggiati da tali sistemi. Difatti, il concetto di prodotto, secondo la Direttiva consiste in ogni bene mobile anche se forma parte di un altro bene mobile o immobile. In tal modo, è possibile far rientrare anche i sistemi di intelligenza artificiale all'interno di tale definizione nonostante taluno[449] ponga gli stessi dubbi che si posero nel considerare il software come bene mobile[450].

Ma allo stesso tempo, si pongono delle problematiche con l'applicazione di tali tipologie di tutele ai sistemi basati sull'intelligenza artificiale a causa delle loro caratteristiche particolari.

Infatti, l'evento lesivo realizzato dalla macchina può derivare dalla semplice elaborazione degli input da parte del sistema, anche in assenza di difetti di fabbricazione e nonostante gli operatori abbiano rispettato gli obblighi cautelari di sicurezza del prodotto. Sarà quindi necessario riformulare il concetto di "difetto" facendoci rientrare anche le variazioni del processo decisionale realizzato dalla macchina rispetto a quanto fu programmato ex ante. E sarà necessario superare l'esclusione della responsabilità del produttore per i difetti sorti successivamente alla messa in commercio del prodotto (art. 7 della Direttiva 85/374/CEE).

Le problematiche in ambito della responsabilità penale riguardano la crisi della causalità connessa alla complessità del prodotto e all'oscurità del meccanismo di machine learning che pone inoltre problematiche in relazione alla colpa a causa dei comportamenti autonomi della macchina non prevedibili ex ante e dunque non evitabili. Infatti, senza un intervento modificativo, si rischierebbe di far rispondere il

[449] PONZANELLI, "Responsabilità per danno da computer: alcune considerazioni comparative", in Resp. Civ. prev., 1991, 650

[450] A. AMIDEI, "Intelligenza artificiale e responsabilità da prodotto", Paragrafo 3 (pag. 131 – 135) "L'A.I. come «prodotto» e l'algoritmo quale suo componente" in U. RUFFOLO, "Intelligenza artificiale – Il diritto, i diritti, l'etica", Parte II "Diritto civile: responsabilità, contratti, persona e privacy", Sezione I "A.I. e responsabilità" Capitolo 2 (pag. 125 – 153).

programmatore a prescindere da una sua possibile previsione o volontà dell'evento dannoso evitabile: di conseguenza, sarebbe utile posse come limitazione della responsabilità il fatto che lo stesso abbia realizzato tutte le condotte idonee al fine di scongiurare la realizzazione dell'evento lesivo comunque realizzatosi. Ed inoltre si amplificano le possibilità di ricorrere ad una responsabilità plurisoggettiva a causa della diversità degli operatori che intervengono nella realizzazione dell'intelligenza artificiale[451].

Le modalità per superare il responsibility gap possono essere inquadrate in due proposte.

Innanzitutto, qualificare il comportamento imprevedibile della macchina come autonoma decisione elaborata dallo stesso andrebbe a limitare la responsabilità degli operatori in quanto causa sopravvenuta capace di provocare da sola l'evento lesivo interrompendo il nesso causale tra la programmazione o utilizzazione del sistema da parte della persona fisica e il fatto-reato[452]. Allo stesso tempo, parte della dottrina[453] solleva critiche a tale prospettiva dato che potrebbe sottovalutare i rischi consapevolmente e volontariamente realizzati dagli operatori e andrebbe a limitare l'analisi di possibili comportamenti colposi commessi dal soggetto. Si andrebbe così ad accentuare la responisiblity gap non considerando responsabile né il sistema intelligente, né l'essere umano.

In secondo luogo, il superamento del responisiblity gap è realizzabile attraverso l'imposizione di obblighi giuridici di comportamento in capo agli operatori in modo da garantire la loro diligenza: qualora tali doveri vengano rispettati, se il sistema intelligente dovesse realizzare illeciti, loro non ne sarebbero responsabili in quanto hanno posto in essere tutte le condotte giuridicamente doverose per evitare e prevenire il reato. La realizzazione di tali eventi lesivi potrà al massimo ricadere in forme di responsabilità esclusivamente civile. Nonostante ciò, i rischi di commissione degli illeciti rimangono e dunque bisognerà circoscrivere un'area di rischio tollerato e adeguato in virtù dei benefici che tali tecnologie portano con loro[454].

Si dovrà quindi fare riferimento non tanto alle azioni emergenti realizzate dal sistema intelligente, quanto agli obblighi di diligenza gravanti sul programmatore al fine di

[451] B. PANATTONI, op. cit., Paragrafo 4 "La responsabilità penale per danno da prodotto"
[452] B. PANATTONI, op. cit., Paragrafo 5.4 "Possibili soluzioni al responsibility gap"
[453] S. GLESS, E. SILVERMAN, T. WEIGEND, "If robots cause harm, who is to blame? Self-driving cars and criminal liability", in New Criminal Law Review, 2016, pag. 430 e ss..
[454] B. PANATTONI, op. cit., Paragrafo 5.4 "Possibili soluzioni al responsibility gap"

realizzare un robot basato sull'intelligenza artificiale che sia conforme ai principi ART (accountability, responsability e transparency). Ciò emerge dagli obblighi di controllo e valutazione delineati all'interno della Proposta di Regolamento dell'Unione Europea[455] del 21 aprile 2021[456].

È necessario elaborare un nuovo concetto di colpa nel quale escludere la colpevolezza qualora vengano rispettate le misure di cautela imposte andando ad applicare in concreto le tre leggi della robotica di Asimov[457]. Si potrà così considerare colposo il comportamento del programmatore che non vada a limitare l'imprevedibilità del sistema intelligente immettendo dei vincoli e dei sistemi di controllo del robot. Si rispetterebbero così i principi di precauzione e di colpevolezza legata al mancato rispetto delle regole cautelari da parte del programmatore che sfociano in una sua colpa eventuale dovuta ad un'astratta prevedibilità del rischio che non viene da lui limitato[458].

Il diritto penale connesso alla c.d. "società del rischio" (dove per rischio si intende la possibilità di realizzazione di un evento lesivo) sta cambiando paradigma: si passa da una prospettiva ex post, di punizione per il delitto connesso, ad una ex ante, in cui vengono individuati obblighi di cautela da dover rispettare prima di realizzare un'azione intrinsecamente pericolosa.

Infatti l'assunto base di tale concetto è connesso all'impossibilità di conoscere tutti i fattori di rischio che ci circondano e di conseguenza il problema della precauzione va valutato caso per caso in una prospettiva influenzata da fattori culturali e sociali. L'evoluzione tecnologica ha sicuramente fornito maggiori strumenti verso l'individuazione dei rischi ma al contempo ne ha generati di nuovi (e di difficile comprensione)[459].

Tale ideologia si sposa perfettamente con la scelta politica di non frenare il progresso scientifico limitandone al contempo i possibili effetti negativi sulla società.

Per bilanciare tali due contrapposte esigenze, bisognerà dunque individuare un margine di rischio consentito allo scopo di tutelare i beni giuridici che possono essere

---

[455] COMMISSIONE EUROPEA, "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", COM2021/206 final, Bruxelles 21/04/2021: https://eur-lex.europa.eu/legal-content/it/txt/?uri=celex:52021pc0206

[456] B. PANATTONI, op. cit., Paragrafo 6 "Responsabilità diffuse eLegal Protection by Design"

[457] I. ASIMOV, "I, Robot", Mondadori, 1950.

[458] M.B. MAGRO, "Biorobotica, robotica e diritto penale", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali".

[459] S. ARCIERI, "Percezione del rischio e attribuzione di responsabilità", in Diritto Penale e Uomo (DPU) – Criminal Law and Human Condition, Fascicolo 10/2020, 28 ottobre 2020, Milano: https://dirittopenaleuomo.org/wp- content/uploads/2020/10/douglas_DPU.pdf

lesi dall'intelligenza artificiale senza frenarne lo sviluppo. Di conseguenza, gli sviluppatori e produttori dovranno realizzare sistemi intelligenti che siano sicuri (testandoli con regolarità e prevedendo aggiornamenti migliorativi) per la collettività: se ciò non fosse possibile, allora il prodotto dovrebbe essere ritirato dal mercato.

Solo qualora il produttore, a conoscenza dei difetti del prodotto intelligente, non si attivi per risolverli, potrà essere ritenuto responsabile a titolo doloso o colposo per non aver impedito l'evento dannoso verificatosi. Qualora invece, abbia osservato tutte le regole cautelari possibili non potremmo considerarlo responsabile: si rientrerebbe in un rischio accettato dalla collettività.

Il margine di rischio consentito dovrebbe essere valutato dal legislatore sulla base di regole cautelari impose e tipizzate: una strada che l'Unione Europea sembra voler imboccare a seguito della citata Proposta di Regolamento[460].

Dovrebbero dunque essere imposti obblighi di informazione nei confronti delle competenti autorità pubbliche, un sistema di autorizzazioni amministrative sulla base di parametri tecnico-precauzionali sottoponendo il prodotto intelligente ad una rigida procedura di controllo sulla sperimentazione e produzione[461].

Serviranno quindi interventi legislativi che possano implementare la materia inserendo specifici obblighi di monitoraggio e di revisione periodica dei sistemi intelligenti. Infatti, si dovrà valutare se l'azione commessa dal robot rientri in un utilizzo dell'intelligenza artificiale difforme rispetto a quello a cui era destinato ma comunque prevedibile; oppure da una modifica del sistema dopo l'immissione sul mercato che incida sull'uso a cui era predisposto. Si fissa così un nuovo confine di diligenza.

Risulterà forse utile implementare il nostro ordinamento con forme di responsabilità intermedia tra l'illecito penale e amministrativo, realizzando modelli in grado di regolare anche i fatti illeciti causati da comportamenti imprevedibili dei sistemi intelligenti, senza gravare gli operatori di una responsabilità oggettiva e al contempo tutelando le vittime di tali reati.

In più, oltre ad ipotizzare una responsabilità in capo ai programmatori e produttori per non aver correttamente gestito il rischio, come abbiamo già ipotizzato precedentemente e come osserveremo nel prossimo paragrafo, i problemi potrebbero

[460] COMMISSIONE EUROPEA, "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", COM2021/206 final, Bruxelles 21/04/2021: https://eur-lex.europa.eu/legal-content/it/txt/?uri=celex:52021pc0206

[461] I. SALVADORI, op. cit., Paragrafo 9 "Considerazioni finali".

astrattamente risolversi inserendo nel sistema un "algoritmo del diritto" contenente al suo interno tutti i principi etici e le norme espresse dal nostro ordinamento.

Tale proposta però sembra più facile a dirsi che a farsi, infatti, ad oggi, i sistemi intelligenti si muovono nel mondo come esseri umani, ma senza di fatto esserlo; di conseguenza, anche l'interpretazione delle norme in modo elastico risulterà fortemente complessa da replicare in un robot. In situazioni complesse, in cui sono posti in gioco differenti fattori di rischio e beni giuridici da tutelare, il programmatore dovrà istituire un criterio nella macchina che la stessa seguirà per giungere ad una decisione: è possibile ipotizzare l'immissione di norme all'interno di detto criterio e la creazione di criteri di razionalità basati ad esempio sul minor danno. Ma fino ad ora, nessuna macchina intelligente è riuscita ad eguagliare la mente umana in questo campo; infatti, solo la sensibilità e la capacità di ragionamento di un essere umano permettono di giungere ad una piena comprensione del mondo esterno e di conseguenza, detteranno i binari da seguire per agire concretamente in situazioni complesse.

Di certo, riuscire ad inserire "l'algoritmo del diritto" nei sistemi intelligenti risulterebbe un notevole passo avanti nel bilanciamento di interessi dato dalla "società del rischio".

CONCLUSIONI

Secondo una definizione accettata a livello internazionale, l'"Intelligenza Artificiale" è quella disciplina, appartenente all'informatica, che studia i fondamenti teorici, le metodologie e le tecniche che consentono di progettare sistemi hardware e sistemi di programmi software capaci di fornire all'elaboratore elettronico delle prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana.

La rapida evoluzione dell'Intelligenza Artificiale e le sempre più numerose applicazioni che essa trova nei diversi settori della vita quotidiana impongono una profonda riflessione sulle sue implicazioni in ambito giuridico. Qui interessano, in particolare, quelle riguardanti il diritto penale (sostanziale), dove lo straordinario sviluppo dell'Intelligenza Artificiale dell'ultimo decennio ha sollevato questioni assai delicate.

Se in ambito civilistico, è possibile rintracciare a livello europeo un quadro legislativo, costituito essenzialmente da tre direttive europee, la Direttiva 2006/42/CE (direttiva "Macchine"), la Direttiva 01/95/CE e infine la Direttiva 99/44/CE, alle quali si aggiunge la proposta di risoluzione approvata dal Parlamento Europeo nel 2017 che prevede "Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica".

Nel settore penale, viceversa, il panorama attuale è ancora acerbo e meritevole di interventi.

Dalle quattro analisi effettuate emerge chiaramente la mancanza di una legislazione di riferimento unitamente a una normativa attuale che, così come è interpretata e applicata, risulta ostativa all'applicazione di queste nuove tecnologie nell'ambito processual-penalistico.

Emerge inoltre come una delle chiavi di lettura del rapporto tra diritto penale e IA sia la responsabilità, intesa sia dal punto di vista dell'uomo, responsabile della programmazione degli algoritmi, del loro uso ed eventuale abuso, sia anche (potenzialmente) della macchina stessa, qualora commettesse (più o meno autonomamente) o fosse usata per commettere un reato.

Deve essere in ogni caso chiaro che l'approccio all'IA non può prescindere dalla considerazione che si tratta di una tecnologia fallibile, che non promette risposte o risultati dotati di assolutezza e certezza; alla luce di ciò, è bene che l'intervento e il controllo umano siano sempre valorizzati e garantiti, anche e soprattutto a livello legislativo.

Per concludere sembra potersi apprezzare lo sforzo del Parlamento, e più in generale delle istituzioni europee, di stare al passo con l'evoluzione tecnologica e con i rischi che essa pone, anche e soprattutto in un settore sensibile come quello penale. Se è vero che i tentativi di regolazione sono ancora in una fase preliminare – e la vaghezza di alcune disposizioni della Risoluzione ne sono una chiara dimostrazione – è indiscutibile la volontà dell'UE di assumere un ruolo da protagonista nella regolazione dell'IA. In particolare, le disposizioni della Risoluzione sembrano andare verso l'accettazione che l'IA, nel settore penale, avrà un ruolo sempre più pregnante, e, di conseguenza, sempre più fertile di rischi per la tutela dei diritti: se questo è inevitabile, appare dunque doveroso preservare un diritto penale costituzionalmente orientato, liberale e democratico, che respinga con forza le tendenze securitarie e di risposta preventiva che l'intelligenza artificiale ha già mostrato di poter pericolosamente favorire.

Un tempo si diceva - e forse era anche un modo per autoconvincersi da parte della società degli umani sulla sua permanente supremazia su qualsiasi altra realtà- che i computer fanno solo quello che sono programmati a fare.

Oggi la frase pare una di quelle cartoline sbiadite che raccontano storie ormai lontane.

Gli enormi progressi fatti nel campo della robotica, in particolare del machine learning, di pari passo con gli sviluppi rapidissimi (e a costi sempre più accessibili) della tecnologia computazionale, ha permesso l'implementazione di sistemi di Artificial Intelligence sempre più evoluti, in grado ormai di raggiungere e superare le capacità umane in molti settori.

Nel presente lavoro si è tentato di dimostrare che quando parliamo di Intelligenza artificiale dobbiamo, ormai, pensare ad una disciplina sostanzialmente ingegneristica, che poco ha a che vedere con l'intelligenza (oltre che con la conformazione del corpo) umana. Ne è una sorta di "copiatura" perfetta, una macchina con un "codice genetico" ingegnerizzato basato sulla razionalità, ovvero la capacità di scegliere sempre la migliore alternativa possibile secondo criteri algoritmici di ottimizzazione delle risorse a disposizione. E muta la forma a seconda dei molteplici ambiti in cui trova applicazione.

L'enorme progresso delle potenzialità computazionali delle macchine, come si diceva, unito al grande aumento di circolazione dei Big Data, e in larga parte anche di dati scambiabili machine to machine (generati dall'interconnessione degli strumenti dell'IoT), sta creando ciò che solo nel secolo scorso sembrava ancora poco più che fantascienza: un software sempre più intelligente che apprende dall'ambiente esterno e modifica le proprie

prestazioni adattandole ad esso, indipendentemente dall'imprinting iniziale del programmatore.
Ciò, ovviamente, crea di pari passo rilevanti implicazioni socio-culturali, etiche (rectius, bioetiche) e giuridiche.

BIBLIOGRAFIA

K. ABNEY, G.A. BEKEY, R. JENKINS, Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence, Oxford, 2017, p. 339.

I. AJUNWA, Genetic Testing Meets Big Data: Tort And Contract Law issues, in 75 Ohio St. L.J. 1225, 2014

I.AJUNWA, Algorithms At Work: Productivity Monitoring Platforms And Wearable Technology As The New Data-Centric Research Agenda For Employment And Labor Law, in 63 St. Louis U. L.J. 2019

E. AL MUREDEN, La responsabilità del fabbricante nella prospettiva della standardizzazione delle regole sulla sicurezza dei prodotti, in AL MUREDEN E., La sicurezza dei prodotti e la responsabilità del produttore: casi e materiali, Giappichelli, Torino, 2017

A.ALESSANDRI E S. SEMINARA, "Diritto penale commerciale", Volume 1 "I principi generali", G. Giappichelli Editore, Torino 2018, Capitolo 2 (pag. 43 – 87) "La responsabilità penale della persona fisica".

A. AMIDEI, "Intelligenza artificiale e responsabilità da prodotto", Paragrafo 3 (pag. 131 – 135) "L'A.I. come «prodotto» e l'algoritmo quale suo componente" in U. RUFFOLO, "Intelligenza artificiale – Il diritto, i diritti, l'etica", Parte II "Diritto civile: responsabilità, contratti, persona e privacy", Sezione I "A.I. e responsabilità" Capitolo 2 (pag. 125 – 153).

S. ARCIERI, "Percezione del rischio e attribuzione di responsabilità", in Diritto Penale e Uomo (DPU) – Criminal Law and Human Condition, Fascicolo 10/2020, 28 ottobre 2020, Milano: https://dirittopenaleuomo.org/wp- content/uploads/2020/10/douglas_DPU.pdf

P. BALDUCCI E A. MACRILLÒ, "Esecuzione penale e ordinamento penitenziario", Giuffrè Francis Lefebvre S.p.a. Milano – 2020

W. BARFIELD, Towards a law of artificial intelligence, in BARFIELD W., PAGALLO U., Research Handbook on the Law of Artificial Intelligence, Ed. Edward Elgar Publishing, 2018

S. BAROCAS - A. D. SELBSDT, Big Data's Disparate Impact, in 104 Calif. L. Rev. 671, 674 N.10, 2016
S. BAROCAS-S. HOOD-M. ZIEWITZ, Governing Algorithms: A Provocation Piece, Governing Algorithms, 29 Mar. 29, 2013

F. BASILE, Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine, in Diritto Penale e Uomo, 2019

M. BASSINI, L. LIGUORI, O. POLLICINO, Sistemi di Intelligenza Artificiale, responsabilità e accountability. Verso nuovi paradigmi?, in F. PIZZETTI (a cura di), Intelligenza artificiale.

G. A. BEKEY, Current Trends in Robotics: Technology and Ethics, in LIN P., ABNEY K., BEKEY G. A., Robot Ethics: The Ethical and Social Implications of Robotics, 2011, p. 17 ss

J. BERNSTEIN, Uomini e macchine intelligenti, Adelphi, Milano, 2013

M.A. BODEN, Intelligenza artificiale, in J. l-Khalili (a cura di), Il futuro che verrà, Bollati Boringhieri, 2018, p. 133

R. BORRUSO, "Riflessioni sull'informatica giuridica", in: L'informatica del diritto, Milano 2004

M.C. BROMBY, M.J HALL, The Development and Rapid Evaluation ofthe Knowledge Model of advokate: anAdvisory System to Assess the Credibility ofEyewitness Testimony, January 2002

E. BURNS, K. BRUSH, In-depth guide to machine learning in the enterprise, What is Deep Learning and How Does It Work? – TechTarget

R. CALÒ, Robotics and the Lessons of Cyberlaw, in California Law Review, 2015

A. CAPPELLINI, "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale", op. cit., Paragrafo 4 (pag. 10 - 13) "La responsabilità diretta dell'IA: la tesi positiva di Gabriel Hallevy"

G. CAPUZZO, "Do Algorithms dream about Electric Sheep?" Percorsi di studio in tema di discriminazione e processi decisori algoritmici tra le due sponde dell'Atlantico" in Saggi - Focus: innovazione, diritto e tecnologia: temi per il presente e il futuro, RDM 3-2020

U. CARNEVALI, Prevenzione e risarcimento nelle direttive comunitarie sulla sicurezza dei prodotti

S. CARRER, Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin, in Giurisprudenza Penale Web, 2019, 4

M.C. CARROZZA, I Robot e noi, Il Mulino, 2017, pp. 3 ss.

V.CASADEI, Dagli automi alla moderna robotica: un'insolenza che dura 2000 anni, https://insolenzadir2d2.it/dagli-automi-alla-moderna-robotica-uninsolenza-dura-2000-anni/4369/ , Novembre 2019

C. CASTELLI, D. PIANA, Giusto processo e intelligenza artificiale, Maggioli, Santarcangelo 2019

G. CIACCI, G. BUONUOMO, Profili di informatica giuridica, Milano 2018 p.75

D. K. CITRON - F. PASQUALE, The Scored Society: Due Process For Automated Predictions, in 89 Wash. L. Rev. 1, 2014

G. CONTISSA, Information technology for the law, Torino 2017 p. 141

J. COPELAND, voce Artificial Intelligence, in S. GUTTENPLAN, A companion to the Philosophy of Mind, 1996, p. 124

K. CRAWFORD - J. SCHULTZ, Big Data And Due Process: Toward A Framework To Redress Predictive Privacy Harms, in 55 B.C. L. Rev. 93, 2014

M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, op. cit., Paragrafo 3.4 (pag. 31 - 36) "L'Intelligenza Artificiale più evoluta considerata come una persona"

A. DAMASIO, Cervelli che parlano: il dibattito su mente, coscienza ed intelligenza artificiale, Mondadori Bruno, Milano, 2003

D. DANKS, Learning, FRANKISH K., RAMSEY W. M. (a cura di), The Cambridge Handbook of Artificial Intelligence, Cambridge University Press, p. 154.

R. DAVIS- D. LENAT, Knowledge Based Systems In Artificial Intelligence, McGraw Hill International Book Co.,New York, 1982

L. EDWARDS-M. VEALE, Slave To The Algorithm? Why A 'Right to an explanation' Is Probably Not The Remedy You Are Looking For, in 16 Duke L. & Tech. Rev. 18, 2017

F.H. EASTERBROOK, "Cyberspace and the Law of the Horse", University of Chicago Legal Forum, 1996, 2017

P. FABBRI, Cos'è l'intelligenza artificiale e quali sono le sue applicazioni attuali e future, https://www.zerounoweb.it/analytics/cognitive-computing/cosa-intelligenza-artificiale/ consultato il novembre del 2019

L. FLORIDI, "Philosophy of computing and information. 5 Questions", (pag. 95) Automatic Press/VIP, 2008

L. FLORIDI, What the Near Future of Artificial Intelligence Could Be, in Philosophy & Technology, fasc. 32, 2019, pp. 3 ss. (online al link https://doi.org/10.1007/s13347-019-00345-yp)

G. FORNERO, Intelligenza artificiale e filosofia, in N. ABBAGNANO, Storia della filosofia, Torino, 1994 p. 523.

S. FRANKLIN, History, motivations, and core themes, in FRANKISH K., RAMSEY W. M. (a cura di), The Cambridge Handbook of Artificial Intelligence, Cambridge University Press, p. 19.

E. FRONZA e C. CARUSO, Ti faresti giudicare da un algoritmo? Intervista ad Antoine Garapon, Rivista trimestrale, Fascicolo 4/2018

V. FROSINI, Cibernetica: Diritto e Società, Ed. di Comunità (Collana "diritto e cultura moderna", n6), Milano 1968

G. GALUPPI, "Libero arbitrio, imputabilità, pericolosità sociale e trattamento penitenziario", Dir. famiglia, fasc.1, 2007, pag. 328; Paragrafo 2 "Cenni sulla negazione del libero arbitrio nei secoli passati"

A.GARAPON e J. LASSEGUE, Justice digitale, Presses Universitaires de France, Paris, 2018

S. GLESS, E. SILVERMAN, T. WEIGEND, "If robots cause harm, who is to blame? Self-driving cars and criminal liability", in New Criminal Law Review, 2016

L.S. GOTTFREDSON, Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography, in Intelligence, Volume 24, Issue 1 , 1997, p. 13-23

R.L. GREGORY, "intelligence", in: The Oxford Companion to the Mind, a cura di R.L. GREGORY, Oxford: Oxford University Press, 1987

G. HALLEVY, The Criminal Liability of Artificial Intelligence Entities, in "Akron intellectual property journal", 2010
G. HALLEVY, op. cit., Paragrafo 3 (pag. 177 – 194) "Three models of the criminal liability of artificial intelligence entities"

D. HEAVEN (a cura di), Macchine che pensano. La nuova era dell'intelligenza artificiale, Edizioni Dedalo, 2018.

E. HILGENDORF, Auf dem Weg zu einer Regulierung des automatisierten Fahrens: Anmerkungen zur jüngsten Reform des StVG, in KriPoZ, 2017, p. 225 ss; ID. Automatisiertes Fahren und Recht – ein Überblick, in JA, 2018

E. HILGENDORF, Automated Driving and the Law, cit., p. 179

P. HUSBANDS, Robotics, in FRANKISH K., RAMSEY W. M. (a cura di), The Cambridge Handbook of Artificial Intelligence, Cambridge University Press, 2014, p. 269 ss.

G.F. ITALIANO, Intelligenza artificiale: passato, presente, futuro, in F. PIZZETTI (a cura di), Intelligenza artificiale, protezione dei dati personali e regolazione, Giappicchelli, 2018.

J. KAPLAN, Intelligenza artificiale. Guida al futuro prossimo, Luiss University Press, II ed., 2018, pp. 81 ss., e pp. 193 ss.;

H. KELSEN,"General Theory of Law and State", (pag. 93 – 95) Harvard University Press, Cambridge, 1945

P. T. KIM, Auditing Algorithms for Discrimination, in 166 U. Pa. L. Rev. Online 189, 2017; Id., Data-Driven Discrimination At Work, in 58 Wm. & Mary L. Rev. 857, 2017;

P. KIM-S. SCOTT, Discrimination In Online Employment Recruiting, in 63 St. Louis U. L.J., 2019

V. KNAPP, L'applicabilità della cibernetica al diritto, Giulio Einaudi editore, Torino 1978.

C. KÖNIG, Gesetzgeber ebnet Weg für automatisiertes Fahren – weitgehend gelungen, in NZV, 2017

J. A. KROLL et al., Accountable Algorithms, in 165 U. Pa. L. Rev. 633, 2017

J. LARSON, S. MATTU, L. KIRCHNER e J. ANGWIN, How We Analyzed the COMPAS Recidivism Algorithm, May 23, 2016

G. LATTANZI, P. SEVERINO, A. GULLO, "Responsabilità da reato degli enti", Volume I "Diritto Sostanziale", G. Giappichelli Editore, Torino 2020

R.C. LAWLOR, What Computers Can Do: Analysis and Prediction of Judicial Decisions, in 'American Bar Association Journal', 49, 1963

C. LEROUX, R. LABRUTO et a., A Green Paper on Legal Issues in Robotics, euRobotics CA, public report, 2012, p. 10

P. LICCARDO, Introduzione al processo civile telematico, in: Rivista trimestrale di diritto e procedura civile (2000)

M. G. LOSANO, Il progetto di legge tedesco sull'auto a guida automatizzata, in Diritto dell'Informazione e dell'Informatica, 2017

M. LUCIANI, La decisione giudiziaria robotica, Rivista N°: 3/2018, data pubblicazione: 30/09/2018

V. LÜDEMANN, C. SUTTER, K. VOGELPOHL, Neue Pflichten für Fahrzeugführer beim automatisierten Fahren – eine Analyse aus rechtlicher und verkehrspsychologischer Sicht, in NZV, 2018

L. LUPARIA, G. ZICCARDI, Investigazione penale e tecnologia informatica, Giuffrè, Milano, 2007

S. LUTZ, Autonome Fahrzeuge als rechtliche Herausforderung, in NJW, 2015

M.B. MAGRO, "Decisione umana e decisione robotica un'ipotesi di responsabilità` da procreazione robotica", op. cit., Paragrafo 3 (pag. 6 - 7) "Gli agenti intelligenti agiscono autonomamente, e sono perciò agenti liberi? La tesi funzionalista"

M. B. MAGRO, "Biorobotica, robotica e diritto penale", in D. Provolo, S. Riondato, F. Yenisey (a cura di), Genetics, Robotics, Law, Punishment, Padova University Press, 2014, pp. 510 s.; Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali"

A.MANTELERO, I Big Data nel quadro della disciplina europea della tutela dei dati personali, in Il Corriere giuridico - Speciali Digitali 2018, 2018, 46 ss.;

G. MARINUCCI, E. DOLCINI, G.L. GATTA, "Manuale di diritto penale. Parte generale", Ottava edizione, Giuffrè Francis Lefebvre S.p.a. Milano – 2019; Sezione III "Il Reato", Capitolo V (pag. 207 – 225) "Analisi e sistematica del reato"

M. MAUGERI, I robot e la possibile «prognosi» delle decisioni giudiziali, Contenuto in A. CARLEO, Decisione robotica, Bologna 2019

L. MC ARTHUR, Machine Learning for Philosophers, Beneficial AI Society, Edinburgh, 2019., p. 4

J. MCCARTHY, M.L. MINSKY, N. ROCHESTER, C. E. SHANNON, A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, in AI Magazine, 2006, p. 12.

P. MELLO, Intelligenza artificiale, in http://disf.org/intelligenza-artificiale (Documentazione Interdisciplinare di Scienza & Fede), 2002

V. MORABITO, Big Data and Analytics. Strategic and Organizational Impacts, New York, 2015, 23 ss;

J. NIEVA FENOLL, Inteligencia artificial y proceso judicial, Marcial Pons Ediciones Juridicas y Sociales, Madrid 2018

E. NISSAN, Digital technologies and artificial intelligence's presentand foreseeable impact on lawyering, judging, policing and law enforcement, Londra, 14 Ottobre 2015

J. NOSTA, L'AI può leggere nel pensiero. E tradurlo in immagini e parole, articolo scritto per Fortune.com, Maggio 2018

M. PALADINI, I rimedi al difetto di conformità nella vendita di beni di consumo, in E. TOSI, La tutela dei consumatori in internet e nel commercio elettronico, 2012

E. PALMERINI, "Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea", Responsabilità Civile e Previdenza, fasc.6, 2016, pag. 1815B; Paragrafo 2 "Il robot sulla scena giuridica"

E. PALMERINI et a., Guidelines on regulating robotics, The RoboLaw Project, 2014, p. 15

B. PANATTONI, "Intelligenza artificiale: le sfide per il diritto penale nel passaggio dall'automazione tecnologica all'autonomia artificiale", Diritto dell'Informazione e dell'Informatica (Il), fasc.2, 1 aprile 2021, pag. 317, Paragrafo 2 "Dall'automazione tecnologica all'autonomia artificiale"

G. PASCUZZI, Il diritto nell'era digitale. Tecnologie informatiche e regole privatistiche, Bologna, 2002

F. PASQUALE, The black Box Society: The Secret algorithms that Control Money And Information, Cambridge, 2015

S. PIACENTI, AI: è possibile predire la capacità intellettiva?, 25 Settembre 2018. https://www.ingenium-magazine.it/ai-possibile-predire-capacita-intellettiva/

C. PIERGALLINI, "Intelligenza artificiale: da 'mezzo' ad 'autore' del reato?" Rivista Italiana di Diritto e Procedura Penale, fasc.4, 1 dicembre 2020, pag. 1745; Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'"

G. PITRUZZELLA, Big Data, Competition and Privacy: A Look from the Antitrust Perspective, in Concorrenza e Mercato, 2016,

F. PIZZETTI, Privacy e diritto europeo nella protezione dei dati personali, Torino, 2016

PONZANELLI, "Responsabilità per danno da computer: alcune considerazioni comparative", in Resp. Civ. prev., 1991, 650

N. POSTERARO, Una risoluzione del Parlamento europeo sull'uso dell'intelligenza artificiale nel diritto penale, in Istituto di Ricerche sulla pubblica amministrazione, 23 novembre 2021, https://www.irpa.eu/una-risoluzione-del-parlamento-europeo-sulluso-dellintelligenza-artificiale-nel-diritto-penale/

D. PROUDFOOT, J. COPELAND, Artificial Intelligence, in MARGOLIS E., SAMUELS R., STICH S.P. (a cura di), The Oxford Handbook of Philosophy of Cognitive Science, Oxford, 2012, p. 166 ss.

S. QUINTARELLI, op. cit., Capitolo 6.2 (pag. 110 – 113) "Personalità giuridica. La convivenza tra uomini e software"

G. RESTA, Diritti esclusivi e nuovi beni immateriali, Torino, 2011

G. RESTA-V. Z. ZENCOVICH (a cura di), La protezione transnazionale dei dati personali, Roma, 2016

S. RODOTA', Elaboratori elettronici e controllo sociale, Bologna, 1973;

S. J. RUSSEL, P. NORVIG, Artificial Intelligence – A modern approach, Third Edition, 2010

J. SALLANTIN, J. JAQUES. Szczeciniarz, Il concetto di prova alla luce dell'intelligenza artificiale, Giuffrè editore (Collana "Epistemologia giudiziaria"), Milano, 2005

I. SALVADORI, "Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale", Rivista Italiana di Diritto e Procedura Penale, fasc.1, 1 marzo 2021, pag. 83; Paragrafo 9, "Considerazioni finali"

A. SANTOSUOSSO, "Diritti, storicità, artificialità – Creature di Dio e artefatti giuridici"

A.SANTOSUOSSO, C. BOSCARATO, F. CAROLEO, Robot e diritto: una prima ricognizione, in Nuova Giurisprudenza Civile Commentata, 2012

A. SANTOSUOSSO, B. BOTTALICO, Autonomous Systems and the Law: Why Intelligence Matters. A European Perspective, in HILGENDORF E., SEIDEL U., Robotics, Autonomics and the Law, 2017

J.R. SEARLE, Minds, Brains and Programs, in The Behavioural and Brain Science, 1980 p.417
J.R. SEARLE, La mente è un programma?, in Le Scienze, Marzo 1990 n. 259

A.D. SELBSDT-S. BAROCAS, The Intuitive Appeal Of Explainable Machines, in 87 Fordham L. Rev. 2018
P. SEVERINO, "Intelligenza artificiale e diritto penale", Paragrafo 2 (pag. 533 – 536) "La responsabilità penale degli agenti intelligenti" in RUFFOLO U., "Intelligenza artificiale – Il diritto, i diritti, l'etica", Collana "Tech and Law" di Giuffrè Francis Lefebvre Edizione, Milano 2020; Parte IV "Diritto processuale, penale, costituzionale, amministrativo, tributario, eurounitario", Sezione V "A.I. e diritto penale" Capitolo 5 (pag. 531 – 547

B. SCHAFER, ‘Closing Pandora’s box? The EU proposal on the regulation of robots’, in The journal of the Justice and the Law Society of the University of Queeensland, 2016, vol. 19

P. STEINERT, Automatisiertes Fahren (Strafrechtliche Fragen), in SVR, 2019

F. STELLA, Giustizia e modernità. La protezione dell’innocente e la tutela delle vittime, Giuffré, 2003

E. STRADELLA, La regolazione della Robotica e dell’Intelligenza artificiale: il dibattito, le proposte, le prospettive

C. A. SULLIVAN, Employing, 2018 (Seton Hall Public Law Research Paper)

A.TURING, Computing Machinery and Intelligence, in Mind, 1950.

S. VALLOR, G. A. BEKEY, Artificial Intelligence and the Ethics of Self-Learning Robots, in LIN P.

M. WEBER, Economia e società, vol. III, Edizioni di Comunità, Milano, 1980

N. WIENER, The Human Use of the Human Beings, trad. it. di D. Persiani con il titolo Introduzione alla cibernetica, Boringhieri, Torino 1953.

H. WITTEN, E. FRANK, M.A. HALL, Data mining. Practical Machin Learning Tools and Techniques, Elsevier, Burligton, 2011, p. 7

L. WÖRNER, Der Weichensteller 4.0 Zur strafrechtlichen Verantwortlichkeit des Programmierers im Notstand für Vorgaben an autonome Fahrzeuge, in ZIS, 2019