LUISS

Dipartimento
di Giurisprudenza

Cattedra di Diritto Penale 2

# Responsabilità penale per i reati commessi dall'Intelligenza Artificiale: *"Machina delinquere non potest"*

Prof. Enrico Gallucci
RELATORE

Prof. Rocco Blaiotta
CORRELATORE

Costanza Corridori
CANDIDATO

Anno Accademico 2020/2021

# Responsabilità penale per i reati commessi dall'Intelligenza Artificiale: *"Machina delinquere non potest"**

* Per il titolo si è debitori di A. CAPPELLINI, "*Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale",* in Criminalia: Annuario di scienze penalistiche, Edizioni ETS, 2018, Pisa: https://discrimen.it/wp-content/uploads/Cappellini-Machina-delinquere-non-potest.pdf

# ***INDICE***

# ***INTRODUZIONE***

"*Le macchine sono in grado di pensare?*"

Si tratta di una delle frasi più celebri di Alan Turing: il padre dell'informatica.

Siamo giunti in quella che viene comunemente denominata la quarta rivoluzione industriale, dominata da *robot* e macchine pensanti. Tali apparecchi informatici ed elettronici ci supportano e aiutano nella quotidianità, dagli assistenti vocali come *Alexa* o *Siri*, agli elettrodomestici che comunicano tra di loro, passando per le vetture dotate di guida autonoma.

Se di punto in bianco tutti i sistemi dotati della c.d. intelligenza artificiale dovessero sparire, la società rimarrebbe paralizzata. Molte attività, prima tipicamente manuali, sono state ora interamente automatizzate. Le macchine hanno sostituito gli esseri umani nel mondo del lavoro ma, al contempo, la tecnologia ha generato una richiesta continua di personale altamente specializzato che si possa occupare di realizzare ed educare l'algoritmo di *machine learning* contenuto all'interno di un *software* di intelligenza artificiale.

Nonostante la quantità di tecnicismi legati all'automazione, di cui si tratterà all'inizio della tesi, l'elaborato mira a comprendere il rapporto tra gli illeciti e l'intelligenza artificiale. Dunque sarà importante comprendere i meccanismi sottostanti a tale tecnologia per trattare dei reati e della responsabilità penale ad essa connessi.

Di recente capita sempre più spesso di assistere a conferenze di giuristi in cui si sente parlare di algoritmi, di digitalizzazione, di *software* e di automatizzazione, in cui si tratta di giudici *robot*, della integrale sostituzione degli avvocati con dei sistemi per il *computer*, ovvero di *cybercrime*, di *chatbot* e di *deep-fake*. Il mondo dell'informatica sta entrando nelle varie branche del diritto e non solo: in ambito sanitario si possono osservare *robot* che operano i pazienti sotto la supervisione del chirurgo e algoritmi che possono suggerire al medico la migliore diagnosi e cura per il singolo soggetto.

Sembra la descrizione di un mondo idilliaco in cui macchine intelligenti e esseri umani possono convivere supportandosi a vicenda.

Ma cosa succede se un veicolo autonomo dotato di intelligenza artificiale investe un pedone? Chi risponde se il *robot* chirurgo sbaglia l'approccio medico? E se una *chatbot* inizia a diffamare un utente su una piattaforma *social*?

Non stiamo parlando di film fantascientifici, ma di realtà attuali che presto il legislatore dovrà affrontare al fine di adottare una disciplina di diritto penale che sia idonea a tutelare i vari interessi in gioco.

In questa tesi tratteremo l'argomento della responsabilità penale per i reati commessi dall'intelligenza artificiale seguendo un approccio che va dal generale al particolare. Affronteremo infatti, inizialmente il concetto stesso di "intelligenza" in relazione alle macchine al fine di studiare il loro funzionamento e le loro possibili applicazioni (senza pretesa di analizzare tecnicamente ogni passaggio di un algoritmo, si è provato a

trattare i punti e i concetti principali con cui il giurista del futuro dovrà convivere per comprendere le modalità di applicazione del diritto a queste nuove entità). Successivamente la trattazione affronta la carenza di una normativa idonea a disciplinare tali fattispecie sia a livello nazionale che sovrannazionale ed europeo (l'unica parvenza di testo vincolante risulta essere una Proposta di Regolamento del Parlamento Europeo del 21 aprile 2021).

Successivamente si entrerà nel vivo del diritto penale, trattando dell'applicazione dei sistemi di intelligenza artificiale nell'ambito delle indagini e del processo penale, le modalità di assistenza alle decisioni giudiziarie e la possibilità di considerare il *robot* come autore, vittima o strumento del reato.

Nel secondo capitolo si tratterà della questione principale concerne il brocardo "*machina delinquere non potest*" (formulato sulla falsa riga di quello prima applicabile alle società). Può un sistema intelligente essere considerato responsabile per gli illeciti che realizza?

Di particolare interesse, sul punto, le teorie di Hallevy, connesse all'ipotesi di istituire una personalità elettronica per i *robot* da affiancare alle comuni categorie dei soggetti di diritto (persona fisica e giuridica) al fine di valutare la sussistenza di margini per configurare una responsabilità penale in capo ai sistemi di intelligenza artificiale. Si affrontano, poi, le critiche a tali teorie essenzialmente fondate sulla impraticabilità dell'analogia tra la mente umana e quella robotica, sul concetto di libero arbitrio e sull'effettiva coscienza e volontà di movimento posseduta da un sistema intelligente.

Per quanto l'algoritmo si basi sul *machine learning* (che permette al sistema di apprendere dall'ambiente e di modificare il proprio comportamento esteriore), allo stato attuale della tecnica, i *robot* sono fortemente vincolati all'algoritmo che li domina, non hanno possibilità di decidere in modo pienamente autonomo e cosciente le azioni da intraprendere. Risultano dunque entità determinate, ma vista la rapidità di evoluzione della tecnologia, non è possibile escludere che in pochi anni l'abilità delle macchine e la loro autonomia superi o equivalga quella dell'essere umano. Di conseguenza è necessario porre fin da subito questi interrogativi.

Assumendo che, per ora, i *robot* intelligenti non possano essere considerati come titolari di diritti e di doveri e dunque non possano essere soggetti al diritto penale, chi risponderà per i reati da loro commessi?

Ipotizzare una responsabilità in capo al programmatore o utilizzatore, se si procede sulla scorta di automatismi, sconta il rischio di violare il principio di personalità e colpevolezza della responsabilità penale, come sancito dall'articolo 27 della Costituzione, ricadendo in inammissibili ipotesi di responsabilità oggettiva che il diritto penale non ammette. Dunque, risulta necessario analizzare le azioni del *robot*, i cui meccanismi sono spesso oscuri (come una *black box*), per valutare se le persone fisiche che lavorano e agiscono intorno, tramite e attraverso i sistemi intelligenti, possano nel caso concreto prevedere a priori la realizzazione dell'evento lesivo.

Nel terzo capitolo, infine, in applicazione degli attuali paradigmi di imputazione penale, sono affrontate le principali fattispecie criminose che l'intelligenza artificiale può commentare seguendo un approccio casistico attinto dalla cronaca e da dottrina specializzata.

Quando si sente parlare di tali sistemi, sembra sempre che riguardino un mondo lontano, futuro e inaccessibile. Invece, non è così. Il futuro è adesso ed è bene che si inizi ad affrontarlo. Il mondo dei reati robotici è già reale. Il diritto (come sempre) arriva in ritardo (e il penale in particolar modo): è normale, esso disciplina le situazioni concrete che di volta in volta si vanno a realizzare e non può ipotizzare situazioni future. Il legislatore rincorre i fenomeni che avvengono nella quotidianità e la funzione stessa del diritto penale impone che sia così, non essendo ammissibile in una democrazia occidentale avanzata l'utilizzo del sistema penale per "correggere" e "indirizzare" le condotte dei consociati verso scopi superindividuali. Il diritto penale, e quello punitivo latamente inteso, deve rimanere l'*ultima ratio*.

Infatti, per quanto possiamo girare la testa dall'altra parte e non pensare alle conseguenze (positive e negative) dell'evoluzione tecnologica, queste ci coinvolgono da vicino.

Vengono, dunque, analizzate tre tipologie di reati "robotici" (sarebbe stato interessante trattarne anche altri, ma purtroppo in una tesi di laurea non risultava possibile): l'omicidio e le lesioni colpose stradali, la responsabilità medica e il *cybercrime* in senso ampio.

Infatti, con i veicoli autonomi i dubbi relativi alla responsabilità per i danni arrecati da un'autovettura non guidata da un essere umano, risultano notevoli: la possibilità di un controllo *ex ante* ed uno realizzato dal "guidatore-passeggero" persona fisica, l'ipotesi di realizzazione di un algoritmo del rischio che scelga lui discrezionalmente quale bene giuridico tutelare maggiormente e il tema dell'affidabilità dei sistemi intelligenti (connesso alla fiducia che la collettività può dargli).

In ambito medico e sanitario, le scoperte tecnologiche permettono ai pazienti di riprendere la mobilità di arti paralizzati, di controllare il tremore del *Parkinson* e di mantenere sotto controllo i parametri vitali del soggetto al fine di somministrargli correttamente i farmaci. Inoltre, gli algoritmi intelligenti supportano il medico nelle sue decisioni e diagnosi. Ma in caso di errore? Le vigenti fattispecie criminose risultano idonee ad affrontare anche tali situazioni?

Da ultimo, il tema dei reati informatici: la diffamazione e le *fake-news* rischiano di essere automatizzate tramite *chatbot* intelligenti che comunicano con gli utenti sui *social network* e imparano dai loro atteggiamenti. Inoltre, un fenomeno alquanto preoccupante riguarda il *deep-fake*, l'abilità di alcuni sistemi intelligenti di realizzare video falsi raffiguranti persone vere: possono consistere in falsi discorsi politici o in falsi video pornografici andando così a "spogliare" virtualmente un soggetto non consenziente. È abbastanza evidente l'impatto che questo potrebbe avere sull'informazione e sulla vita personale delle vittime: una nuova frontiera del *revenge porn*.

Gli interrogativi sono molti e questa tesi non ha la presunzione di rispondere in modo esaustivo e categorico a tali problematiche. Lo scopo principale è far riflettere il lettore, mettere in dubbio le sue certezze ponendolo difronte al futuro per vedere come reagirà. È solo tramite le domande che sarà possibile, nel prossimo futuro, costruire un assetto normativo tale da poter tutelare concretamente tutti i beni giuridici in gioco. La tesi ha, quindi, l'obbiettivo di evidenziare le difficoltà del momento e i possibili sviluppi futuri di una disciplina su cui dovrà mettere mano il legislatore nazionale ed europeo.

L’intelligenza artificiale non deve preoccupare l’umanità: si tratta di un progresso tecnologico fondamentale e necessario, ma dobbiamo essere pronti per affrontarlo al meglio. Anche e soprattutto in ambito penale.

# ***CAPITOLO 1:***

# ***L'INTELLIGENZA ARTIFICIALE E IL DIRITTO PENALE***

---

**SOMMARIO**: 1. Definizione scientifica – 1.1 Storia dell'Intelligenza Artificiale: da Alan Turing a Elon Musk – 1.2 L'algoritmo alla base dell'Intelligenza artificiale – 1.3 Tecniche di apprendimento: *machine learning* e *deep learning*. Il *Black box algorithm* – 1.4 Ambiti di applicazione dell'Intelligenza artificiale – 2. Definizione normativa – 2.1 Quadro normativo Europeo – 2.2 Assenza di una normativa Nazionale – 2.3 Approccio all'estero – 2.4 Tutela dei diritti fondamentali – 3. Intelligenza Artificiale e Sistema Penale – 3.1 Polizia predittiva – 3.2 Digitalizzazione della giustizia e giudice *robot* – 3.3 Giustizia predittiva: Valutazione della pericolosità – 3.4 Intelligenza artificiale come strumento, vittima o autore del reato

---

## ***1. Definizione scientifica***

Che cos'è l'intelligenza artificiale? La realtà attuale è piena di dispositivi (ad esempio le applicazioni degli *smartphone*, i veicoli che si guidano da soli e i *robot* domestici) in grado di imparare autonomamente e di evolversi grazie all'innovazione digitale con algoritmi di apprendimento sempre più sofisticati. Allo stesso tempo, fornirne una definizione universalmente accettata non è poi così semplice proprio a causa del fatto che alla base dell'intelligenza artificiale esiste il concetto di "intelligenza" che di per sé è fortemente complesso da inquadrare. L'intelligenza artificiale può intervenire in qualsiasi attività senza la nostra consapevolezza. Siamo giunti in quella che viene definita come "Quarta Rivoluzione Industriale".[1]

Qual è la differenza tra l'intelligenza umana e quella animale? Le macchine possono pensare?

Filmografia e letteratura hanno cercato di definire e di affrontare il tema dell'automazione, spesso e volentieri in circostanze inquietanti e fantascientifiche in cui la macchina prende il sopravvento sull'uomo e inizia a dominarlo: una paura comprensibile ma completamente irrazionale se si comprende il fatto che la macchina è progettata, costruita e utilizzata dall'uomo, in funzione dell'uomo stesso.

[1] M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, *"AI: profili tecnologici Automazione e Autonomia: dalla definizione alle possibili applicazioni dell'Intelligenza Artificiale"* in Bio Law Journal – Rivista di Bio Diritto, n. 3/2019; Paragrafo 1 (pag. 2) "Introduzione".

Film come "*2001 Odissea nello spazio*"[2], "*Matrix*"[3], "*Her*"[4] e molti altri sono fondamentali per poterci rendere conto di come la robotica e soprattutto l'intelligenza artificiale siano entrate nelle nostre vite ancora prima di una sua piena evoluzione e definizione.

L'opera fantascientifica di Čapek "R.U.R."[5] può essere considerata la madre della parola "*robot*": fu in tale contesto teatrale che, la stessa, venne utilizzata per la prima volta.[6]

Asimov, scrittore e biochimico russo, considerato il padre della fantascienza, in "*Runaround*" del 1942, contenuto nella raccolta di racconti pubblicata nel 1950 "*I, Robot*"[7], formula le tre leggi della robotica:[8]

1. *"Un robot non può recar danno a un essere umano né può permettere che, a causa del suo mancato intervento, un essere umano riceva danno.*
2. *Un robot deve obbedire agli ordini impartiti dagli esseri umani, purché tali ordini non vadano in contrasto alla Prima Legge.*
3. *Un robot deve proteggere la propria esistenza, purché la salvaguardia di essa non contrasti con la Prima o con la Seconda Legge."*

Aggiungendo successivamente la legge zero: "*Un robot non può recare danno all'umanità, né può permettere che, a causa del proprio mancato intervento, l'umanità riceva danno."*

Secondo Stephen Hawking, nell'arco dei prossimi cento anni, l'intelligenza dei *computer* supererà quella umana.[9]

Con tali ragionamenti si rischia di sfociare in ambiti filosofici e sociologici che porterebbero l'elaborato a percorrere una via completamente diversa anche se estremamente interessante. Ritornando ad una spiegazione scientifica di cosa sia l'intelligenza artificiale, bisogna partire dal presupposto che non esiste una definizione univoca della stessa. Secondo alcuni può essere genericamente considerata come settore dell'informatica con ad oggetto la teoria, le tecniche e le metodologie che permettono di progettare sistemi di *hardware* e *software* in grado di elaborare prestazioni assimilabili all'intelligenza umana. [10]

La macchina non è intelligente ma lo sembra perché emula il comportamento e il ragionamento dell'uomo. L'intelligenza umana non ha eguali per versatilità, capacità di ragionare, di raggiungere obiettivi, e di comprendere il linguaggio. Lo scopo è, quindi, quello di permettere alla macchina di compiere operazioni

---

[2] S. KUBRICK, "*2001: Odissea nello spazio*", 1968.
[3] A. E L. WACHOWSKI, "*Matrix*", serie di tre film, 1999 – 2003.
[4] S. JONZE, "*Her*", 2014.
[5] K. ČAPEK, "*R.U.R.*", 1920.
[6] A. LONGO E G. SCORZA, "*Intelligenza Artificiale: impatto sulle nostre vite, diritti e libertà",* Mondadori Education Università, Firenze, 2020; Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[7] I. ASIMOV, "*I, Robot*", Mondadori, 1950.
[8] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[9] Intervento di S. Hawking durante la Conferenza *Zeitgeist*, Londra, maggio 2015 (citazione riportata da Redazione, "*Do You Trust This Computer?"* in F. BASILE, "*Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine",* in Diritto Penale Contemporaneo online, Milano, 2019: https://archiviodpc.dirittopenaleuomo.org/upload/3089-basile2019.pdf , 15 maggio 2019.
[10] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, "*Intelligenza artificiale",* Istituto della Enciclopedia Italiana, Treccani, Enciclopedia della Scienza e della Tecnica, Roma, 2008: http://www.treccani.it/enciclopedia/intelligenza-artificiale_(Enciclopedia-della-Scienza-e-della-Tecnica)/; "Introduzione" (pag. 1).

fondate su conoscenze e su criteri di elaborazione dei dati tipici del ragionamento umano, fornendo prestazioni qualitativamente equivalenti e quantitativamente superiori a quelle dell'essere vivente.[11] Le macchine fino ad oggi realizzate non possono essere considerate intelligenti: i loro meccanismi cognitivi sono poveri rispetto agli umani e anche se il loro funzionamento gli consente di accumulare esperienza, questo non significa comprendere. Inoltre, non possono generare processi creativi o costruire una rappresentazione del mondo. Si preferisce, dunque, parlare di "razionalità" piuttosto che di "intelligenza" al fine di eliminare i riferimenti antropomorfici.[12]

Nils J. Nilsson, informatico statunitense, definisce l'intelligenza artificiale come "*quell'attività dedicata a rendere intelligenti le macchine, e l'intelligenza è quella qualità che consente a un'entità di funzionare in modo appropriato e con lungimiranza nel suo ambiente*."[13] Inoltre, la considera al tempo stesso come scienza e ingegneria: scienza per quanto riguarda lo studio scientifico dell'intelligenza dell'uomo e le modalità per emularla; ingegneria per lo studio legato alla realizzazione di una macchina basata sull'intelligenza artificiale che possa offrire un miglioramento nella vita dell'uomo.[14]

Nel 1968, Marvin Minsky, scienziato e matematico statunitense specializzato nel campo dell'intelligenza artificiale, la definisce come "*la scienza di far fare alle macchine cose che richiederebbero intelligenza se fatte dall'uomo*"[15]. Nel regno dell'intelligenza artificiale rientra quindi ogni tipo di comportamento intelligente: dal gioco degli scacchi alla comprensione dei racconti, dalle scoperte matematiche alle diagnosi mediche.[16]

È inoltre possibile distinguere una definizione scientifica da una filosofica, di intelligenza artificiale: secondo gli scienziati, la stessa è lo studio degli agenti che ricevono percezioni dall'ambiente ed eseguono azioni. Lo scopo è realizzare diverse tecniche per permetterlo. Secondo la filosofia, invece, l'intelligenza artificiale è il campo dedicato alla costruzione di animali/persone artificiali che appaiono come animali o come persone. La definizione scientifica è inclusa in questa.[17]

In linea generale, secondo l'esperto americano Roger Schank, è possibile riconoscere una macchina basata sull'intelligenza artificiale considerando cinque requisiti fondamentali: la comunicazione, più sarà facile comunicare con un'entità, più l'entità sembrerà intelligente; la conoscenza interna di sé da parte

---

[11] C. PARODI E V. SELLAROLI, "*Sistema penale e intelligenza artificiale: molte speranze e qualche equivoco*", in Diritto Penale Contemporaneo, Fascicolo 6/2019: https://archiviodpc.dirittopenaleuomo.org/pdf-viewer/?file=%2Fpdf-fascicoli%2FDPC_6_2019.pdf#page=47; Paragrafo 2 (pag. 49 – 51) "Intelligenza artificiale: di cosa stiamo parlando?".

[12] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit.

[13] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, "*Artificial intelligence and life in 2030*", Report of the 2015 Study Panel, Standford University, 2016: https://apo.org.au/sites/default/files/resource-files/2016-09/apo-nid210721.pdf; Section 1 (pag. 12 – 17) "What is artificial intelligence?".

[14] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., "Introduzione" (pag. 1).

[15] M.L. MINSKY, "*Semantic information processing*", Cambridge, 1969.

[16] E. L. RISSLAND, "*Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning*", The Yale Law Journal, Vol. 99 no. 8, The Yale Law Journal Company, Inc., 1990, pp. 1957–81: https://doi.org/10.2307/796679; Paragrafo 1A, "Background on Atrificial Intelligence – What is Artificial Intelligence?".

[17] A. SANTOSUOSSO, "*Intelligenza artificiale e diritto: perché le tecnologie di IA sono una grande opportunità per il diritto*", Scienza e Filosofia (collana diretta da A. MASSARENTI), Mondadori Education Università, Firenze, 2020; Capitolo 1 (pag. 16 – 31) "Umano e artificiale".

dell'entità; la conoscenza esterna del mondo; il comportamento guidato da un obiettivo, le azioni dell'entità intelligente devono essere effettuate per raggiungere i suoi obiettivi; la creatività, intesa come capacità di trovare strade alternative quando l'azione iniziale fallisce.[18] Ma alcuni tipi di macchine dotate di intelligenza artificiale possiedono caratteristiche maggiori, potendo così agire secondo metodi più sofisticati.[19]

È possibile, inoltre, distinguere il concetto di "reale" in due componenti: "naturale" e "artificiale". Il primo si divide in "uomo" e "mondo esterno all'uomo"; il secondo è collegato alla nozione di macchina (elaboratore per l'informatica e *robot* per la robotica). Ed è proprio nella relazione tra queste tre componenti, "uomo", "mondo" e "macchina", che dobbiamo focalizzare il nostro studio. [20]

È possibile considerare come data di nascita dell'espressione "intelligenza artificiale" il convegno tenutosi nell'estate del 1955 al *Dartmounth College* di Hanover, New Hamphsire, dove John McCarthy, considerato il padre fondatore dell'intelligenza artificiale, come assistente universitario di matematica, organizzò un evento, insieme ad altri suoi colleghi (Marvin Minsky, Nathaniel Rochester e Claude Shannon), sull'intelligenza artificiale descrivendola così: "*lo studio procederà sulla base della congettura che tutti gli aspetti dell'apprendimento o qualsiasi altra caratteristica dell'intelligenza possa essere di principio descritta in modo così preciso che una macchina la possa simulare. Si tenterà di scoprire come si possa fare in modo che le macchine usino il linguaggio, formulino astrazione e concetti, risolvano tipi di problemi ora riservati agli esseri umani, e migliorino sé stesse*".[21]

### *1.1 Storia dell'Intelligenza Artificiale: da Alan Turing a Elon Musk*

Sebbene il convegno di Dartmouth sia stato il punto di svolta ufficiale per lo studio sull'intelligenza artificiale, molte delle idee tecniche, riguardo la stessa, esistevano da tempo.

Procediamo con ordine.

Le prime modalità di calcolo meccanico nascevano con l'abaco nel 400 a.C. e nel 1642 Pascal inventò la Pascalina, una macchina per aiutare il padre, esattore delle tasse, ad effettuare addizioni e sottrazioni sfruttando un sistema di ruote.[22] Tali strumenti, però, avevano il difetto di permettere lo svolgimento di una

---

[18] R.C. SCHANK, "*What's IA, Anyway?*", in *IA Magazine*, Winter 8(4), 1987, pp. 59 ss.

[19] 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: https://ideaexchange.uakron.edu/akronintellectualproperty/vol4/iss2/1; Paragrafo 2 (pag. 175 - 177) "What is an Artificial Intelligence Entity? – Machina sapiens – A "Thinking" Machine or a Thinking Machine".

[20] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 2 (pag. 4 – 5) "Intelligenza Artificiale e l'informatica in una struttura culturale unitaria".

[21] J. MCCARTHY, M. L. MINSKY, N. ROCHESTER E C.E. SHANNON, "*A proposal for the Dartmouth summer research project on artificial intelligence*", Dartmouth College, Hanover, New Hamphsire, 1955: http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf .

[22] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".

sola operazione alla volta e di non poter essere programmati in anticipo tramite indicazioni, e quindi, di non poter agire autonomamente.[23]

Nel XIX secolo Charles Babbage progettò la prima macchina analogica programmabile composta da ruote a dieci denti che rappresentavano i numeri in sistema decimale. In base alla posizione di tali ruote metalliche, impilate in delle colonne, era possibile compiere dei calcoli evitando la commissione dei tipici errori umani. [24] Tali dispositivi permettevano di elaborare algoritmi scritti su delle schede perforate che incorporavano istruzioni eseguibili dalla macchina. L'idea delle schede perforate nasceva molti anni prima, all'inizio del XIX secolo, con il "Telaio di Jacquard": lo stesso le utilizzava per automatizzare la produzione dei prodotti tessili codificando i ricami e i decori da realizzare.[25]

I progetti di Babbage furono la "macchina alle differenze" e la "macchina analitica". Con la prima, si potevano rappresentare fino a venti numeri di diciotto cifre ciascuno. La seconda, invece, serviva per lo *step* successivo utilizzando i numeri risultanti dalla macchina alle differenze come dati di ingresso per il calcolo susseguente. La macchina analitica, inoltre, avrebbe dovuto possedere una memoria contenente i dati risultanti delle operazioni e avrebbe dovuto avere un sistema di controllo, in grado di eseguire automaticamente una sequenza di operazioni secondo il meccanismo delle schede perforate utilizzato nei telai.[26] Purtroppo, a causa dell'elevato costo di tali apparecchiature, le stesse furono realizzate a partire dal Ventesimo secolo.

È solo negli anni Venti del XX secolo che nacque la macchina computazionale (*computer machine*), uno strumento in grado di svolgere il lavoro di un umano, effettuando calcoli tramite metodi efficaci e rapidi.[27]

All'inizio degli anni Quaranta con il termine "cibernetica" si indicava lo studio dei processi riguardanti la comunicazione e il controllo sia negli animali sia nelle macchine.

Durante la Seconda guerra mondiale è possibile notare un'evoluzione nella tecnologia grazie agli studi realizzati da scienziati al fine di contrastare il nemico sfruttando soprattutto l'intercettazione delle trasmissioni radio. Nacque in tale periodo lo sviluppo di tecniche che porteranno ai futuri "*Wi-Fi*" e "*GPS*".[28]

Tra gli anni Cinquanta e Settanta la ricerca sull'intelligenza artificiale prese piede e grazie alle idee di Charles Babbage, nacquero i primi *computer* elettronici e digitali: *robot* primitivi in grado di svolgere autonomamente delle sequenze di istruzioni.[29] La *computer machine* divenne esclusivamente *computer*.[30]

L'intuizione di Donald Hebb (1949) legata al fatto che i neuroni comunicano inviandosi scariche elettriche e che esse rappresentano l'attività di base dell'apprendimento e della memoria, permise di

---

[23] A. SANTOSUOSSO, op. cit., Capitolo 1 (pag. 16 – 31) "Umano e artificiale".
[24] *Idem.*
[25] S. QUINTARELLI, "*Intelligenza artificiale: cos'è davvero, come funziona, che effetti avrà",* Bollati Boringhieri, Torino, 2020; Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".
[26] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 1: pag. 2 - 4 "Storia dell'intelligenza Artificiale".
[27] A. SANTOSUOSSO, op. cit., Capitolo 1 (pag. 16 – 31) "Umano e artificiale".
[28] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".
[29] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Appendix 1 (pag. 50 – 52) "A short History of AI".
[30] A. SANTOSUOSSO, op. cit., Capitolo 1 (pag. 16 – 31) "Umano e artificiale".

paragonare il cervello ad un sofisticato *computer* dando la possibilità di replicare l'intelligenza umana tramite un cervello artificiale.[31]

Su queste basi, Frank Rosenblatt, scienziato impegnato in un progetto finanziato dalla Marina statunitense presso il Laboratorio aeronautico della Cornell *University*, sviluppò il "percettrone", modello computazionale ispirato ai neuroni biologici. [32] Vennero ideate delle reti, con lo scopo risolvere problemi complessi grazie al collegamento degli *input* e *output*, realizzate grazie ai neuroni artificiali.[33]

Nello stesso periodo, Warren McCulloch e Walter Pitts proponevano il primo modello di neuroni artificiali con lo scopo di studiare i meccanismi dell'autoregolazione e del controllo presenti sia negli organismi viventi sia nelle macchine che possiedano la capacità di modificare il proprio comportamento in base alle sollecitazioni dategli dall'ambiente.[34]

Ma è solo nel 1952 che apparve per la prima volta il termine "*Machine Learning*" (apprendimento della macchina) grazie ad Arthur Samuel, informatico dell'IBM che pubblicò un articolo dal titolo "Alcuni studi sul *Machine Learning* usando il gioco della dama".[35] Il *machine learning* permette ai *computer* di imparare senza una programmazione capillare sulle modalità di apprendimento, alimentando un algoritmo con dei dati, in modo che la macchina impari a eseguire una certa operazione automaticamente.[36] Lo scienziato informatico sviluppò un programma di gioco della dama con la capacità di automigliorarsi apprendendo nuove tecniche automaticamente.[37] Una funzione matematica determinava il punteggio in base alla posizione delle pedine sulla scacchiera e emetteva una stima sulla probabilità di vincita calcolata su ogni singola mossa. Il sistema imparava memorizzando il valore della funzione per ogni posizione.[38]

Giungiamo ora a colui che possiamo chiamare il padre del *computer* e dell'intelligenza artificiale grazie alle sue idee influenti alla base dell'informatica: Alan Turing, che ha immaginato e progettato un *computer* con la possibilità di simulare l'intelligenza umana. Purtroppo non riuscì mai a realizzare le sue idee per mancanza delle risorse necessarie per poterle portare a termine.[39]

Nel 1950, con la pubblicazione di "*Computing Machinery and Intelligence*", Turing ipotizzò che le machine potessero pensare e realizzò un *test* per appurarlo: "*The imitation game*" (il gioco dell'imitazione). Il *test* misurava l'intelligenza artificiale della macchina in base a quanto fosse facile distinguerla da quella umana: era necessario fornire ad una persona una tastiera per scrivere delle domande e all'altro capo del

---

[31] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 1: pag. 2 - 4 "Storia dell'intelligenza Artificiale".
[32] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Appendix 1 (pag. 50 – 52) "A short History of AI".
[33] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".
[34] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 1: pag. 2 - 4 "Storia dell'intelligenza Artificiale".
[35] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[36] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".
[37] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Appendix 1 (pag. 50 – 52) "A short History of AI".
[38] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".
[39] A. M. TURING, "*Computing Machinery and Intelligence*," *Mind* 59, no. 236, 1950 (pag. 433–460).

terminale bisognava porre una macchina e un essere umano. Se il soggetto che poneva le domande riscontrava difficoltà a comprendere la provenienza del messaggio di risposta (se dalla macchina o dalla persona), allora il test poteva essere considerato superato.[40] Secondo Alan Turing, le macchine non pensano ma imitano il pensiero umano.[41]

Ma ancora prima, nel 1936, di maggiore rilevanza per gli studi informatici, è, invece, la "Macchina di Turing"; un modello astratto di *computer* composto da una striscia infinita di carta che l'elaboratore utilizzava per scrivere e leggere simboli al fine di procedere con le istruzioni impartitegli.[42] Il concetto di algoritmo può essere ricondotto alla sequenza di operazioni svolte dalla macchina.[43]

Negli anni Sessanta i *robot* iniziavano ad essere utilizzati nelle industrie. Il primo a lavorare in una catena di montaggio, trasportando pezzi e scaldando parti delle automobili, fu "*Unimate*", nel 1961, progettato da George Devol al fine di sostituire l'uomo nelle mansioni più pericolose.

Un grande progresso nel campo dell'intelligenza artificiale fu "*Eliza*", di Joseph Weizenbaum del 1965. Possiamo definirlo il prototipo di quello che oggi viene chiamato *chatbot*: un sistema informatico che interagisce con un umano in una conversazione. "*Eliza*" fu soprannominata "*The Doctor*" perché fingeva di essere uno psicoterapeuta generando risposte basate su determinate parole chiave.[44] Nonostante la sua efficienza, "*Eliza*" non superò mai il *test* di Turing. [45] Sempre nel campo delle *chatbot,* ma meno noto di "*Eliza*", fu "*Parry*", realizzato nel 1972 dallo psichiatra Kenneth Colby, della Stanford *University*. Il programma simulava una persona affetta da schizofrenia paranoide virando sul complottismo quando non riusciva a rispondere contestualmente in modo semplice. "*Parry*" e "*Eliza*" si incontrarono varie volte sostenendo conversazioni medico-paziente e in una conferenza sulle reti informatiche nel settembre del 1972 "*Eliza*" terminò l'incontro richiedendo un compenso di 399,29 dollari per la seduta. [46]

La prima persona elettronica è "Shakey" del 1966, prodotto nel 1969 dalla SRI International. Un *robot* su ruote per uso generale, con autonoma mobilità, capacità di interpretazione delle istruzioni e capacità di azione,[47] che poteva essere azionato tramite una console con dei comandi specifici.[48] Non si negano i difetti che lo stesso possedeva, ma allo stesso tempo aprì il campo alla ricerca sulla robotica mobile.[49]

---

[40] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[41] A. M. TURING, op. cit.
[42] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".
[43] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 1: pag. 2 - 4 "Storia dell'intelligenza Artificiale".
[44] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[45] *Idem.*
[46] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".
[47] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[48] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Appendix 1 (pag. 50 – 52) "A short History of AI".
[49] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".

Gli anni Settanta vengono considerati come un decennio buio per l'intelligenza artificiale. Dopo un periodo di successo e innovazione arrivarono i primi insuccessi a causa di metodi inadeguati in contesti sempre più complessi e ampi. A tal proposito, infatti, i finanziamenti provenienti dal governo americano e inglese furono sospesi. [50]

Nel 1970 nascevano i primi "sistemi esperti"; la conoscenza non era più legata alla comprensione teorica del problema e dunque era necessario anche un approfondimento delle regole dell'esperienza, al fine di poter paragonare il sistema esperto ad un umano esperto. I primi furono "*Dendral*", programmato per desumere la struttura delle molecole organiche in base alle formule chimiche e "*Mycin*", competente per diagnosi mediche e prescrizioni di trattamenti per infezioni batteriche del sangue a partire da informazioni, anche parziali, sui sintomi. [51]

In contemporanea, in Giappone, all'Università di Waseda, veniva progettato il primo *robot* antropomorfo ("WABOT-1"). E nel 1979, James L. Adams realizzava "Standford *Cart*", progettato nel 1961; un *robot* mobile telecomandato e dotato di TV che può essere considerato come uno dei primi veicoli autonomi.[52]

All'inizio degli anni Ottanta si sviluppava l'intelligenza artificiale come industria. E parallelamente, riprendendo in mano lo studio delle reti neurali, nel 1985, quattro gruppi di ricerca scoprirono un nuovo algoritmo di apprendimento basato sulla retro-programmazione dell'errore e lo applicarono ai problemi informatici e ingegneristici fin all'ora irrisolvibili. Nacque, inoltre, la disciplina delle scienze cognitive in cui far confluire la psicologia e l'intelligenza artificiale considerando la macchina come strumento per lo studio della mente. Infine, veniva affiancato, al tradizionale studio della conoscenza, l'approccio sub-simbolico per permettere ai sistemi di utilizzare intelligenza artificiale anche senza avere una rappresentazione dettagliata della conoscenza di base. Venivano concepiti dei metodi probabilistici e *fuzzy* per poter realizzare un ragionamento efficiente sulla base di informazioni incerte.[53]

In Giappone, invece, a partire dal 1982, venne lanciato un programma pubblico per l'evoluzione dell'intelligenza artificiale creando la "*Fifth Generation Computer Systems*". L'idea era di superare le generazioni precedenti di *computer*: *computer* a valvole, a *transistor* e diodi, a circuiti integrati e quelli basati sui microprocessori. La quinta generazione avrebbe dovuto essere più potente delle precedenti facendo lavorare un grande numero di processori utilizzando, però, un linguaggio di programmazione logica.[54]

[50] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 1: pag. 2 - 4 "Storia dell'intelligenza Artificiale".
[51] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".
[52] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[53] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 1: pag. 2 - 4 "Storia dell'intelligenza Artificiale".
[54] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".

Tali idee, purtroppo, tardarono a diventare successi pratici significativi, anche a causa del prosciugamento dei finanziamenti. Prosegue, quindi, il periodo denominato da Nilsson come "inverno dell'intelligenza artificiale" che durerà fino agli anni 2000.[55]

Negli anni Novanta si poté assistere ad un inizio di rinascita grazie al progresso tecnologico che aveva reso più accessibile la realizzazione di tali sistemi. Gli *hardware* erano diventati più economici e affidabili; *Internet* aveva fornito la possibilità di raccogliere enormi quantità di dati e avere una forte potenza di calcolo e di archiviazione per analizzare gli stessi. Tali sviluppi permisero una grande evoluzione dell'intelligenza artificiale dandole la possibilità di influenzare fortemente la vita quotidiana degli esseri umani.[56]

A questo punto, i ricercatori applicarono in concreto il concetto di "imparare giocando". Il campo dei giochi era un'ottima palestra per allenarsi sull'evoluzione e sull'autoapprendimento delle macchine.

Iniziarono dagli scacchi. Nel 1996, "*Deep Blue*" di IBM fu il primo *computer* a sconfiggere il campione del mondo utilizzando un algoritmo denominato "forza bruta" che analizzava milioni di sequenze prima di realizzare la mossa, che con maggiori probabilità, avrebbe portato alla vittoria. Tale metodo, però, non fu efficace per i giochi da tavolo più complessi come il Go. Infatti, nel 2016, *DeepMind,* oggi di proprietà di *Google,* ha realizzato "*AlphaGo*" battendo il vincitore, allora indiscusso, di Go e sostituendo l'algoritmo della "forza bruta" con uno basato sull'estrazione di correlazioni statistiche tra le azioni effettuate dai giocatori al fine di predire la mossa successiva.[57]

Ma il mondo dei giochi basati sull'intelligenza artificiale non termina qui. Nel 1998 venne messo in commercio "*Furby*" il giocattolo da compagnia per bambini di Dave Hampton e Caleb Chung. L'utente poteva insegare al *robot* a parlare, ma il giocattolo, in realtà, poteva imparare solo le parole già iscritte dai programmatori, non avendo, quindi, una vera e propria capacità di apprendimento. Nel 1999 *Sony* realizzò "*AIBO*" (*Artificial Intelligence RoBOt*): un cane *robot* con la capacità di imparare tramite l'interazione con l'ambiente e con i proprietari rispondendo a più di 100 comandi vocali.

Negli anni 2000 l'evoluzione permise a Breazeal di sviluppare "*Kismet*", un *robot* che riconosceva e simulava le emozioni con il suo volto. E a *Honda* di lanciare "*ASIMO*": un *robot* umanoide basato sull'intelligenza artificiale. Ma il vero successo arrivò nel 2002, quando *I-Robot* lanciò "*Roomba*", un *robot* aspirapolvere autonomo con la capacità di pulire casa evitando gli ostacoli.

Nel 2004 fu possibile per la *NASA* mandare un dispositivo basato sull'intelligenza artificiale ad esplorare Marte senza l'intervento dell'uomo.

La nuova frontiera fu, però, il riconoscimento automatico degli oggetti.

Nel 2007 Fei-Fei Li assemblò "*ImageNet*": un *database* di immagini al fine di creare *software* automatico di riconoscimento di oggetti. E nel 2010 *Microsoft* realizzò il "*Kinect*" per "*Xbox* 360": il primo

---

[55] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Appendix 1 (pag. 50 – 52) "A short History of AI".
[56] *Idem.*
[57] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".

dispositivo di gioco che monitorava i movimenti del corpo umano utilizzando una fotocamera 3D e il rilevamento a infrarossi.[58] Una scoperta, che può far sorridere ma risultata fondamentale per l'evoluzione di tale ambito, è dovuta a *Google* che nel 2012 riuscì ad addestrare un algoritmo al fine di riconoscere i gattini nei video di *YouTube* grazie ad una tecnologia in fase di sviluppo denominata "*Deep Learning"*.[59]

Nel 2011 ci fu il debutto degli assistenti vocali e delle *chatbot* in grado di riconoscere il linguaggio naturale, grazie a "*Siri*" di *Apple* che, con un'interfaccia *ad hoc*, aveva (e ha) la capacità di rispondere all'utente in modo coerente in base al contesto tramite comandi vocali, dando la possibilità di usufruire di un'esperienza personalizzata per il singolo soggetto.[60]

Ma la *Apple* non è stata l'unica a muoversi nel campo degli assistenti vocali perché nel 2014 venne seguita da *Microsoft* che lanciò "*Cortana*" e da *Amazon* che produsse "*Alexa*": altoparlanti intelligenti utilizzati come assistenti personali, dando informazioni a richiesta e potendo controllare apparati domestici compatibili e connessi alla rete.[61] Prima ancora di "*Siri*", *Google* nel 2008 aveva lanciato la funzione di riconoscimento vocale per la ricerca *on-line*.[62] Nel 2017, *Facebook* realizzò due *dialog agent* per comunicare tra loro al fine di imparare a negoziare, ma i due si discostarono sempre di più dal linguaggio umano inventando una nuova lingua per dialogare tra loro. [63]

Le evoluzioni non sono terminate qui.

Solo per citarne alcune: nel 2014 *Google* realizzò l'auto senza conducente che superò il *test* di guida autonoma nel Nevada. Nel 2016, *Hanson Robotics* progettò la prima cittadina robotica ("*Sophia*") che vedeva, comunicava e si esprimeva anche tramite espressioni facciali. Nel 2018, *Alibaba*, colosso cinese, con un sistema di intelligenza artificiale superò l'intelletto umano in un *test* di lettura e comprensione. E nello stesso anno *Google* lanciò "*Bert*", un algoritmo realizzato al fine di migliorare la comprensione di ciò che gli utenti vogliono cercare su *Google*.

Ma la tecnologia porta con sé anche aspetti negativi e usi distorti della stessa come la realizzazione di armi autonome basate sull'intelligenza artificiale ai fini della guerra, che possono decidere, senza l'intervento dell'uomo, se sparare o meno. Infatti, nel 2015 Elon Musk, Stephen Hawking e Stave Wozniak firmarono una lettera chiedendo di vietare lo sviluppo di tale tecnologia bellica. [64]

Come abbiamo potuto notare in questa rapida disamina delle fasi storiche più importanti in tema di automazione e intelligenza artificiale, si ha avuto la possibilità di assistere ad un'alternanza tra periodi di grande crescita, che potremmo chiamare "estati" e periodi di delusioni che portarono ad un arresto

---

[58] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[59] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".
[60] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[61] *Idem.*
[62] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".
[63] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[64] *Idem.*

dell'innovazioni denominati "inverni". Ad oggi, possiamo affermare di trovarci nella piena estate dell'intelligenza artificiale grazie ai grandi risultati ottenuti in sede di apprendimento automatico e all'applicazione dello stesso nella quotidianità al fine di migliorare la vita degli esseri umani. [65]

L'ultima novità? Elon Musk, il CEO di *Tesla* ha annunciato il progetto di un *robot* umanoide che verrà lanciato il prossimo anno, "*Tesla Bot*". Il *robot*, dell'altezza di un metro e ottanta, sfrutterà quanto appreso dall'azienda in relazione alle macchine automatizzate per svolgere compiti ripetitivi o pericolosi.[66] Grazie alla sua capacità di sollevare fino a 68kg e di correre fino ad 8 km/h avrà profonde ricadute sull'economia e nel mondo del lavoro. Secondo Elon Musk è necessario dare un volto umano all'automazione e ha dichiarato che, proprio per questo, "*Tesla Bot*" sarà un *robot* amichevole.[67]

### *1.2 L'algoritmo alla base dell'Intelligenza artificiale*

Dopo aver provato a dare una definizione di intelligenza artificiale e dopo aver percorso i tratti salienti della sua storia, è giunto il momento di analizzare il suo funzionamento per poi collegarlo, nei capitoli successivi, alla responsabilità penale.

Tentando di fornire una spiegazione semplice ad un argomento complesso, si considerano intelligenza artificiale i sistemi *software* e *hardware* progettati dall'uomo che, dato un obiettivo, agiscono percependo l'ambiente circostante attraverso l'acquisizione e l'interpretazione dei dati raccolti e che, elaborando le informazioni da questi derivate, decidono le azioni migliori per raggiungere lo scopo prefissato, adattando il loro comportamento e analizzando gli effetti che le loro azioni avranno sull'ambiente.[68]

Ma è necessario soffermarci su ognuno di tali aspetti per comprendere a pieno tale definizione, partendo dal concetto di dato per passare alla spiegazione di cosa sia l'informazione e successivamente ad analizzare l'algoritmo.

Come abbiamo notato nell'evoluzione storica, i progressi dell'intelligenza artificiale sono dovuti a due fattori: l'aumento delle capacità computazionali grazie ai *computer* più veloci e con maggiore memoria e l'aumento dei dati digitali: dati *people to people*, *people to machine* e *machine to machine*. I primi riguardano la digitalizzazione dei documenti, la realizzazione di foto o video e lo scambio di messaggi tramite *social*; i

---

[65] S. QUINTARELLI, op. cit., Capitolo 2 (pag. 33 – 54) "Intelligenza Artificiale: Applicazioni e tecnica".

[66] ANSA, "*Tesla: Elon Musk annuncia l'arrivo del robot umanoide*", 31/08/2021: https://www.ansa.it/sito/notizie/tecnologia/hitech/2021/08/20/tesla-elon-musk-annuncia-larrivo-del-robot-umanoide_a46baca4-a623-472f-9440-2603d965fb47.html .

[67] S. CAMPANELLI, "*2022 anno del Tesla Bot, il robot umanoide da lavoro di Elon Musk*", Huffpost, 20/08/2021: https://www.huffingtonpost.it/entry/tesla-bot-il-robot-umanoide-metallico-di-elon-musk_it_611f881ae4b0e8ac791d153d .

[68] F. BASILE, "*Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*", in Diritto Penale Contemporaneo online, Milano, 2019: https://archiviodpc.dirittopenaleuomo.org/upload/3089-basile2019.pdf; Paragrafo 2 (pag. 4 – 8) "Che cosa intendiamo per intelligenza artificiale?".

secondi, invece, si riferiscono a dati raccolti da istituzioni pubbliche o private come dati fiscali, sanitari o bancari; gli ultimi sono dati generati indipendentemente dall'intervento umano come la geo-localizzazione.[69]

Ogni intelligenza artificiale ha, quindi, per oggetto e presupposto un complesso di dati e informazioni, reperiti *on-line* automaticamente dalla macchina, immessi nel sistema e destinati ad essere elaborati.[70]

Secondo l'Enciclopedia "Treccani", il dato è un elemento di un'informazione costituito da simboli che devono essere elaborati secondo un determinato programma.[71]

Esistono vari tipi di dati: osservazionali, sperimentali, relativi alle simulazioni, al calcolatore o ad analisi di testi e oggetti fisici. Per rendere il dato fruibile è necessario creare meccanismi che consentano all'intelligenza artificiale di analizzarli e di utilizzarli in maniera corretta garantendo: consistenza, qualità, intelligibilità e quantità.[72]

Il rilievo dei dati è generato dall'insieme di varie interazioni che possono essere effettuate tra di loro in modo strutturato creando informazioni con un alto valore aggiunto e dipende da due fattori: la buona qualità del dato e la quantità. Per avere una buona qualità bisogna evitare quelli che vengono denominati "*bias*" cioè gli errori di valutazione che possono incidere negativamente sugli *output* dell'analisi e che derivano anche da sviste nella generazione del dato e nelle annotazioni realizzate dagli esseri umani. [73]

La raccolta massiva dei dati è possibile grazie alla diffusione e all'incremento dell'uso di *Internet*. Ad oggi, infatti, ogni 60 secondi vengono inviati 42 milioni di messaggi tramite le applicazioni di messaggistica e vengono spedite 188 milioni di *e-mail*. [74] È così che nasce il concetto di *bigdata*: enormi quantità di dati che vengono scambiati e utilizzati soprattutto dalle aziende per analizzare i clienti e personalizzare appositamente i servizi offerti. I *bigdata* hanno alcune caratteristiche fondamentali denominate le "cinque V": l'elemento quantitativo, il "Volume"; la "Velocità", la capacità dei dati di essere raccolti ed estratti in tempo reale; la "Varietà" della tipologia dei dati estratti; la "Variabilità" del contenuto dei dati che muta di significato a seconda dell'analisi a cui è sottoposto; il "Valore" legato all'estrazione di un significato rilevante dall'analisi degli stessi. Alcuni studiosi aggiungono anche la "Visualizzazione", necessità di specifiche competenze nella realizzazione di strumenti che possano presentare i risultati dell'analisi e la "Veridicità" dei dati in un determinato contesto.

Analizzando il concetto di "velocità" è importante definire cosa significhi latenza e velocità di trasmissione. La prima è il tempo di risposta che intercorre dal momento in cui un dato viene richiesto fino a quando lo stesso viene reso disponibile all'utente. (misurato in secondi "*s*"); la seconda è la quantità di dati che si possono trasferire in un'unità di tempo ("*bytes/s*"). Nelle reti *internet* ad alta velocità si ha latenza misurabile in "*ms*" (millisecondi) e velocità di trasmissione misurabile in "*Mbytes/s*".

---

[69] *Idem.*
[70] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 3 (pag.52 – 55) "Il "dominio" della macchina: un falso problema".
[71] Enciclopedia Treccani, definizione di "*dato*": https://www.treccani.it/vocabolario/dato/ .
[72] S. QUINTARELLI, op. cit., Capitolo 2 (pag. 33 – 54) "Intelligenza Artificiale: Applicazioni e tecnica".
[73] M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, op. cit., Paragrafo 2 (pag. 2 – 6) "Il dato".
[74] S. QUINTARELLI, op. cit., Capitolo 2 (pag. 33 – 54) "Intelligenza Artificiale: Applicazioni e tecnica".

Questa enorme quantità di informazioni, la possibilità di immagazzinarla in sistemi a costi contenuti e una capacità computazionale cresciuta molto rapidamente negli ultimi anni ha reso l'intelligenza artificiale una tecnologia finalmente fruibile grazie alla realizzazione di algoritmi basati su una notevole potenza di calcolo applicata ad una vasta quantità di dati.[75]

Le informazioni generate dall'intelligenza artificiale si basano sui *bigdata* e dovranno essere contenute all'interno di un *database* organizzato nel modo più simile possibile al ragionamento umano per poter creare algoritmi di intelligenza artificiale. All'interno del *database* sarà inoltre necessario creare relazioni tra i vari *set* di dati per trovare correlazioni al fine di prevedere situazioni del futuro grazie ad algoritmi predittivi.[76]

Per informazione, intesa come oggetto dell'attività dell'elaboratore, si intende quella classe di modelli del reale che hanno la proprietà di essere sottoposti alle attività di elaborazione svolte dalle macchine.[77]

Inoltre, ogni elaboratore basato sull'intelligenza artificiale viene realizzato per uno scopo e gli algoritmi che vengono generati hanno l'obiettivo di far raggiungere al sistema tale fine. Ma sarà sempre il programmatore esterno a definire gli algoritmi e a selezionare i dati e le informazioni da usare collegando appositi criteri di valutazione degli stessi al fine di permettere alla macchina di prendere una decisione. [78]

L'intelligenza artificiale è quindi un insieme di algoritmi: un *software* capace di migliorarsi imparando dai suoi stessi errori.[79] E l'algoritmo viene definito, dall'Enciclopedia "Treccani" come sequenza finita di operazioni elementari eseguibili facilmente da un elaboratore che a partire da un insieme di dati (*input*) produce un altro insieme di dati (*output*) che soddisfano un preassegnato insieme di requisiti.[80] È necessario che i passaggi siano elementari, chiari e non ambigui.[81]

I requisiti che l'algoritmo deve rispettare sono basati su vincoli e obiettivi. I primi devono essere assicurati in ogni caso mentre i secondi devono essere raggiunti al meglio possibile secondo criteri specifici.

L'algoritmo è caratterizzato, inoltre, da due elementi: la complessità computazionale, numero di operazioni elementari necessarie per produrre l'*output*; e l'approssimazione, grado di soddisfazione degli obiettivi secondo il criterio specificato. Da questo dipende il tempo e l'accuratezza della risposta fornita. L'efficacia di un algoritmo è legata alla qualità con cui l'operatore umano trasforma il contesto reale in cui lavora la macchina e le sue finalità, in operazioni matematiche.[82] Tale attività viene definita come "modellizzazione (o rappresentazione) del problema" che porta a quella che può essere denominata come "risoluzione automatica dei problemi" ossia l'obiettivo di dotare l'elaboratore di metodi generali e di programmi efficienti che gli consentano di costruire l'algoritmo di risoluzione del problema. È necessario

---

[75] *Idem.*
[76] M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, op. cit., Paragrafo 2 (pag. 2 – 6) "Il dato".
[77] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 2 (pag. 4 – 5) "Intelligenza Artificiale e l'informatica in una struttura culturale unitaria".
[78] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 3 (pag.52 – 55) "Il "dominio" della macchina: un falso problema".
[79] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[80] Enciclopedia Treccani, definizione di "*algoritmo*": https://www.treccani.it/vocabolario/algoritmo/ .
[81] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[82] M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, op. cit., Paragrafo 4 (pag. 7 – 9) "L'algoritmo".

quindi costruire un meta-algoritmo ("l'algoritmo inferenziale"), in grado di generare un secondo algoritmo che possa fornire le soluzioni indispensabili per raggiungere lo scopo prefissato. Analizzando le fasi eseguite dallo stesso possiamo distinguere tre stadi del "problema": il problema intuitivo, il problema rappresentato e il problema risolto. I dispositivi dotati di intelligenza artificiale hanno l'obiettivo di passare dal problema rappresentato a quello risolto tramite una serie di attività: la rappresentazione, il passaggio dal fenomeno osservabile come *input* al modello computazionale (da problema intuitivo a problema rappresentato); la risoluzione ed esecuzione, ossia il passaggio dal modello alla legge risolutiva (dal problema rappresentato al problema risolto); il confronto e soddisfacimento, ovvero il passaggio dalla legge al fenomeno realizzato come *output* (dal problema risolto al problema intuitivo). Poniamo due esempi pratici: l'algoritmo può essere eseguito dalla macchina o dall'uomo.

Nel primo caso è l'elaboratore a svolgere l'algoritmo inferenziale e a costruire l'algoritmo risolvente il problema per poi eseguirlo e ottenere la soluzione. L'esempio è il gioco degli scacchi: la macchina realizza la costruzione "dell'albero del gioco", ossia l'espressione gerarchica delle sequenze di tutte le mosse convenienti, contromosse e mosse successive; visualizza il percorso da svolgere lungo l'albero del gioco partendo dalla base fino ad arrivare alla posizione più vantaggiosa possibile (tendenzialmente lo scacco matto); ed esegue la prima mossa lungo il cammino individuato.

Nel secondo caso, invece, la macchina realizza gli algoritmi inferenziale e risolvente ma è l'uomo ad eseguirli per giungere alla soluzione. L'esempio è il summenzionato "*Mycin*": la macchina si occupa della costruzione "dell'albero delle concatenazioni" (coppie causa-effetto) per giungere ad una diagnosi di una malattia e alle possibili terapie; è sempre la macchina ad individuare il cammino lungo l'albero delle concatenazioni causa-effetto fino a raggiungere l'identificazione diagnostica della malattia; infine, interverrà il medico a rispondere alle domande che gli vengono poste o ad effettuare il primo esame sul paziente. Questo vale per la maggior parte dei sistemi esperti realizzati.[83]

L'intelligenza artificiale arriva ad una scelta razionale grazie alla percezione dell'ambiente tramite i sensori (fotocamere, microfoni, tastiera o siti *internet*) da cui raccoglie i dati, che devono necessariamente essere logicamente corretti e completi, da cui desume le informazioni che verranno analizzate tramite un algoritmo che le acquisisce come *input* per poi decidere l'azione migliore e agire di conseguenza grazie ai suoi attuatori (*software* e elementi fisici come braccia o ruote) modificando l'ambiente circostante.[84]

I dispositivi quindi percepiscono attraverso tali sensori, i quali trasformano ogni cosa in segnali elettrici codificati usando i *bit*. Qualsiasi informazione venga trasferita o memorizzata in un sistema digitale, viene trasformata in una sequenza di zeri e di uno, perché all'interno del *computer* l'informazione viene rappresentata attraverso il passaggio di corrente elettrica a due diversi livelli di tensione.[85]

---

[83] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 3 (pag. 5 – 6) "Intelligenza Artificiale e risoluzione dei problemi".
[84] F. BASILE, op. cit., Paragrafo 2 (pag. 4 – 8) "Che cosa intendiamo per intelligenza artificiale?".
[85] S. QUINTARELLI, op. cit., Capitolo 2 (pag. 33 – 54) "Intelligenza Artificiale: Applicazioni e tecnica".

Un sistema di intelligenza artificiale deve essere in grado di acquisire, rappresentare ed elaborare conoscenza relativa al compito da eseguire e di applicarla in meccanismi di elaborazione intelligenti, al fine di fornire le prestazioni richieste. Deve quindi essere capace di saper gestire, elaborare e scambiare conoscenza mediante i meccanismi tipici dell'intelligenza umana, quali: l'inferenza, la deduzione, il ragionamento con incertezza, il ragionamento analogico, la generalizzazione, la particolarizzazione, la valutazione di ipotesi e l'apprendimento. Le tecniche e i modi per realizzare tali programmi sono studiati da diverse aree di ricerca. Per citarne alcune: "la ricerca nello spazio degli stati", per programmi che determinano automaticamente una sequenza di azioni per risolvere un problema; "la formalizzazione e il ragionamento probabilistico e *fuzzy*", per la progettazione di programmi che agiscono in condizioni di incertezza riguardo all'ambiente esterno al programma; "la pianificazione", per determinare automaticamente una sequenza di azioni per raggiungere un obiettivo; "l'apprendimento automatico", per la realizzazione di programmi che migliorano le loro prestazioni in base all'esperienza; e "le reti neurali", per applicazioni basate su complesse funzioni non lineari apprese automaticamente dai dati. [86]

Possono quindi essere definite le tre caratteristiche principali dell'intelligenza artificiale: il forte utilizzo dei *bigdata*, l'elevata capacità logico-computazionale e l'uso di algoritmi come il *deep learning* e il *machine learning* per poter prendere decisioni in base ai dati che gli vengono forniti potendo anche modificare gli algoritmi originari man mano che la macchina acquisisce conoscenze.[87]

Sono quindi necessari due elementi all'interno della macchina: il *software* (o l'algoritmo) caratterizzato da meccanismi *machine learning* e il *database* costituito dall'insieme di dati raccolti e analizzati.[88]

### *1.3 Tecniche di apprendimento: machine learning e deep learning. Il Black box algorithm*

All'interno del concetto di intelligenza artificiale bisogna segnalare due aspetti fondamentali: l'autonomia e l'auto-apprendimento. La prima è la capacità di prendere decisioni e di metterle in atto anche senza un controllo o un'influenza esterna; essa dipende dal grado di complessità con cui è stato progettato il rapporto tra la macchina e l'ambiente. Il secondo tratto fondamentale è rappresentato dalla capacità di apprendere dall'esperienza.[89]

Le tecniche di apprendimento utilizzate dall'intelligenza artificiale sono di vario tipo, ad esempio, l'apprendimento automatico (apprendimento profondo e per rinforzo), il ragionamento meccanico

[86] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 4 (pag. 6 – 7) "Aspetti scientifici all'intelligenza artificiale".
[87] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 2 (pag. 76 – 77) "Indeterminatezza delle nozioni di diritto e di intelligenza artificiale" .
[88] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 3 (pag.52 – 55) "Il "dominio" della macchina: un falso problema".
[89] D. IMBRUGLIA, "*L'intelligenza artificiale (IA) e le regole. Appunti"* in rivista quadrimestrale di Media Laws, Rivista di Diritto dei Media, 3/2020 Milano: https://www.medialaws.eu/wp-content/uploads/2020/12/RDM-3-2020-Imbruglia-18-31.pdf; Paragrafo 4 (pag. 26 – 27) "Afferrare il nuovo: l'IA, oggi".

(pianificazione, programmazione, rappresentazione delle conoscenze e l'ottimizzazione) e la robotica (controllo, percezione, sensori e attuatori).[90]

È possibile, quindi distinguere due diverse categorie di intelligenza artificiale: l'intelligenza artificiale "ristretta", che descrive tutti quei sistemi progettati e utilizzati per affrontare compiti specifici, e l'intelligenza artificiale "generale", che indica un sistema che riesce ad adattarsi in modo autonomo e a risolvere qualsiasi compito gli venga assegnato. Sono approcci distinti, il ragionamento logico e l'apprendimento automatico. Il primo metodo si basa su una visione "*top-down*" del problema, che permette di descrivere precisamente un contesto determinato ma la macchina necessita di conoscere molte regole e di allenarsi su molti casi e per questo risulta poco congeniale in contesti più complessi. Il secondo metodo, invece, è un approccio "*bottom-up*", che parte dai dati disponibili per conoscere l'ambiente che gli stessi descrivono al fine di poter agire in modo coerente. Tale metodo utilizza algoritmi particolari e necessita di una notevole capacità di calcolo per poter analizzare un'enorme quantità di dati.[91]

Con il termine "apprendimento automatico" si indicano gli approcci con cui i sistemi possono migliorare le loro prestazioni imparando automaticamente dall'esperienza come eseguire compiti futuri.[92]

Il sistema, costruisce automaticamente un suo modello a partire dall'analisi dei dati, sulla base di un algoritmo di apprendimento automatico. Usando tale modello, il sistema genera classificazioni, valutazioni e previsioni sui nuovi casi che gli sono sottoposti. Ampliando i dati forniti si migliora il modello e di conseguenza le previsioni realizzate dalla macchina.[93]

Come abbiamo potuto osservare fin ora, è possibile rendersi conto di come tutti gli algoritmi delle intelligenze artificiali si basano su enormi quantità di dati e sono dotati di una veloce capacità logico-computazionale e dalla capacità di riconoscere e di interagire con l'ambiente circostante al fine, addirittura, di comunicare e cooperare con altre macchine o esseri umani. Inoltre, le forme più evolute possiedono anche una spiccata capacità di *problem solving*, utilizzando algoritmi che implementano automaticamente la banca dati con nuove informazioni: il c.d. Auto *Machine Learning*.[94] Esso corrisponde all'automazione del processo di applicazione della tecnologia di apprendimento automatico in quanto, dato che risulta estremamente complicato, se non impossibile, codificare manualmente tutte le conoscenze necessarie per far affrontare alla macchina situazioni complesse, l'evoluzione dell' intelligenza artificiale ha spinto sempre di più verso sistemi di autoapprendimento per far si che le macchine potessero accedere autonomamente ai dati conoscitivi.[95]

---

[90] F. BASILE, op. cit., Paragrafo 2 (pag. 4 – 8) "Che cosa intendiamo per intelligenza artificiale?".
[91] S. QUINTARELLI, op. cit., Capitolo 2 (pag. 33 – 54) "Intelligenza Artificiale: Applicazioni e tecnica".
[92] S. J. RUSSELL E P. NORVIG, "*Artificial Intelligence. A Modern Approach*", 3 ed. Prentice Hall, Englewood Cliffs, N. J., 2010.
[93] G. CONTISSA, G. LASAGNI, G. SARTOR, "*Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo"* in Pacini Giuridica, rivista trimestrale di Diritto di Internet: Digital Copyright e Data Protection n. 4/2019, Pisa: https://dirittodiinternet.it/wp-content/uploads/2019/12/2_Sartor.pdf; Paragrafo 1 (pag. 619 – 629) "Introduzione".
[94] M.B. MAGRO, "*Decisione umana e decisione robotica un'ipotesi di responsabilità da procreazione robotica",* in La Legislazione penale, Giustizia Penale e nuove tecnologie, 2020, Dipartimento di Giurisprudenza, Università degli Studi di Torino, Torino: http://www.lalegislazionepenale.eu/wp-content/uploads/2020/05/Magro-Giustizia-penale-e-nuove-tecnologie.pdf; Paragrafo 1 (pag. 1 – 3) "L'automazione della tecnologia di apprendimento automatico o Auto Machine Learning".
[95] J. STEINHOFF**, "***The Automation of Automating Automation: Automation in the AI Industry*", Comunicazione al *Canadian Communication Association Annual Conference*, in *UBC* 2-6 giugno 2019.

In questo modo vengono utilizzati processi di automazione in grado di inglobare dati non elaborati dal modello di apprendimento iniziale, e a prescindere dal programmatore umano.

Si tratta dell'ultima frontiera in campo di automazione. Nella programmazione classica si costruisce un programma per risolvere un problema sulla base di una pregressa ed esatta conoscenza della soluzione dello stesso, con l'apprendimento automatico, invece, il programmatore costruisce un modello che "trova" e "impara" la soluzione a quel problema stesso, senza che il programmatore conosca la soluzione. Viene quindi superato il problema della mancanza di conoscenza dell'umano sia nel campo in cui verrà utilizzata la macchina, sia nell'utilizzo stesso dello strumento. Tutto ciò consente ai sistemi di *cognitive computing* di risolvere problemi di grande complessità con risultati ottimali che migliorano sempre più nel tempo anche di fronte a dati informativi non noti. Tali tecnologie sono in grado di conservare le informazioni acquisite e di utilizzarle in successivi processi decisionali le cui regole vengono elaborate dopo una ricerca statistica fatta dalla macchina. Inoltre, trattandosi di sistemi che hanno l'abilità di imparare dall'esterno, le modalità di azione non possono essere *ex ante* pienamente determinate dai programmatori che hanno però il compito di impartire i *task* alla macchina, dando campo libero all'evoluzione del *robot* che non viene limitata fin dall'inizio. Non tutto è pienamente controllabile dall'uomo sia per numero di operazioni e di dati inseriti ed elaborati, sia per quanto concerne le possibilità di apprendimento. Grazie agli algoritmi appartenenti al sistema di *deep learning* e di apprendimento automatico i *robot* intelligenti possono prendere "decisioni" individuali del tutto autonome, sottratte al controllo umano.[96]

Il *software*, dunque, impara autonomamente dall'ambiente esterno tramite dati che immagazzina e elabora, modificando le proprie prestazioni per adattarle agli esiti del procedimento di apprendimento.[97]

Il *machine learning* è uno dei modi più utilizzati. La macchina apprende similarmente agli esseri umani, attraverso errori e tentativi tramite algoritmi matematico-computazionali per apprendere informazioni dal mondo esterno senza necessariamente essere stata prima programmata perché migliora progressivamente grazie all'esperienza. Gli elaboratori devono, quindi, essere addestrati, come un bambino che impara grazie all'imitazione degli adulti fino a non dover più essere sorvegliato.[98]

Esistono quattro livelli funzionali dell'intelligenza artificiale: la comprensione, capacità della macchina di riconoscere immagini, testi o suoni; il ragionamento, capacità di collegare tra loro le informazioni raccolte per prendere determinate decisioni, l'apprendimento automatico (*machine learning*); e l'interazione, la capacità della macchina di interagire con gli esseri umani.

Ed esistono due categorie di modelli: i simbolici e i non-simbolici. I primi utilizzano rappresentazioni simboliche del problema e dei passi necessari per risolverlo. Tale modello si appoggia ad una costruzione del reale effettuata dai programmatori. Gli *input* sono rappresentati da simboli che poi genereranno *output*. Il

---

[96] M.B. MAGRO, "*Decisione umana e decisione robotica un'ipotesi di responsabilità da procreazione robotica",* op. cit., Paragrafo 1 (pag. 1 – 3) "L'automazione della tecnologia di apprendimento automatico o Auto Machine Learning".
[97] F. BASILE, op. cit., Paragrafo 2 (pag. 4 – 8) "Che cosa intendiamo per intelligenza artificiale?".
[98] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".

sistema esperto si basa su tale modello e su istruzioni *if-then* (se succede A, allora deve succedere B) da ciò, l'intelligenza artificiale, tramite la deduzione logica, riesce a comprendere e ricavare fatti nuovi. I sistemi esperti riescono quindi a fornire soluzioni ottimali a problemi specifici anche con dati incompleti grazie alla *fuzzy logic*, tramite ragionamenti approssimativi che portano alla soluzione più probabile. Tale modello è utile quando la macchina si muove in un ambito dai confini certi e i cui passaggi per giungere alla conclusione sono chiari e trasparenti. Ma le regole devono essere codificate a mano dai programmatori, il che crea problemi di tempo e di efficienza.

I modelli non-simbolici includono anche i modelli basati su connessioni, le reti neutrali e il *deep learning*. Essi vengono utilizzati nei casi in cui la realtà è più mutevole, immettendo nella macchina dati grezzi che verranno analizzati dall'intelligenza artificiale al fine di generare proprie conoscenze implicite grazie alle regole dell'esperienza. Non è necessario quindi l'intervento dei programmatori perché attinge informazioni dai *bigdata* disponibili. L'unica problematica evidentemente riscontrabile è la loro necessità di essere addestrate nel tempo, dato che non possiedono dei binari e si muovono grazie a calcoli probabilistici prima di comprendere pienamente i concetti. Il sistema, quindi, non è pienamente trasparente e rischia di generare *bias* nell'algoritmo.

Per tali motivazioni si spinge verso modelli misti simbolici-non-simbolici: la macchina impara secondo un modello non-simbolico ma sfrutta anche concetti inseriti a mano con regole rigide dagli sviluppatori. [99]

Nell'analizzare il modello non-simbolico è stato toccato il concetto del *deep learning*: in cosa consiste?

Esso sussiste in un tipo di rete neurale artificiale adattiva, particolarmente profonda, addestrata utilizzando un metodo chiamato *back-propagation*.[100]

La rete neurale è composta da diversi strati: un livello di *input*, uno o più *layer* nascosti (*hidden*) e uno di *output*. Ogni strato è costituito da nodi, che rappresenterebbero i neuroni di un cervello umano.

Lo strato di *input* è quello che ha il compito di ricevere ed elaborare i segnali e i dati in ingresso adattandoli alle richieste dei nodi della rete. Lo strato *hidden,* invece, si occupa del processo di elaborazione delle informazioni e può essere costituito da più strati. Lo strato di *output,* infine, raccoglie i risultati dell'elaborazione e li adatta alle richieste del successivo livello-blocco della rete neurale.

Come accennato poco fa, la rete neurale è un sistema adattivo, il che significa che lo stesso ha la capacità di modificare la sua struttura grazie ai numerosi tentativi ed errori che commette. Ogni volta che la macchina "sbaglia" rivede le sue connessioni utilizzando algoritmi di *back-propagation*. Lo scopo dell'adattamento è quello di migliorare l'*output* finale. [101]

---

[99] *Idem.*

[100] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Section 1 (pag. 12 – 17) "What is artificial intelligence?".

[101] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".

Esistono tre modalità di apprendimento per la macchina: *supervised learning* o apprendimento supervisionato, *unsupervised learning* o apprendimento non supervisionato e apprendimento per rinforzo. Il primo, utilizzato principalmente per compiti di classificazione, è basato sul fatto che chi addestra la rete fornisce all'algoritmo esempi di *input* e *output* per permetterle di capire quali sarebbero i risultati desiderati con l'obiettivo di arrivare al momento in cui la rete potrà indentificare automaticamente una regola che potrà utilizzare per compiti simili.[102] Quindi nella fase di *training* si conoscono sia i dati in *input* che i dati in *output*.[103] Inoltre, l'algoritmo viene controllato da macchine addestrate o da esseri umani che supervisionano l'apprendimento.[104] Il secondo, invece, viene utilizzato per realizzare raggruppamenti tra i dati e delle informazioni acquisite, ma non viene fornita, dal programmatore, un'indicazione sul risultato desiderato. È la macchina a dover autonomamente individuare una struttura logica che possa sorreggere l'*input* confrontando dati e ricercando similitudini o differenze.[105] Nella fase di *training* vengono formati dei gruppi di riferimento utilizzabili per classificare nuovi dati.[106] È necessario, poi, un modello interpretativo che spieghi il fenomeno in base al contesto interessato.[107]

L'apprendimento per rinforzo, infine, è molto efficace negli ambienti dinamici e molto variabili. Il sistema apprende direttamente da un ambiente e il suo scopo è raggiungere un obiettivo decidendo autonomamente il percorso per arrivarci. Impara quindi dall'esperienza diretta invece che dai dati fornitigli. L'algoritmo funziona con un sistema di ricompense e punizioni. L'intelligenza umana funziona diversamente? Anch'essa impara attraverso l'esperienza e gli errori commessi nel corso della vita. La macchina emula l'apprendimento umano.

Nelle attuali applicazioni di *deep learning* si tende ad addestrare per molto tempo l'algoritmo con dati etichettati (*supervised learning*), e solo in un secondo momento gli vengono somministrati dati grezzi da cui può continuare a migliorarsi.[108]

Come si è giunti a comprendere tutto ciò?

Fin dagli anni Cinquanta, i giochi da tavolo sono stati una palestra di addestramento estremamente quotata per l'intelligenza artificiale e per quindi giungere ad un'evoluzione del settore. Nei giochi si necessita di fornire le regole (poche e semplici), il contesto (ben delineato e strutturato, senza rischio di imprevisti) e le condizioni per vincere la partita grazie all'intuizione della mossa migliore da compiere.

Un *computer*, partendo da una conoscenza base del gioco, può giungere a livelli elevati grazie all'esperienza che acquisisce continuando a giocare più volte. La conoscenza viene memorizzata e aggiornata

---

[102] *Idem.*
[103] S. QUINTARELLI, op. cit., Capitolo 2 (pag. 33 – 54) "Intelligenza Artificiale: Applicazioni e tecnica".
[104] M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, op. cit., Paragrafo 4 (pag. 7 – 9) "L'algoritmo".
[105] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[106] S. QUINTARELLI, op. cit., Capitolo 2 (pag. 33 – 54) "Intelligenza Artificiale: Applicazioni e tecnica".
[107] M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, op. cit., Paragrafo 4 (pag. 7 – 9) "L'algoritmo".
[108] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".

dopo ogni partita al fine di codificare la funzione espressiva della strategia appresa dalla macchina. Serve solo che i programmatori le forniscano il metodo di apprendimento, il resto lo calcolerà l'intelligenza artificiale.

Con l'evoluzione del tempo i metodi sono migliorati e "*AlphaGo*" ha sostituito il metodo della forza bruta di "*Deep Blue*" perché la prima, invece di analizzare ogni mossa possibile e le sue combinazioni, studia come gli esseri umani abbiano giocato in passato a Go realizzando delle statistiche sulle modalità di gioco per trovare la strategia vincente. Con "*AlphaGo Zero*" ci si è evoluti ancora creando un modello che imparava non più dagli esseri umani ma da sé stesso tramite l'auto-gioco. Qui ci troviamo in un contesto in cui la macchina ha la piena conoscenza di tutte le variabili di gioco ma esistono giochi in cui le informazioni non sono complete. È necessario imparare a gestire anche l'incertezza.

Ciò avvenne grazie al sistema chiamato "*Libratus*", per il gioco del *Poker Texas Hold 'Em*, in grado di agire anche in situazioni in cui ha accesso solo ad informazioni parziali, analizzando, con metodi statistici, le tattiche degli avversari umani. La prossima frontiera saranno i videogiochi in cui la macchina dovrà agire in tempo reale in situazioni imprevedibili e incomplete.[109]

Se quanto visto sinora è, senza presunzione di completezza e accuratezza tecnica, il modo di funzionamento degli algoritmi di *machine learning*, ruolo rilevante rivestono anche i *robot*, ovverosia le macchine impiegate in sostituzione dell'uomo per l'adempimento di determinate operazioni ripetitive, complesse, pesanti o pericolose.

È possibile individuarne tre categorie: i *robot* tele-operati, le cui azioni sono completamente controllate dall'uomo e sono quindi meri strumenti; *robot* autonomi, che dopo la programmazione dell'azione, hanno la capacità di svolgere un compito senza alcun intervento umano; e i *robot* cognitivi, dotati di un sistema per auto-programmarsi e auto-apprendere dalla propria esperienza.[110]

Giunti a questo punto, è importante comprendere il perché di tale parentesi informatica.

È necessario analizzare il funzionamento della macchina per poter poi realizzare un esame dettagliato del tema della responsabilità penale in caso di commissione di illeciti da parte della stessa. Si pone in questo contesto il problema del *black box algorithm*.

Il comportamento dell'intelligenza artificiale è imprevedibile quando deve affrontare una situazione per cui non è stata programmata una risposta e quando grazie a nuove esperienze, la stessa inizi a modellare i dati acquisiti autonomamente. L'uomo quindi non può pienamente controllare o prevedere il comportamento della macchina in situazioni non pianificate. Più gli algoritmi sono capaci di migliorarsi interagendo con il mondo esterno, maggiore sarà la quantità di informazioni elaborate e quindi la capacità di operare autonomamente sui dati.[111] Il loro comportamento può quindi essere *ex ante* imprevedibile. Si parla di *black box algorithms* per indicare l'opacità del meccanismo che porta dall'*input* all'*output*, che riscontra l'osservatore esterno, ma anche

[109] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".
[110] D. IMBRUGLIA, op. cit., Paragrafo 4 (pag. 26 – 27) "Afferrare il nuovo: l'IA, oggi".
[111] 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 algorithms".

il programmatore stesso, che non può avere una piena comprensione di ciò che accade in tale fase. Sono quindi imprevedibili anche i rischi associati a terzi.[112] Quindi, per utilizzare tale tecnologia in ambiti ad alto rischio è necessario che gli stessi siano identificabili e trasparenti.[113]

Per quanto sia vero che la macchina si muova autonomamente e che l'umano non possa essere a conoscenza dell'intera sfera di azioni che il *computer* può realizzare, l'algoritmo rimane condizionato dalle interazioni di chi ne commissiona la creazione, di chi lo crea e di chi lo utilizza, oltre che dall'autoapprendimento della macchina.[114]

### *1.4 Ambiti di applicazione dell'Intelligenza artificiale*

L'intelligenza artificiale è un fenomeno che interessa esclusivamente gli scacchi e i giochi da tavolo? È circoscritto ad una setta di esperti in informatica?

Com'è possibile immaginare, la risposta è certamente negativa. Tutto parte da esperimenti settoriali ma come già abbiamo anticipato nel sotto-paragrafo relativo alla storia dell'intelligenza artificiale, la stessa ha preso sempre più piede nella quotidianità della società. Tante attività che prima venivano svolte autonomamente dall'essere umano, ora non sapremmo più come realizzarle. Le macchine ci hanno viziato e noi ci siamo adattati ad un mondo interamente (o quasi) informatizzato, basato su tecnologie che assecondano i nostri desideri.

Analizziamo rapidamente i settori di maggior interesse nel campo dell'intelligenza artificiale, anche perché è proprio in questi ambiti che si riscontreranno i più importanti risvolti penali di tale materia.

Un sistema basato sull'intelligenza artificiale ha una varietà di applicazioni che vanno dalle macchine a guida autonoma ai *robot* chirurghi passando per i droni da combattimento[115], imitando e migliorando le capacità umane.[116]

Partiamo da dove tutto è iniziato: il mondo dei giochi. È una delle aree tradizionali dell'intelligenza artificiale che ha permesso lo sviluppo di numerose tecniche.[117] Ma l'intrattenimento non si ferma con essi e

---

[112] G. CONTISSA, G. LASAGNI, G. SARTOR, op. cit., Paragrafo 1 (pag. 619 – 629) "Introduzione".
[113] A. CAPPELLINI, "*Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale",* in Criminalia: Annuario di scienze penalistiche, Edizioni ETS, 2018, Pisa: https://discrimen.it/wp-content/uploads/Cappellini-Machina-delinquere-non-potest.pdf; Paragrafo 3 (pag. 5 – 10) "L'imprevedibilità dell'IA avanzate e la crisi del modello vicario".
[114] G.UBERTIS, "*Intelligenza artificiale, giustizia penale, controllo umano significativo*", in rivista trimestrale Diritto Penale Contemporaneo, Milano, 4/2020: https://dpc-rivista-trimestrale.criminaljusticenetwork.eu/pdf/DPC_Riv_Trim_4_2020_Ubertis.pdf; Paragrafo 3 (pag. 78 – 79) "Solo apparente neutralità dell'Intelligenza artificiale e sua opacità".
[115] A. CAPPELLINI, "*Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale",* op. cit., Paragrafo 3 (pag. 5 – 10) "L'imprevedibilità dell'IA avanzate e la crisi del modello vicario".
[116] G. HALLEVY, op. cit., Paragraph 2 (pag. 175 – 177) "What is an Artificial Intelligence Entity? – Machina sapiens – A "Thinking" Machine or a Thinking Machine".
[117] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 5 (pag. 8 – 9) "Applicazioni dell'intelligenza artificiale".

anzi, ultimamente, *Internet* ha reso i contenuti realizzati dagli utenti, una fonte di informazione e di svago. Il "*Kindle*" di *Amazon*, sta tentando di sostituire i libri cartacei permettendo di avere un'intera libreria nel palmo della mano. Inoltre, l'intelligenza artificiale ha permesso una crescita esponenziale dei *social network* a partire da *Facebook*, *WhatsApp* e *Snapchat*, che permettono agli utenti di *smarphone* di comunicare con altre persone. Le comunità *on-line* hanno permesso agli esseri umani di realizzare una propria vita in un mondo virtuale e esterno al reale. Tutto ciò è collegato anche al grande tema della *chatbot*.[118] La comunicazione dal punto di vista linguistico e psicologico è estremamente complessa da realizzare in una macchina. Esistono sistemi che funzionano solo per compiti semplici come impartire ordini, consultare sistemi esperti o tradurre semplici frasi da una lingua all'altra, e ne esistono altri che permettono di riprodurre artificialmente il linguaggio parlato. Ma l'ambito è ancora in via di evoluzione. [119]

Continuando con il campo della sanità, le applicazioni basate sull'intelligenza artificiale, potrebbero migliorare fortemente i risultati sulla salute e sulla qualità della vita. L'unico freno è la fiducia dei medici e dei pazienti e gli ostacoli politici e normativi. Alcuni esempi sono le applicazioni che supportano le decisioni cliniche e monitorano il paziente, le apparecchiature per la chirurgia e le cartelle cliniche elettroniche.[120]

È possibile portare ad emblema due *start up*, una spagnola e una francese. La prima, "$eB^2$", migliora le condizioni di vita dei pazienti con patologie psichiatriche modificando il loro monitoraggio attraverso un'applicazione che raccoglie le informazioni su di essi, sulla loro giornata e sullo stato di avanzamento della loro patologia. I dati elaborati tramite algoritmi verranno inviati ai medici che si occuperanno di una cura *ad hoc* per il soggetto. La seconda, "*Diabeloop*", personalizza le dosi di insulina in base al bisogno reale del paziente affetto dal diabete di tipo 1. È dotata di tre strumenti: un sensore per monitorare il livello di zucchero nel sangue, un erogatore di insulina e un *software* con algoritmi di intelligenza artificiale che analizza le informazioni e dosa l'insulina al momento giusto.[121]

L'ultima frontiera, che porta a discutere sul significato di essere umano e sul trans-umanesimo, e la realizzazione di innesti artificiali basati sull'intelligenza artificiale all'interno del corpo umano per fini terapeutici (ad esempio, *microchip intra* o *extra* cranici).[122]

Nell'ambito finanziario e commerciale è interessante notare le piattaforme economiche digitali che mettono in contatto più utenti (il proprietario della piattaforma, i *providers* che rendono la piattaforma disponibile agli utenti, i *creators* che creano il servizio offerto sulla piattaforma e gli utenti semplici che usufruiscono di tali servizi). Queste piattaforme di mercato virtuale sono basate su sistemi di intelligenza

---

[118] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[119] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 5 (pag. 8 – 9) "Applicazioni dell'intelligenza artificiale".
[120] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Section 2 (pag. 18 – 41) "AI by domain".
[121] M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, op. cit., Paragrafo 6 (pag. 13 – 19) "Le applicazioni dell'intelligenza artificiale tra potenzialità e rischi".
[122] G. UBERTIS, op. cit., Paragrafo 2 (pag. 76 – 77) "Indeterminatezza delle nozioni di diritto e di intelligenza artificiale".

artificiale che mettono in contatto gli utenti in base ai bisogni specifici degli stessi.[123] Inoltre, strumenti di intelligenza artificiale possono essere utilizzati nell'ambito finanziario contro le frodi, nella gestione del rischio bancario e dell'andamento del mercato.[124]

Con il processo di socializzazione della robotica, la macchina è entrata nelle case e ha iniziato ad interagire con l'umano che delega alla stessa lo svolgimento di alcuni compiti.[125] Le case sono sempre più domotiche grazie alla crescita dei *robot* di uso domestico che, anche se lenta, ha portato ad un'ampia diversità di applicazione degli stessi. Gli automi hanno la capacità di consegnare pacchi, pulire uffici e migliorare la sicurezza.[126] Tutto è nato con "*Roomba*", l'aspirapolvere *robot* autonoma, ma è proseguito anche con "*Alexa*", l'assistente vocale intelligente e di controllo della casa.[127]

In ambito di istruzione, le applicazioni sono ampiamente utilizzate da educatori e studenti. È assolutamente necessario che gli insegnanti rimangano umani ma l'intelligenza artificiale permette di personalizzare i metodi di apprendimento e ampliarne le possibilità. Nascono quindi sistemi di tutoraggio intelligente tramite macchine interattive che aiutano lo studente nello studio di determinate materie. L'insegnamento *on-line* ha permesso di ampliare le dimensioni delle classi di studenti. Le difficoltà sono dovute alla mancanza di fondi e alla mancanza di certezza che tali strumenti aiutino gli studenti a raggiungere gli obiettivi richiesti. [128]

La sicurezza pubblica è fondamentale e le tecnologie di intelligenza artificiale sono in grado di migliorarla e ampliarla. Le applicazioni includono l'implementazione di telecamere di sorveglianza che imparano a rilevare anomalie criminali, l'utilizzo di droni e di applicazioni di polizia predittiva[129] (di cui parleremo nei paragrafi successivi). Tali strumenti permettono alla polizia di agire in modo mirato e solo se necessario ma, allo stesso tempo, la sfiducia a cui si assiste negli ultimi anni riguardo i corpi di polizia ha portato anche a tesi opposte che li vedono come applicazioni troppo potenti che rischierebbero di avere effetti negativi sulla società. Questa tecnologia ha riscontrato un grosso successo nell'ambito dei *white collar crime* e nella sicurezza informatica. Saranno utili, inoltre, nell'analisi delle scene del crimine o nelle azioni di ricerca

[123] M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, op. cit., Paragrafo 6 (pag. 13 – 19) "Le applicazioni dell'intelligenza artificiale tra potenzialità e rischi".
[124] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[125] M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, op. cit., Paragrafo 6 (pag. 13 – 19) "Le applicazioni dell'intelligenza artificiale tra potenzialità e rischi".
[126] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Section 2 (pag. 18 – 41) "AI by domain".
[127] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[128] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Section 2 (pag. 18 – 41) "AI by domain".
[129] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".

e soccorso riuscendo a creare una gerarchia di priorità dei compiti da svolgere in un dato contesto. Lo scopo sarà non solo utilizzarla per la risoluzione dei crimini ma anche e soprattutto per la prevenzione degli stessi.[130]

Una disfunzione di utilizzo dell'intelligenza artificiale la notiamo, come avevamo accennato precedentemente, nell'esercito. Molte *ONG* si stanno battendo per far bandire a livello internazionale l'utilizzo di *robot* che hanno la capacità di uccidere senza l'intervento umano.[131]

Ci troviamo nella "Quarta rivoluzione industriale" ed infatti notiamo *robot* basati sull'intelligenza artificiale che collaborano con gli esseri umani al fine di essere più flessibili e precisi grazie a modelli adattivi.[132] La pianificazione autonoma di attività della produzione e della logistica sono realizzate grazie a sistemi informatizzati che operano generando un piano composto da una sequenza di operazioni al fine di raggiungere l'obiettivo assegnato e per poi monitorarne l'esecuzione. Grazie alla robotica intelligente si ha l'obiettivo di realizzare macchine autonome capaci di sostituirsi all'uomo nell'esecuzioni di attività ripetitive e nocive. Tutto ciò avviene con l'ausilio di sensori vicini per riconoscere le forme e gli oggetti e braccia meccaniche per realizzare i movimenti nello spazio. [133]

Come si relaziona tutto ciò con l'occupazione lavorativa degli esseri umani? Il mondo del lavoro sta cambiando sotto i nostri occhi grazie alla globalizzazione e alla crescita della produttività dei grandi colossi mondiali. L'intelligenza artificiale influenza la domanda di lavoro e le richieste di competenze. Le macchine sostituiranno le attività, elimineranno alcuni posti di lavoro ma al contempo ne creeranno di nuovi. [134]

Secondo McKinsey nel 2030, ottocento milioni di posti di lavoro saranno a rischio e il 6% verrà automatizzato del tutto. È necessario formare una nuova generazione di lavoratori legata al mondo dell'intelligenza artificiale dato che la stessa è fondamentale per la produttività delle aziende e competitività dei paesi. [135] Elon Musk, nell'estate del 2021, ha dichiarato, durante il lancio del prototipo "*Tesla Bot*", sul palco dell'evento "*A.I. Day*" che nel futuro, il lavoro fisico sarà esclusivamente una scelta.[136]

Molte decisioni delegate all'intelligenza artificiale hanno impatto sulla vita quotidiana della società, ciò è possibile notarlo anche nelle vetture senza conducente che oramai esistono e non sono più solo un immaginario fantascientifico. Con l'intelligenza artificiale, le auto potranno diventare conducenti migliori delle persone. Già esistono sistemi di rilevazione del traffico in tempo reale e di calcolo del percorso le cui

---

[130] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Section 2 (pag. 18 – 41) "AI by domain".

[131] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".

[132] *Idem.*

[133] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 5 (pag. 8 – 9) "Applicazioni dell'intelligenza artificiale".

[134] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Section 2 (pag. 18 – 41) "AI by domain".

[135] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".

[136] S. CAMPANELLI, op. cit.

informazioni vengono fornite dal *GPS* ai conducenti.[137] Le automobili possiedono sensori per controllare la velocità, la luce esterna e l'umidità così da poter decidere se azionare fari o i tergicristalli: e questi sono solo alcuni esempi delle abilità che una macchina media può svolgere.

Le auto a guida autonoma sono utilizzate in mare e nello spazio ma non ancora in città perché gli ostacoli dati dal traffico e dalle persone sono maggiori. [138]

*Google* ha realizzato un veicolo autonomo senza l'intervento umano e *Tesla* uno semi-autonomo in cui l'umano ha la possibilità di prenderne il controllo in caso di problematiche. Il rischio del semi-autonomo è la possibile distrazione degli utenti che acquisiscono maggiore fiducia nella capacità del veicolo. Il primo incidente stradale che ha coinvolto un'auto autonoma, verificatosi in America nel giugno del 2016, ha messo maggiormente a fuoco tale questione.[139] È proprio in questo campo specialistico in cui i rapporti tra intelligenza artificiale e diritto penale si stringono fortemente.[140]

## ***2. Definizione normativa***

L'aumento della potenza di calcolo e il progresso nell'implementazione degli algoritmi hanno reso l'intelligenza artificiale una delle tecnologie più importanti dell'ultimo secolo.[141]

A causa della rapida e costante evoluzione della ricerca nel campo della robotica, risulta complesso fornire, come abbiamo spiegato nel precedente paragrafo, una definizione univoca e universalmente accettata di cosa sia l'intelligenza artificiale (oltre a chi genericamente la considera come la tecnologia che permette ad un *computer* di analizzare grandi quantità di dati e adottare, sulla base della conoscenza e dell'esperienza acquisita, comportamenti intelligenti o proporre decisioni[142]). Ciò comporta, inoltre, una grande difficoltà di definizione normativa, generando lacune che difficilmente possono essere colmate da interpretazioni estensive di altri testi normativi.

Possiamo quindi notare due diverse posizioni: una conservatrice, che reputa inopportuno intervenire normativamente sulla tecnologia a causa della sua elevata dinamicità, e una che potremmo considerare più

---

[137] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[138] M. SOMALVICO, F. AMIGONI, V. SCHIAFFONATI, op. cit., Paragrafo 5 (pag. 8 – 9) "Applicazioni dell'intelligenza artificiale".
[139] P. STONE, R. BROOKS, E. BRYNJOLFSSON, R. CALO, O. ETZIONI, G. HAGER, J. HIRSCHBERG, S. KALYANAKRISHNAN, E. KAMAR, S. KRAUS, K. LEYTON-BROWN, D. PARKES, W. PRESS, A.L. SAXENIAN, J. SHAH, M. TAMBE, A. TELLER, op. cit., Section 2 (pag. 18 – 41) "AI by domain".
[140] A. CAPPELLINI, "*Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale",* op. cit., Paragrafo 3 (pag. 5 – 10) "L'imprevedibilità dell'IA avanzate e la crisi del modello vicario".
[141] COMMISSIONE EUROPEA, "*Comunicazione della Commissione - Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni - Piano coordinato sull'intelligenza artificiale*" Bruxelles, 7.12.2018, COM (2018) 795 final; Paragrafo 1 (pag. 1 - 2 ) "Introduzione - La strategia europea per l'intelligenza artificiale".
[142] F. DONATI, "*Intelligenza artificiale e giustizia*", AIC: Associazione Italiana dei Costituzionalisti, Rivista N°: 1/2020 (02/03/2020); Paragrafo 1 (pag. 415 - 418) "Premessa".

innovativa in quanto punta verso un aumento, estensione o modifica degli attuali istituti giuridici per adattarli ad un mondo digitalizzato.[143]

È importante ricordare e considerare la dinamicità del diritto, quale capacità di mutare nel tempo adattandosi alle innovazioni del momento, come punto chiave di questo elaborato. La nascita di un ordinamento e di un corpo normativo sono di per sé frutto di evoluzioni storiche e di riunioni di diverse ideologie. Il mondo cambia sotto i nostri occhi ed è necessario che il diritto si sappia adattare al fine di non diventare obsoleto. Ciò può avvenire tramite azioni mirate del legislatore o tramite il lavoro degli interpreti; ma rimane necessaria una certezza normativa che possa garantire il rispetto e la tutela dei diritti fondamentali degli esseri umani.

Gli aspetti da analizzare per giungere ad una definizione normativa di intelligenza artificiale sono molti: i modelli, l'algoritmo, l'automazione, i *bias*, la correttezza, la responsabilità, la trasparenza e la riparabilità. Inoltre è importante disciplinare i problemi di sicurezza, l'interazione uomo-macchina e i rischi per i terzi. Oltre ovviamente agli effetti e alle ricadute che l'utilizzo di sistemi di intelligenza artificiale possono avere sulla società come, ad esempio, la perdita dei posti di lavoro, l'impatto sulla democrazia e sui diritti civili o i rapporti sociali tra gli esseri umani.[144] Uno degli aspetti più complessi da definire è il concetto di "intelligenza" utilizzato in relazione alla macchina; in realtà, tali sistemi, come accennavamo nel primo paragrafo, non sono dotati di intelletto ma allo stesso tempo hanno l'abilità di simularlo e di interagire come tali con l'essere umano.[145] Essendo, questi, concetti sfuggenti in ambito filosofico e scientifico, di conseguenza, lo sono anche per il diritto.

Il mondo del diritto è più lento del mondo della robotica.

### *2.1 Quadro normativo Europeo*

La tecnologia corre più velocemente dei parlamenti e governi.

Lo si nota con il *software* che nella giurisprudenza italiana è stato prima assimilato alle opere cinematografiche e poi alle letterarie. Ed ha ottenuto uno statuto giuridico autonomo solo con la Direttiva 91/250/CEE[146] del Consiglio del 14 maggio 1991 relativa alla tutela giuridica dei programmi per elaborare.

---

[143] D. IMBRUGLIA, op. cit., Paragrafo 1 (pag. 20 – 21) "Regole e rivoluzioni scientifiche".
[144] A. SANTOSUOSSO, op. cit., Paragrafo 2 (pag. 32 – 55) "Il diritto e l'intelligenza artificiale".
[145] C. TREVISI, "*La regolamentazione in materia di Intelligenza artificiale, robot, automazione: a che punto siamo*", in rivista quadrimestrale di Media Laws, Law and Policy of the Media in a Comparative Perspective, Rivista di Diritto dei Media, Milano 25 Giugno 2018, 2/2018 (pag. 447 - 458); Paragrafo 1 (pag 447 – 449) "Il tentativo di una definizione univoca di robot e intelligenza artificiale": https://www.medialaws.eu/la-regolamentazione-in-materia-di-intelligenza-artificiale-robot-automazione-a-che-punto-siamo/
[146] DIRETTIVA 91/250/CEE DEL CONSIGLIO, DEL 14 MAGGIO 1991, relativa alla "*Tutela giuridica dei programmi per elaboratore*": https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:31991L0250&from=MT

L'intelligenza artificiale incide fortemente sull'economia mondiale e di conseguenza è importante analizzare i possibili rischi che la stessa può causare: lo scopo è quello di realizzare una normativa che riesca a bilanciare i vari interessi in gioco, tutelare i diritti fondamentali degli esseri umani senza limitare il progresso e il commercio di sistemi basati su tale tecnologia. Per farlo è necessario definire la tipologia di bene giuridico-economico in cui rientra l'intelligenza artificiale e se la macchina possa essere o meno titolare di diritti e di doveri al fine di comprendere di chi sia la responsabilità in caso di reati commessi dalla stessa. Negli ordinamenti nazionale e sovranazionale l'intelligenza artificiale non forma oggetto di specifiche disposizioni di legge ma si ha la possibilità di estendere l'applicazione di altri istituti giuridici dando risposte parziali a questioni che la riguardano.[147]

L'Unione Europea ha sempre mostrato un forte interesse verso le tecnologie intelligenti anche al fine di far si che le aziende le sviluppino in modo etico e trasparente.[148]

Se l'Unione riuscisse a combinare i suoi punti di forza industriali e tecnologici con un'infrastruttura digitale di elevata qualità e un quadro normativo basato sui suoi valori fondamentali potrebbe diventare un *leader* mondiale nell'innovazione e nell'economia dei dati. Si garantirebbe, così, ai cittadini di usufruire di vantaggi, alle imprese di avere prodotti di ultima generazione e ai servizi di interesse pubblico di ridurre i costi di fornitura migliorando la sostenibilità dei prodotti.

Per permettere tutto ciò, è necessario che l'intelligenza artificiale europea sia fondata sui diritti fondamentali e che inoltre, possa essere considerata affidabile. [149]

Ciò viene provato dalla strategia "Europa 2020" [150] approvata nel 2010 dalla Commissione europea in cui è possibile trovare, tra le iniziative proposte, l'"Agenda europea per il Digitale"[151] con lo scopo di permettere agli Stati Membri di individuare le strategie volte al raggiungimento degli obiettivi comunitari. L'innovazione, come si sa, è rapida, anche più del previsto; di conseguenza, gli Stati Membri hanno dovuto affrontare nuove urgenti sfide senza una politica mirata da parte dell'Unione Europea.[152]

Ad oggi non abbiamo un quadro normativo organico, necessario per garantire maggiore chiarezza rispetto ai doveri e responsabilità dei soggetti coinvolti nel processo di innovazione. Sussistono, invece, molti documenti di *soft law* non vincolanti per gli Stati Membri. La svolta è avvenuta il 21 aprile 2021 con la

---

[147] A. LONGO E G. SCORZA, op. cit., Paragrafo 7 (pag. 198 – 240) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela".
[148] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 5 (pag. 59 – 61) "Le indicazioni dell'Unione Europea sui principi dell'i.a.".
[149] F. DONATI, op. cit., Paragrafo 8 (pag. 433 - 435) "Norme etiche e regolazione in materia di IA".
[150] COMMISSIONE EUROPEA, "*Comunicazione della Commissione - Europa 2020: Una strategia per una crescita intelligente, sostenibile e inclusiva*", COM(2010) 2020 definitivo, Bruxelles, 3.3.2010: http://publications.europa.eu/resource/cellar/6a915e39-0aab-491c-8881-147ec91fe88a.0008.02/DOC_1
[151] COMMISSIONE EUROPEA, "*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni - Un'agenda digitale europea*", COM(2010)245, Bruxelles, 19.5.2010: https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52010DC0245&from=IT
[152] M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, "*Intelligenza artificiale e imaging diagnostico. Implicazioni per il Tecnico sanitario di radiologia medica*", Federazione nazionale Ordini TSRM PSTRP, 8 novembre 2020; Paragrafo 2.1 (pag.13 - 14) "In particolare: in Italia": https://www.congressonazionaletsrm.it/wp-content/uploads/2020/11/Intelligenza-artificiale-e-TSRM-8-novembre-2020-FNO-TSRM-e-PSTRP.pdf

"Proposta di Regolamento dell'intelligenza artificiale"[153] pubblicata dalla Commissione Europea al fine di introdurre regole per prodotti basati sull'intelligenza artificiale con obiettivi di correttezza, sicurezza e trasparenza.

La Proposta esprime la presa di coscienza da parte dell'Unione Europea delle potenzialità dell'intelligenza artificiale nei procedimenti giudiziari ma al contempo ne sottolinea gli impatti negativi sui diritti fondamentali nel caso di utilizzo per valutazioni sulla pericolosità di un soggetto o per la valutazione dell'attendibilità delle prove.[154] L'intelligenza artificiale viene considerata nella relazione di accompagnamento della proposta, come una famiglia di tecnologie in rapida evoluzione in grado di apportare una vasta gamma di benefici economici e sociali in tutto lo spettro delle attività industriali e sociali. [155]

L'oggetto della Proposta riguarda la regolamentazione armonizzata per l'immissione sul mercato, la messa in servizio e l'uso dei sistemi di intelligenza artificiale, il divieto di determinate pratiche, i requisiti specifici per i sistemi considerati ad alto rischio e le regole di trasparenza delle macchine intelligenti. I destinatari delle norme saranno sia pubblici che privati (fornitori di servizi e utenti).

La definizione di intelligenza artificiale fornita, dall'articolo 3 della proposta, è la seguente "*un software sviluppato con una o più delle tecniche e degli approcci elencati nell'allegato I, che può, per una determinata serie di obiettivi definiti dall'uomo, generare output quali contenuti, previsioni, raccomandazioni o decisioni che influenzano gli ambienti con cui interagiscono.*" L'allegato I comprende l'apprendimento automatico (supervisionato o no, per rinforzo) con l'utilizzo del *deep learning*, gli approcci basati sulla logica, conoscenza, programmazione induttiva, ragionamento simbolico e sui sistemi esperti e gli approcci statistici (stima *bayesiana*, metodi di ricerca e ottimizzazione.

Il quadro prospettato si fonda su due livelli di rischio: le intelligenze artificiali a rischio intollerabile (in quanto contrario ai valori dell'Unione) come le pratiche *on-line* di manipolazione cognitiva subliminale e senza consenso che causano danni fisici o psicologici, i sistemi di valutazione sociale dei cittadini con effetti dannosi sproporzionati ovvero l'identificazione facciale biometrica remota "in tempo reale" usata indiscriminatamente dalle forze dell'ordine in luoghi aperti (articolo 5); i sistemi ad alto rischio compreso il riconoscimento facciale o l'utilizzo in contesti rischiosi come l'educazione, l'assistenza sociale, i tribunali e le funzioni di contrasto (valutazioni individuali dei rischi reato o recidiva, poligrafi, affidabilità probatoria, accertamento e indagini), la migrazione e le frontiere (articolo 6 e allegato III che contiene un numero limitato, ma ampliabile, di sistemi di intelligenza artificiale con rischi già concretizzati o concretizzabili in futuro). Il

---

[153] 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

[154] A. M. MAUGERI, "*L'uso di algoritmi predittivi per accertare la pericolosità sociale: una sfida tra evidence based practices e tutela dei diritti fondamentali*", in *Archivio penale* 2021, online, in https://archiviopenale.it/File/DownloadArticolo?codice=79045524-aaff-4497-88fc-be501eff1988&idarticolo=27135 , pp. 1-37; Paragrafo 1 (pag. 1 - 6) "Premessa".

[155] 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, Relazione (pag. 1 – 18) op. cit.

primo tipo è da considerare vietato, mentre il secondo sottostà a delle regole per garantire che gli utilizzatori e fornitori di tali sistemi evitino i rischi e assicurino la sorveglianza umana dell'algoritmo al fine di poterlo modificare e correggere. In realtà, la realizzazione di un elenco completo delle pratiche proibite o ad alto rischio risulta complessa data la versatilità di tali sistemi di intelligenza artificiale.

Esistono due tipologie di sistemi di intelligenza artificiale ad alto rischio: i sistemi destinati ad essere utilizzati come componenti di sicurezza di prodotti (ad esempio in ambito medico o automobilistico) soggetti a valutazione della conformità *ex ante* da parte di terzi (organismi ad hoc che controllano la sussistenza dei requisiti per immetterli nel mercato) e altri sistemi indipendenti che presentano implicazioni principalmente in relazione ai diritti fondamentali come l'identificazione e categorizzazione biometrica delle persone fisiche, la gestione delle infrastrutture critiche, i sistemi utilizzati per l'assunzione e selezione di persone fisiche per i posti di lavoro o per prevenire il rischio di reato tramite poligrafi e lo studio della personalità degli individui. (elencati nell'allegato III).

I requisiti dei sistemi di intelligenza artificiale ad alto rischio riguardano: l'istituzione di un sistema di gestione dei rischi, l'utilizzo di dati per l'addestramento, la redazione di una documentazione tecnica, la registrazione automatica degli eventi, la trasparenza, il controllo umano e l'accuratezza, la robustezza e la *ciber*-sicurezza.

Per quanto concerne l'istituzione di un sistema di gestione dei rischi, esso deve essere basato un percorso continuo durante il ciclo di vita della macchina diviso in quattro fasi (identificazione e analisi dei rischi noti e prevedibili, stima di quelli che possono emergere se il sistema viene usato correttamente per le finalità previste e in caso di uso improprio, valutazione di altri rischi riscontrabili a seguito del monitoraggio successivo all'immissione sul mercato, adozione delle misure adeguate per la gestione del rischio per garantire l'eliminazione o riduzione del rischio tramite la progettazione adeguata del sistema o adottando misure di controllo per i rischi ineliminabili) con l'aggiunta di prove successive per valutare le misure di gestione del rischio più appropriate (articolo 9).

L'utilizzo di dati per l'addestramento dei modelli deve essere sviluppato sulla base di *set* di dati di addestramento, convalida e prova che rispettino determinate caratteristiche: scelte progettuali pertinenti, raccolta e trattamento di dati adeguata alla preparazione degli stessi, formulazione di ipotesi coerenti, valutazione della disponibilità, quantità e adeguatezza del *set*, valutazione di possibili distorsioni o lacune. Tali dati devono essere pertinenti, rappresentativi, completi ed esenti da errori, devono inoltre tener conto della finalità prevista e delle caratteristiche del sistema usato (articolo 10).

L'articolo 11 sancisce la necessità della redazione, prima dell'immissione sul mercato, di una documentazione tecnica in modo da dimostrare che il sistema intelligente sia conforme a tali requisiti.

All'articolo 12, invece, viene espressa l'importanza dello sviluppo di una funzione di registrazione automatica degli eventi all'interno del sistema per garantire un livello di tracciabilità del funzionamento del dispositivo durante il suo ciclo di vita.

Il requisito, forse più importante riguarda la trasparenza del sistema, che deve essere adeguata a consentire agli utenti di interpretare *l'output* del sistema e utilizzarlo adeguatamente. Le informazioni devono riguardare l'identità e i dati di contatto del fornitore, i limiti e capacità delle prestazioni del sistema come la finalità prevista, il livello di accuratezza, robustezza e sicurezza, circostanze note connesse al suo uso e informazioni pertinenti al set di dati (articolo 13). L'importanza di tale principio si denota anche dal fatto che al Titolo IV della Proposta vengono espressi obblighi aggiuntivi di trasparenza per determinati sistemi di intelligenza artificiale (che interagiscono con gli esseri umani, che sono utilizzati per rilevare emozioni o un'associazione con categorie sociali sulla base di dati biometrici e che generano o manipolano contenuti) al fine di tenere conto dei rischi specifici di manipolazione che essi comportano. Quando tali sistemi interagiscono con la collettività, le persone devono esserne informate come ad esempio rivelare la manipolazione automatizzata effettuata a contenuti foto o video che assomigliano a contenuti autentici (salvo le finalità legittime come le attività di contrasto e la libertà di espressione). L'idea di fondo è permettere alle persone di compiere scelte informate e consapevoli.

La sorveglianza umana del sistema serve per garantire la prevenzione e riduzione dei rischi per la salute, sicurezza o diritti fondamentali nei casi di utilizzo conforme del sistema in relazione alla finalità prevista o in caso di uso improprio. Le misure devono essere individuate dal fornitore prima dell'immissione sul mercato, se possibile, e adattate dall'utente successivamente, se necessario (articolo 14).

Infine, l'articolo 15 esprime i requisiti di accuratezza, robustezza (ad errori interni al sistema o nell'ambiente in cui opera a causa dell'interazione con le persone fisiche o altri sistemi) e *ciber*-sicurezza (in relazione ad attacchi che possono avvenire da terzi non autorizzati a modificarne l'uso e prestazioni sfruttando le vulnerabilità del sistema) dell'intelligenza artificiale alla luce della finalità prevista. I livelli devono essere dichiarati nelle istruzioni per l'uso che accompagnano il sistema.

La Proposta non deve essere concepita come un freno all'innovazione, ma anzi, ha lo scopo di incentivarla all'interno di un quadro giuridico favorevole e armonizzato. Vengono incoraggiate le autorità nazionali competenti a creare spazi di sperimentazione in un ambiente controllato dove sottoporre a prova le tecnologie (Titolo V).

Nella Proposta emerge la necessità di istituire un Comitato europeo per l'intelligenza artificiale al fine di contribuire alla cooperazione delle autorità nazionali di controllo competenti (istituite presso ciascuno Stato membro al fine di garantire l'applicazione e attuazione del regolamento nel rispetto dell'obiettività e imparzialità) e della Commissione per quanto concerne le materie disciplinate dal regolamento (articolo 56). Ed è inoltre importante sottolineare come la Proposta vada ad imporre anche il rispetto alla riservatezza delle informazioni e dei dati ottenuti dalle autorità al fine di tutelare i diritti di proprietà intellettuale e le informazioni commerciali riservate o sottoposte a segreto (articolo 71).

Tale progetto tiene fede all'impegno politico della Presidente von der Layen che ha annunciato, alla Commissione 2019-2024 "Un'Unione più ambiziosa", una normativa per un approccio europeo coordinato alle implicazioni umane ed etiche dell'intelligenza artificiale. A seguito dell'annuncio è stato pubblicato il

Libro bianco sull'intelligenza artificiale – un approccio europeo all'eccellenza e alla fiducia (19 febbraio 2020).[156] Tale Libro bianco evidenzia i pregi dell'impiego dell'intelligenza artificiale e l'importanza di effettuare investimenti in tale settore. Non esistono però soltanto gli aspetti positivi e di conseguenza, il Libro si sofferma anche sui rischi che la stessa comporta per i diritti fondamentali, insistendo sulla necessità di un nuovo quadro regolamentare volto a garantire uno sviluppo dei sistemi intelligenti tutelando, al contempo, i diritti umani.[157] Gli elementi costitutivi dello stesso riguardano: la realizzazione di un quadro strategico per allineare gli sforzi a livello europeo, nazionale e regionale al fine anche di giungere all'eccellenza dell'innovazione; e l'individuazione degli elementi chiave per un futuro quadro normativo sull'intelligenza artificiale in Europa nel rispetto dei diritti fondamentali dell'Unione Europea. L'idea è quella di creare fiducia nell'intelligenza artificiale antropocentrica.[158]

Quella del 21 aprile 2021 risulta essere una Proposta molto avanzata che andrà discussa dal Parlamento Europeo con lo scopo di garantire sistemi di intelligenza artificiale affidabili. Allo stesso tempo, la stessa non contempla il riconoscimento di nuovi diritti individuali per i cittadini.[159]

Ma già nel 2017, l'Unione Europea, con normative non vincolanti (come Risoluzione del Parlamento Europeo del 16 Febbraio 2017, recante le raccomandazioni alla Commissione riguardanti le norme civili sulla robotica)[160] tentò di anticipare i legislatori nazionali degli Stati Membri fornendogli una disciplina armonizzata. La Risoluzione aveva lo scopo di differenziare il concetto di "Intelligenza artificiale" e di "*Robot*" suggerendo una disciplina civilistica delle implicazioni giuridiche dell'uso di tali strumenti nelle attività industriali e private.[161]

Nei Considerando viene inoltre affermato che "*è possibile che a lungo termine l'intelligenza artificiale superi la capacità intellettuale dell'uomo*" [162] dato che già ora "*l'umanità si trova (…) 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, rendendo imprescindibile che la legislazione ne consideri le implicazioni e le conseguenze legali ed etiche, senza ostacolarne l'innovazione*".[163] La Risoluzione sottolinea, inoltre, gli aspetti problematici dovuti

---

[156] EUROPEAN COMMISSION, "*White Paper on Artificial Intelligence - A European approach to excellence and trust*", Brussels, 2020: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

[157] F. DONATI, op. cit., Paragrafo 8 (pag. 433 - 435) "Norme etiche e regolazione in materia di IA".

[158] EUROPEAN COMMISSION, "*White Paper on Artificial Intelligence - A European approach to excellence and trust"*, Brussels, 2020; Paragrafo 1 (pag. 1 – 3) "Introduzione" op. cit.

[159] G. MALGIERI E F. PASQUALE, "*L'Europa regola l'Intelligenza Artificiale ad altro rischio Lezione per gli USA*", LUISS University Press, 2 agosto 2021: https://luissuniversitypress.it/l-europa-regola-l-intelligenza-artificiale-ad-alto-rischio/

[160] RISOLUZIONE DEL PARLAMENTO EUROPEO DEL 16 FEBBRAIO 2017 recante "*Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*" (2015/2103(INL)), Gazzetta ufficiale dell'Unione Europea: https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=CELEX%3A52017IP0051

[161] E. BURGIO E L. DE SIMONE, "*Intelligenza Artificiale e responsabilità civile*", 15 aprile 2021, Medialaws, Law and Policy of the Media in a Comparative Perspective; Paragrafo 2 (pag. 1- 2) "Modelli di riferimento applicabili": https://www.medialaws.eu/intelligenza-artificiale-e-responsabilita-civile/

[162] RISOLUZIONE DEL PARLAMENTO EUROPEO DEL 16 FEBBRAIO 2017 recante "*Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*" (2015/2103(INL)), op. cit., Introduzione – Considerando P (pag. 3).

[163] RISOLUZIONE DEL PARLAMENTO EUROPEO DEL 16 FEBBRAIO 2017 recante "*Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*" (2015/2103(INL)), op. cit., Introduzione – Considerando B (pag. 1).

all'innovazione: "*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*"[164] e in più, "*l'apprendimento automatico offre enormi vantaggi economici e innovativi per la società migliorando notevolmente le capacità di analisi dei dati, sebbene ponga nel contempo alcune sfide legate alla necessità di garantire la non discriminazione, il giusto processo, la trasparenza e la comprensibilità dei processi decisionali*".[165]

L'accento posto sulla trasparenza, da parte dalla Risoluzione del Parlamento Europeo sulla robotica, sottolinea la necessità di rendere sempre chiara la logica di base di ogni decisione presa dall'intelligenza artificiale nei casi in cui la stessa possa avere un impatto rilevante sulla vita delle persone.

Il Considerando Z della Risoluzione, nonostante si occupi di responsabilità civile, fa trapelare spunti di configurabilità di una possibile responsabilità penale: "*considerando che, grazie agli strabilianti progressi tecnologici dell'ultimo decennio, non solo oggi i robot sono in grado di svolgere attività che tradizionalmente erano tipicamente ed esclusivamente umane, ma lo sviluppo di determinate caratteristiche autonome e cognitive – ad esempio la capacità di apprendere dall'esperienza e di prendere decisioni quasi indipendenti – li ha resi sempre più simili ad agenti che interagiscono con l'ambiente circostante e sono in grado di alterarlo in modo significativo; che, in tale contesto, la questione della responsabilità giuridica derivante dall'azione nociva di un robot diventa essenziale*".[166] Emergono, quindi, profili fondamentali come il considerare i *robot* come agenti che interagiscono con l'ambiente, collegando ad una loro azione nociva, una possibile responsabilità giuridica considerandoli autonomi e capaci di prendere decisioni senza controlli esterni. Risulta, di conseguenza, impossibile considerarli come meri strumenti nelle mani dell'autore. Le problematiche legate alla natura giuridica di tali entità dipendono proprio dal loro grado di autonomia. Vengono posti dei dubbi sulla necessità di creare nuove categorie giuridiche specifiche. Ci si è resi conto quindi, di come l'attuale assetto normativo sia insufficiente.

L'Allegato di suddetta Risoluzione suggerisce una definizione di *robot* autonomo intelligente come dotato delle seguenti caratteristiche: "*la capacità di acquisire autonomia grazie a sensori e/o mediante lo scambio di dati con il proprio ambiente (inter-connettività) e l'analisi di tali dati; la capacità di apprendimento attraverso l'esperienza e l'interazione; la forma del supporto fisico del robot; la capacità di adeguare il suo comportamento e le sue azioni all'ambiente*."[167] L'elenco permette di comprendere come

[164] RISOLUZIONE DEL PARLAMENTO EUROPEO DEL 16 FEBBRAIO 2017 recante "*Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*" (2015/2103(INL)), op. cit., Introduzione – Considerando G (pag.2).
[165] RISOLUZIONE DEL PARLAMENTO EUROPEO DEL 16 FEBBRAIO 2017 recante "*Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*" (2015/2103(INL)), op. cit., Introduzione – Considerando H (pag.2).
[166] RISOLUZIONE DEL PARLAMENTO EUROPEO DEL 16 FEBBRAIO 2017 recante "*Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*" (2015/2103(INL)), op. cit., Responsabilità – Considerando Z (pag.4).
[167] RISOLUZIONE DEL PARLAMENTO EUROPEO DEL 16 FEBBRAIO 2017 recante "*Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*" (2015/2103(INL)), op. cit., Allegato alla Risoluzione – Raccomandazioni Concernenti il Contenuto della Proposta Richiesta – "Definizione e classificazione dei robot intelligenti" (pag.14).

siano necessarie la fisicità della macchina e l'apprendimento automatico grazie all'interazione con l'ambiente, per poter produrre effetti all'esterno del *robot*.[168]

La Risoluzione del 16 febbraio 2017 pone l'accento su quattro temi di rilevante importanza: la responsabilità per danni, lo *status* giuridico dei *robot* come "persone elettroniche", le possibili dipendenze emotive degli esseri umani rispetto alle macchine dotate di un'autonoma capacità di apprendimento e il problema della disoccupazione a causa dell'automazione.

Nel primo caso, una delle modalità per avere un controllo sui sistemi intelligenti è quella di dotare i *robot* più sofisticati di una scatola nera che possa registrare i dati di ogni operazione compresi i passaggi logici per giungere alle decisioni (idea ripresa nella Proposta di Regolamento del 2021).

Per quanto concerne il secondo punto, si ipotizza una capacità giuridica e di agire completamente nuova, imponendo ai *robot* autonomi, considerati come persone elettroniche, di risarcire essi stessi i danni da loro causati[169] (articolo 59).[170] Secondo il parere del Comitato Economico e Sociale Europeo sull'Intelligenza artificiale 2017/C 288/01, l'introduzione di una personalità dei *robot* comporterebbe un rischio inaccettabile perché si eliminerebbe la funzione preventiva posta alla base della correzione di un comportamento oltre alla possibilità di un abuso di uno *status* giuridico di questo tipo; [171] inoltre, dal punto di vista civilistico, il *robot* è una *res* e come tale non ha un patrimonio proprio capace di rispondere personalmente dei danni da lui causati e sul fronte penalistico, le pene esistenti, nei confronti dei *robot*, non svolgerebbero alcuna funzione special o general preventiva a causa dell'incapacità di una macchina di provare emozioni, paure e inibizioni.

Il terzo punto risulta connesso al il legame che si può instaurare tra un uomo e una macchina nel caso di *robot* utilizzati in ambienti familiari o per l'assistenza agli anziani. Esso è connesso ad aspetti psicologici del singolo individuo e, nonostante i benefici che i *robot* concedono, risulta problematico in quanto può instaurare nell'uomo una dipendenza verso la macchina diventando incapace di agire in assenza della stessa.

Della questione lavorativa abbiamo già accennato nel primo paragrafo e in questo ci limiteremo a sottolinearne la forte rilevanza attuale e la difficolta di giungere ad una soluzione immediata e al passo con l'evoluzione delle macchine. Nelle fasi di sviluppo dei *robot* è fondamentale ricordare come gli stessi debbano essere realizzati al fine di aiutare l'umano nelle mansioni più complesse e pericolose senza, però, sostituirsi completamente allo stesso così da non peggiorarne l'esistenza.

Si spera quindi che il Parlamento Europeo nell'approvare la Proposta di Regolamento del 2021, che finora ha omesso di trattare l'aspetto della personalità e responsabilità propria della macchina, consideri le varie implicazioni legali, soprattutto nei casi in cui la stessa abbia la capacità di assumere autonomamente delle decisioni, regolamentandone la responsabilità giuridica derivante da azioni nocive realizzate dal *robot* stesso. È fondamentale che il progresso in tale ambito avvenga nel rispetto della dignità, autonomia e

---

[168] A. LONGO E G. SCORZA, op. cit., Paragrafo 7 (pag. 198 – 240) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela"
[169] RISOLUZIONE DEL PARLAMENTO EUROPEO DEL 16 FEBBRAIO 2017 recante "*Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*" (2015/2103(INL)), op. cit., Responsabilità – Articolo 59 f) (pag.12).
[170] C. TREVISI, op. cit., Paragrafo 4 (pag. 453 – 455) "L'impulso alla regolamentazione europea dato dalla Risoluzione di febbraio"
[171] A. SANTOSUOSSO, op. cit., Paragrafo 7 (pag. 171 – 208) "Diritti, storicità, artificialità".

autodeterminazione degli individui proprio perché, nel futuro si avrà la possibilità che l'intelligenza artificiale superi l'intelletto umano.[172] Come espresso dalla[173] Comunicazione della Commissione Europea del 2018 "*Artificial Intelligence for Europe*" che inoltre fornisce la seguente definizione: "*l'intelligenza artificiale indica sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere specifici obiettivi. I sistemi basati sull'intelligenza artificiale possono consistere solo in software che agiscono nel mondo virtuale, oppure incorporare l'intelligenza artificiale in dispositivi hardware*".[174]

Per realizzare un quadro normativo coerente e rispettoso dei diritti fondamentali, si necessita di un approccio coordinato, tra gli Stati Membri e la Commissione, al fine di sfruttarne le opportunità e ridurne i rischi. È importante, quindi, dare impulso alla capacità tecnologica e industriale dell'Unione Europea, adottando l'intelligenza artificiale nei settori economici, pubblici e privati, preparandosi ai cambiamenti socio-economici apportati dalla stessa. È necessario realizzare macchine basate sull'intelligenza artificiale in un ambiente improntato sulla fiducia e sulla responsabilità nel rispetto dei principi fondamentali al fine di mitigare le sfide etiche e giuridiche le la stessa propone (come i profili di responsabilità di tali sistemi o il loro utilizzo ai fini della giustizia predittiva). Per generare fiducia è fondamentale che le persone comprendano le modalità di azione della tecnologia intelligente aumentandone la trasparenza e riducendone il rischio di errori. I sistemi di intelligenza artificiale devono, quindi, essere sviluppati in modo da permettere agli esseri umani di conoscerne la loro logica sottostante.[175]

L'affidabilità dell'intelligenza artificiale è stata oggetto del documento, del 2018 e revisionato nell'aprile 2019, "Orientamenti etici per un'IA affidabile" [176], realizzato da un gruppo di cinquantadue Esperti di alto livello (provenienti da industrie, università, istituzioni pubbliche di varie nazioni), nominati dalla Commissione Europea. Esso forniva le linee guida per eliminare l'oscurità e la non trasparenza delle macchine raccomandando la legalità, l'eticità, la robustezza e la sicurezza dei sistemi, ma, soprattutto, la centralità dell'essere umano nel rapporto con la macchina. L'autonomia delle persone deve prevalere sull'autonomia del *robot*. Gli esseri umani devono sempre essere in grado di controllare le macchine in modo tale da poterne limitare le decisioni prese autonomamente.[177] Inoltre tale documento definisce l'intelligenza artificiale come "*sistemi software (ed eventualmente hardware) progettati dall'uomo che, dato un obiettivo complesso, agiscono nella dimensione fisica o digitale percependo il proprio ambiente attraverso l'acquisizione di dati, interpretando i dati strutturati o non strutturati raccolti, ragionando sulla conoscenza o elaborando le*

[172] C. TREVISI, op. cit., Paragrafo 4 (pag. 453 – 455) "L'impulso alla regolamentazione europea dato dalla Risoluzione di febbraio".
[173] A. SANTOSUOSSO, op. cit., Paragrafo 2 (pag. 32 – 55) "Il diritto e l'intelligenza artificiale".
[174] COMMISSIONE EUROPEA, "*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni - L'intelligenza artificiale per l'Europa*", Bruxelles, 25.4.2018, COM (2018) 237 final; Introduzione – Accettare il cambiamento - "Cos'è l'intelligenza artificiale?" (pag. 1).
[175] COMMISSIONE EUROPEA, "*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni - L'intelligenza artificiale per l'Europa*", Bruxelles, 25.4.2018, COM (2018) 237 final; Paragrafo 3.3 (pag. 15 – 18) "Assicurare un quadro etico e giuridico adeguato", op. cit.
[176] EUROPEAN COMMISSION'S HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, "*Ethics Guidelines for Trustworthy AI"*, 2019.
[177] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 5 (pag. 59 – 61) "Le indicazioni dell'Unione Europea sui principi dell'i.a.".

*informazioni derivate da questi dati e decidendo le migliori azioni da intraprendere per raggiungere l'obiettivo dato. I sistemi di IA possono usare regole simboliche o apprendere un modello numerico, e possono anche adattare il loro comportamento analizzando gli effetti che le loro azioni precedenti hanno avuto sull'ambiente".*[178]

Tale definizione è utile per distinguere l'intelligenza artificiale da altri sistemi evoluti ma non artificialmente intelligenti.[179] Possono, dunque, essere contrapposte due diverse teorie: chi considera che la macchina rimanga tale anche se intelligente, e ciò induce a considerare sufficienti le norme oggi vigenti in materia anche per disciplinare il mondo dei *robot* intelligenti, in quanto, dietro ad una macchina è sempre possibile individuare un essere umano che l'ha prodotta o utilizzata; e altri che, invece, ritengono utile garantire un'autonoma soggettività giuridica alle macchine intelligenti puntando, quindi, alla realizzazione di una disciplina normativa *ad hoc*.

Nello stesso anno, la Commissione Europea emanò una Comunicazione legata alla realizzazione di un piano coordinato sull'intelligenza artificiale definendo la stessa come "*quei sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere obiettivi specifici*".[180] L'obiettivo è di porre le persone al centro dell'innovazione tecnologica per risolvere le sfide mondiali più complesse: dall'evoluzione medica alla lotta alla criminalità.[181]

L'applicazione della normativa nazionale ed europea in vigore risulta complessa a causa delle caratteristiche principali dell'intelligenza artificiale: l'opacità e la costante evoluzione dei *software*. Ciò genera difficoltà nell'individuare le violazioni e i responsabili oltre a poter dar luogo a nuovi pericoli non immaginabili.

È necessario un approccio comune europeo all'intelligenza artificiale per evitare una frammentazione del mercato unico e che, l'introduzione di iniziative nazionali, comprometta la certezza del diritto.

Per permetterlo è necessaria la collaborazione tra Stati Membri attuata tramite la strategia sull'intelligenza artificiale adottata ad aprile 2018 dove la Commissione ha presentato un piano coordinato al fine di promuovere lo sviluppo e l'utilizzo dell'intelligenza artificiale in Europa. Tale quadro è composto da settanta azioni comuni per una cooperazione tra Stati e Commissione. Esso durerà fino al 2027 seguendo monitoraggi e revisioni periodiche. L'obiettivo è di massimizzare l'impatto degli investimenti in ricerca, innovazione e diffusione dell'intelligenza artificiale, valutare le strategie nazionali in materia, nonché sviluppare ed estendere il piano coordinato.[182]

---

[178] EUROPEAN COMMISSION'S HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, "*Ethics Guidelines for Trustworthy AI"*, 2019; Glossario – "Intelligenza artificiale o sistemi di IA" (pag.45), op. cit.

[179] S. QUINTARELLI, op. cit., Capitolo 1 (pag. 12 – 32) "Storia dell'Intelligenza Artificiale".

[180] COMMISSIONE EUROPEA, "*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni - Piano coordinato sull'intelligenza artificiale*" Bruxelles, 7.12.2018, COM(2018) 795 final; Introduzione – "La strategia europea per l'intelligenza artificiale" (pag.1), op. cit.

[181] *Idem.*

[182] EUROPEAN COMMISSION, "*White Paper on Artificial Intelligence - A European approach to excellence and trust"*, Brussels, 2020; Paragrafo 4 (pag. 5 – 10) "Un ecosistema di eccellenza", op. cit.

Com'è possibile notare, i rischi posti dalle applicazioni di intelligenza artificiale sono molti (violazione della *privacy*, discriminazione, danni fisici o morali).

È necessario che l'Unione Europea realizzi un quadro normativo europeo per un'intelligenza artificiale affidabile a tutela dei cittadini e del mercato interno, al fine scongiurare il pericolo di frammentazione del mercato interno e di lesione della certezza del diritto.[183]

È fondamentale che il progresso tecnologico sia sostenibile e questo dipende dal rispetto dei diritti fondamentali e dal divieto di pratiche rischiose per gli individui. L'Unione Europea grazie alla Proposta di Regolamento del 21 aprile 2021 sta fornendo le basi solide e programmatiche per realizzare un modello che permetta tale protezione. [184]

### *2.2 Assenza di una normativa Nazionale*

Com'è possibile osservare in quasi tutte le fasi evolutive della società, il diritto giunge sempre in ritardo. Dopo un consistente lasso di tempo che permette di far sedimentare i mutamenti nella collettività, allora il diritto inizia a farsi strada nella sua sfrenata corsa verso la tutela delle libertà degli individui.

Se il giurista non dovesse riuscire a confrontarsi con la nuova realtà limitandosi a ripetere principi propri di altri sistemi senza creare nulla di nuovo, si giungerebbe ad una perdita di capacità regolativa e di lesione dei diritti fondamentali.[185] Il tema dell'intelligenza artificiale ha sempre interessato particolarmente la dottrina giuridica che da anni tenta di inquadrare e di descrivere meglio il fenomeno oggetto di discussione. Gran parte dei contributi giuridici relativi alla ricerca di una disciplina dell'intelligenza artificiale fanno ricorso alle modalità con cui l'uomo ha provato a descrivere tale tipologia di macchina. Nasce, quindi, una forte discussione circa la personalità elettronica e la piena soggettività dei *robot*: una categoria giuridica nuova e con caratteristiche specifiche modellabile o meno sulla personalità giuridica delle persone fisiche o giuridiche. La dottrina maggioritaria reputa sufficienti le attuali regole generali applicabili, altri, invece, riconoscono che le tecniche digitali hanno la capacità di innovare il discorso giuridico e che quindi sia necessario un intervento del legislatore. [186]

Nel caso in cui si volesse considerare il *robot* come persona giuridica, ci si scontrerebbe con la necessità di una rappresentanza e dirigenza svolta da parte di un soggetto; com'è possibile immaginare, ciò risulta complesso per le macchine. Se, invece, venissero considerati come persone fisiche si rischierebbe un'eccessiva umanizzazione della tecnologia. Per non essere più considerati come *res* ma come persone è necessaria

[183] EUROPEAN COMMISSION, "*White Paper on Artificial Intelligence - A European approach to excellence and trust"*, Brussels, 2020; Paragrafo 5 (pag. 10 – 28) "Un ecosistema di fiducia: Quadro normativo per l'IA", op. cit.
[184] G. MALGIERI E F. PASQUALE, op. cit.
[185] D. IMBRUGLIA, op. cit., Paragrafo 2 (pag. 21 – 23) "La lezione di Rodotà: afferrare il nuovo per darvi la forma giusta".
[186] E. BURGIO E L. DE SIMONE, op. cit., Paragrafo 2 (pag. 1 – 2) "Modelli di riferimento applicabili".

l'autocoscienza umana compresa di capacità cognitive ed emotive. Di conseguenza i rischi etici sono molteplici ponendo la questione di cosa significhi essere un umano. L'intelligenza artificiale, come spiegato precedentemente, è un processo di riproduzione del pensiero che imita le capacità di apprendimento al fine di giungere a delle decisioni: la macchina non svolge una funzione indipendente e originale al punto tale da poter essere considerata come un essere umano inteso come centro autonomo di imputazione giuridica. Nessuna delle attuali applicazioni dell'intelligenza artificiale può considerarsi pienamente autonoma, dato che nessuna delle recenti applicazioni può fare a meno dell'essere umano (come programmatore o utilizzatore) e inoltre nessuna macchina ha mai raggiunto quella che può essere considerata come intelligenza umana. [187]

Le difficoltà che riguardano l'identificazione di una definizione universalmente accettata di cosa sia l'intelligenza artificiale si ripercuotono anche sul giurista che trova complesso individuare il catalogo di caratteristiche specifiche che gli consentono di applicare una determinata disciplina. Tale problematica è dovuta da due fattori: la forte interdisciplinarità dell'intelligenza artificiale che rende impossibile l'individuazione di una definizione idonea a rappresentarla in ogni ambito e la forte rapidità di innovazione del settore. È dunque necessario avere una conoscenza approfondita di ciascun tipo di intelligenza artificiale in modo da poterne individuare le caratteristiche per qualificarla in una determinata categoria giuridica. La forte opacità di funzionamento tipica dell'intelligenza artificiale e la sua imprevedibilità di comportamento ha una rilevante ricaduta giuridica nel caso di realizzazione di un danno, ponendo problematiche in relazione alla responsabilità giuridica. Spesso i *robot* sono realizzati al fine di garantire un margine di discrezionalità all'umano sulla decisione finale ma molto spesso, in realtà, la persona si sente vincolata da quanto raccomandato dalla macchina dato che la stessa realizza un calcolo probabilistico sul miglior risultato ottenibile ed eventuali errori sono difficilmente riconoscibili dall'uomo. È dunque indispensabile comprendere se le norme vigenti siano idonee o meno a disciplinare il fenomeno e ad adeguarsi alle nuove sfide che si presenteranno.

I problemi a cui si va incontro sono di due categorie: filosofici, relativi alla dicotomia uomo-macchina che viene lesa dalla realizzazione di *robot* intelligenti con la possibilità di acquisire personalità giuridica e pratici, circostanti la possibilità che la macchina sostituisca l'uomo.

L'analisi di tale tema risulta complessa a causa della necessità di un intervento di rivisitazione delle normative già esistenti nel nostro ordinamento, sussistendo dei dubbi sull'idoneità delle norme giuridiche tradizionali per la risoluzione delle questioni emergenti. [188]

L'OCSE (Organizzazione per la Cooperazione e lo Sviluppo Economico) ha stipulato un Rapporto sulla "*Digital Transformation*" in cui evidenzia l'approccio utilizzato dall'Italia nella gestione dell'intelligenza artificiale. L'Italia si pone al quinto posto dei Paesi G20 per numero di articoli pubblicati in relazione al *machine learning*, dopo gli Stati Uniti, la Cina, l'India e la Gran Bretagna, ed è tra i *leader* europei per la

[187] D. IMBRUGLIA, op. cit., Paragrafo 2 (pag. 21 – 23) "La lezione di Rodotà: afferrare il nuovo per darvi la forma giusta".
[188] E. BURGIO E L. DE SIMONE, op. cit., Paragrafo 2 (pag. 1 – 2) "Modelli di riferimento applicabili".

realizzazione di macchine intelligenti, ma, allo stesso tempo, gli investimenti governativi sono stati fortemente ridotti negli ultimi anni.

Il sistema di intervento pubblico nei settori sensibili viene armonizzato dall'Unione Europea ma è importante che il legislatore italiano adotti strumenti in grado di regolamentare l'intelligenza artificiale, renderla sicura e al contempo garantire all'Italia la possibilità di essere al passo con lo sviluppo tecnologico. In un ambiente in costante trasformazione, è necessario che vengano realizzati interventi finalizzati a rendere più flessibile e sostenibile la disciplina.[189]

### *2.3 Approccio all'estero*

Le difficoltà riguardo l'individuazione di una definizione giuridicamente riconosciuta di intelligenza artificiale, non colpiscono esclusivamente l'Italia.

Gli impatti etici, legali e sociali dovuti alla comparsa di macchine autonome titolari di autonoma soggettività, saranno notevoli e di conseguenza i principi fondamentali dei diritti nazionali saranno soggetti a radicali modifiche per integrare tali nuovi agenti nella teoria generale del diritto.

I problemi in questo ambito rimangono comuni a tutti gli Stati, nonostante le differenze di approccio dei vari governi.

In Giappone, nel febbraio del 2017, il comitato etico della *Japanese Society for Artificial Intelligence* (JSAI), ha approvato le linee guida sull'intelligenza artificiale che dovranno essere osservare dai membri della società. Sono nove articoli interessanti che mettono in relazione tecnologia e società visto il rilevante impatto che l'intelligenza artificiale avrà in futuro in vari settori e i rischi che possono derivarne. A tal fine si spinge per far si che i professionisti specializzati in materia agiscano in modo etico e morale e che siano consapevoli delle loro responsabilità sociali. I nove articoli incidono sul rispetto dell'umanità e della tutela dei diritti fondamentali al fine di evitare minacce alla sicurezza umana nei progetti di intelligenza artificiale; sul rispetto delle leggi e dei regolamenti relativi alla ricerca e sviluppo e alla proprietà intellettuale per non recare danno attraverso la violazione di tali normative; sul rispetto della *privacy* altrui; sull'equità e sull'eliminazione di disuguaglianze e discriminazioni nella società dovute all'ingresso nella collettività di macchine intelligenti; sulla sicurezza al fine di riconoscere la propria responsabilità nel mantenere l'intelligenza artificiale sotto controllo; sull'integrità e affidabilità del progetto; sulla responsabilità sociale dei ricercatori e sviluppatori nel caso di utilizzo a loro conoscenza al fine di danneggiare gli altri; sulla comunicazione con la società e sullo sviluppo personale migliorando la comprensione della società dell'intelligenza artificiale; sul rispetto delle linee guida etiche da parte dell'intelligenza artificiale come se fosse un membro della società. Il fine di tali

[189] M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, op. cit., Paragrafo 2.1 (pag.13 - 14) "In particolare: in Italia".

linee guida è quello di riflettere le diverse opinioni riguardo l'impatto positivo e negativo che avrà l'intelligenza artificiale nella società. È interessante come l'articolo 9 vada a riconoscere una soggettività alle creazioni di intelligenza artificiale considerandole autonome al punto tale che le stesse debbano rispettare le linee guida: la macchina intelligente viene considerata come membro della società. Come notato nel sotto-paragrafo precedente, la questione dell'attribuzione della personalità giuridica alla macchina è un tema caldo che porterebbe un grosso cambiamento nella concezione di soggetto di diritto.

Sempre nel febbraio 2017, il governo Sud-Coreano ha annunciato che i ministeri avrebbero presentato nuovi *standard* legali per definire le responsabilità e i criteri etici delle industrie basate sull'intelligenza artificiale per far si che la società Sud-Coreana sia preparata per la "Quarta Rivoluzione Industriale".[190]

Nel 2018 il rappresentante del Bronx *City*, James Vacca, nel Consiglio della città di New York ha fatto approvare la legge n.49 del 2018 in materia di responsabilità dell'Intelligenza artificiale. Tale introduzione è avvenuta causa dell'incisività degli strumenti di intelligenza artificiale e dei rischi di *bias* raziali e discriminatori nell'utilizzo della stessa in attività amministrative. Fu istituita un *task force*, composta da esperti informatici e funzionari amministrativi, al fine di mappare l'utilizzo delle macchine intelligenti all'interno del Municipio di New York e di elaborare raccomandazioni per l'utilizzo trasparente di tali sistemi. L'idea è quella di attribuire all'intelligenza artificiale una personalità giuridica al fine di ottenere un risarcimento per danni causati dalla macchina.[191]

Alla fine del 2020 l'Ufficio di gestione e *budget* dell'amministrazione Trump ha pubblicato delle linee guida richiedendo alle agenzie federali di presentare un piano per regolare le aziende che utilizzano intelligenza artificiale. Lo scopo non era la protezione dei diritti fondamentali ma la riduzione delle barriere sullo sviluppo e uso dell'intelligenza artificiale. Ad oggi, nessuna agenzia federale ha adempiuto a tale richiesta, dunque, i principi esposti nelle linee guida (accuratezza, sicurezza e affidabilità degli algoritmi) non sono stati supportati da nessuna indicazione sul come possano essere rispettati nella pratica.

In California, il Parlamento discute di una nuova legge per la regolamentazione dei sistemi, utilizzabili in ambito pubblico, di intelligenza artificiale. La pubblica amministrazione deve valutare l'impatto degli algoritmi sui diritti dei cittadini. E le società che forniscono tali servizi devono rendere gli algoritmi spiegabili anche ai non esperti e prevenire i *bias* discriminatori. [192]

Nel 2015, l'ONG "*The Future of Life Institute*" pubblicò una lettera in cui venivano denunciati i rischi legati all'intelligenza artificiale; tra i firmatari è possibile notare Bill Gates, Elon Musk e Stephen Hawking. Secondo tale documento sono state prese sottogamba le questioni legate alla sicurezza per quanto riguarda l'evoluzione e l'utilizzo di *robot* autonomi. [193]

---

[190] C. TREVISI, op. cit., Paragrafo 2 (pag 449 – 452) "Riconoscere una soggettività alle creazioni di AI".
[191] E. BURGIO E L. DE SIMONE, op. cit., Paragrafo 4: (pag. 4- 2) "Cenni all'esperienza Americana (*A Local Law on automated decision system used by agencies*)".
[192] G. MALGIERI E F. PASQUALE, op. cit.
[193] U. PAGALLO, "*Intelligenza Artificiale e diritto. Linee guida per un oculato intervento normativo*", (doi: 10.1422/88512), Sistemi intelligenti (ISSN 1120-9550) Fascicolo 3, dicembre 2017, Il Mulino – Rivisteweb; Paragrafo 1 (pag.615 – 617) "Introduzione".

In campo di sperimentazioni in materia di intelligenza artificiale, il governo giapponese, per affrontare le sfide della robotica, ha realizzato delle zone speciali giuridicamente deregolamentate al fine di effettuare verifiche empiriche sullo sviluppo dei sistemi intelligenti: un laboratorio all'aperto chiamato "*Tokku*". L'idea è quella di creare un'interfaccia *robot*-società in cui gli scienziati possano verificare le abilità di tali entità intelligenti. In questo modo è possibile che gli esseri umani e le macchine inizino a vivere in modo coordinato e rispettoso gli uni degli altri per far si che l'intelligenza artificiale diventi affidabile e trasparente. Le questioni giuridiche incorse a seguito di tali sperimentazioni sono legate al codice della strada, alle trasmissioni radio, alla tutela della *privacy* e alle misure di sicurezza.

In Svizzera la sperimentazione non è da meno, essendoci un programma *Guidami* – "*Autoveicoli automi per mobilità sostenibile*" tra la "*Volvo*", la pubblica amministrazione e altri *partner* con lo scopo di utilizzare le strade pubbliche, in condizioni di guida ordinarie, per i veicoli autonomi o semi-autonomi. Un approccio simile è avvenuto anche in Germania, nei Paesi Bassi e in Francia. [194]

Abbiamo evidenziato precedentemente come l'Europa si stia impegnando per far si che i suoi Stati Membri siano i *leader* indiscussi in tema di intelligenza artificiale garantendo un mercato unico, libero e competitivo ma bisogna sottolineare che la maggior parte di tali strumenti vengono prodotti negli Stati Uniti d'America e in Cina. Nonostante ciò, la normativa europea, precedentemente analizzata, si applicherebbe a tutti i prodotti di intelligenza artificiale impiegati nell'Unione Europea anche se sviluppati all'estero.[195]

### *2.4 Tutela dei diritti fondamentali*

Come abbiamo più volte sottolineato, l'intelligenza artificiale permette all'essere umano di superare notevoli sfide in campo di rapidità, precisione ed efficienza. Emerge, sempre di più, la consapevolezza dell'utilità dell'impiego di tali tecnologie anche al servizio della giustizia. Ma, al contempo, la stessa rischia di ledere fortemente i diritti fondamentali dei singoli individui (come ad esempio la protezione dei dati personali, la *privacy*, la non discriminazione, la libertà di espressione e riunione, la dignità umana, diritto ad un ricorso giurisdizionale effettivo e a un giudice imparziale e la tutela dei consumatori). Nonostante le problematiche evidenziate, i sistemi di intelligenza artificiale trovano una sempre più diffusa applicazione.[196] Proprio per questo, è necessario comprendere come garantire l'utilizzo dei sistemi intelligenti nel rispetto dei diritti umani.[197]

[194] U. PAGALLO, op. cit., Paragrafo 5.3 (pag. 630 - 632) "Sperimentazioni".
[195] G. MALGIERI E F. PASQUALE, op. cit.
[196] F. DONATI, op. cit., Paragrafo 4 (pag. 423 - 425) "Decisioni algoritmiche e tutela dei diritti".
[197] F. DONATI, op. cit., Paragrafo 8 (pag. 433 - 435) "Norme etiche e regolazione in materia di IA".

I rischi possono derivare da difetti di progettazione dei sistemi di intelligenza artificiale o da un errato utilizzo dei dati in entrata o in uscita. Come sappiamo, gli algoritmi permettono di analizzare grandi quantità di dati e di realizzare connessioni tra gli stessi ma questo potrebbe ledere la tutela della *privacy* e del trattamento dei dati personali in base alle modalità con cui ciò avviene, incidendo fortemente sui diritti umani. Le distorsioni sono un forte rischio non solo quando la decisione viene presa dalla macchina ma anche quando dipende dell'intervento umano. Il problema, però, si amplifica esponenzialmente se le discriminazioni avvengono all'interno di *robot* intelligenti, dato l'impatto maggiore che potrebbero avere sulla società senza, per giunta, avere meccanismi di controllo adeguati. Tali distorsioni avvengono soprattutto nelle fasi di apprendimento automatico della macchina dato che i risultati non possono essere previsti in fase di progettazione e i rischi non dipendono da difetti del sistema ma da conseguenze dovute all'assimilazione di dati dall'ambiente. Tutto ciò rientra nel più ampio problema dell'opacità dell'intelligenza artificiale: nella difficoltà connessa alla comprensione dei meccanismi di apprendimento e nell'imprevedibilità delle decisioni dovute ad un comportamento apparentemente autonomo. La disinformazione che colpisce tale settore di studio non permette ai soggetti interessati di comprendere a pieno e di poter verificare la correttezza di una determinata decisione posta in essere dai sistemi di intelligenza artificiale e di conseguenza il rispetto della normativa di riferimento.[198]

Dunque, risulta di rilevante importanza, come evidenziato dal Comitato per gli affari legali del Parlamento europeo nelle raccomandazioni alla Commissione per la normativa civile sulla robotica del 31 maggio 2016, testare i *robot* nella vita di tutti i giorni al fine di identificare e valutare i rischi che da essi possono dipendere anche per via dell'ulteriore sviluppo tecnologico che va al di là della semplice fase sperimentale nei laboratori. Bisogna comprendere i tipi di pericoli che la possibile perdita di controllo sulle macchine possa causare, al fine di riparare i possibili danni.[199]

I principi a riguardo sono legati ai diritti fondamentali dell'uomo codificati nella Dichiarazione universale dei diritti umani, nella Carta dei diritti fondamentati dell'Unione europea e nella Costituzione: la dignità umana, possibilità di limitare la potenza dell'intelligenza artificiale a tutela di un forte antropocentrismo; libertà e diritti civili, libertà di pensiero, opinione, coscienza, uguaglianza, sicurezza, lavoro, salute, capaci di limitare anche l'azione dello Stato; non discriminazione e uguaglianza di trattamento senza distinzioni; prevenzione del danno, con lo scopo di sfruttare le capacità predittive dell'intelligenza artificiale per massimizzarne la potenza di raggiungimento del risultato riducendo i rischi di danni alle persone e alle cose; sostenibilità nella relazione costi-benefici per la società.[200]

Bisogna quindi considerare che i sistemi di intelligenza artificiale non operano in un mondo senza leggi a livello internazionale, europeo e nazionale. In Europa abbiamo sia il diritto primario dell'Unione Europea

[198] EUROPEAN COMMISSION, "*White Paper on Artificial Intelligence - A European approach to excellence and trust"*, Brussels, 2020; Paragrafo 5 (pag. 10 – 28) "Un ecosistema di fiducia: Quadro normativo per l'IA", op. cit.
[199] U. PAGALLO, op. cit., Paragrafo 5.3 (pag. 630 - 632) "Sperimentazioni".
[200] A. LONGO E G. SCORZA, op. cit., Paragrafo 7 (pag. 198 – 240) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela".

(Trattati e Carta dei Diritti Fondamentali dell'Unione Europea) che il diritto derivato (regolamenti sulla protezione dei dati, direttive antidiscriminazione, direttiva sulla responsabilità dei prodotti e regolamento sulla libera circolazione dei dati non personali), la Convenzione Europea dei diritti dell'uomo e le leggi degli Stati membri dell'Unione Europea (soprattutto normative specifiche di settore come il regolamento sui dispositivi medici nel settore sanitario). L'UNESCO, inoltre, nella primavera del 2019 ha elaborato un documento grazie al gruppo COMASET: una piattaforma globale per il dialogo sull'etica dell'intelligenza artificiale riunendo paesi sviluppati e in via di sviluppo con diverse prospettive culturali e morali. La natura giuridica di tale documento è in realtà una raccomandazione proprio al fine di garantire una maggiore flessibilità legata ad un tema etico complesso. Esso punta al rispetto dei diritti umani, al miglioramento della qualità e dell'autonomia della vita degli esseri umani, alla trasparenza del funzionamento dell'intelligenza artificiale e all'informazione dei cittadini sugli aspetti di intelligence e sicurezza coinvolti dall'intelligenza artificiale.[201]

Ma la mancanza di norme chiare causa un'incertezza giuridica che risulta essere deleteria nell'affrontare tali problematiche. Ciò può comportare quindi una riduzione della tutela della società e al contempo una riduzione della competitività delle imprese europee nello scenario globalizzato del mondo contemporaneo.

Di conseguenza chi subisce danni potrebbe non avere accesso ad elementi probatori necessari per giungere in giudizio o per avere la riparazione dell'interesse leso.[202]

## ***3. Intelligenza Artificiale e Sistema Penale***

La giustizia penale rischia il completo stravolgimento dovuto al dirompente ingresso dell'intelligenza artificiale negli strumenti processuali e nelle fattispecie penali.

Gli scenari di cui parleremo nei prossimi sotto-paragrafi sono accumunati dall'assenza di una regolamentazione normativa che, se non colmata, rischia, di generare conseguenze drammatiche.[203]

Come già sottolineato, è necessaria un'innovazione del diritto.

Nell'attuale situazione normativa sostanziale, l'intelligenza artificiale ha la capacità di incidere sia nel diritto civile ed in particolare nel diritto dei contratti e della responsabilità extracontrattuale, che nel diritto penale, sia comune che militare.[204]

Per quanto riguarda il diritto penale militare, il dibattito dell'uso della tecnologia in tale ambito concerne, come precedentemente accennato, l'utilizzo di armi autonome e la *cyber*-guerra. Nel rapporto del 2010 dell'Assemblea Generale in materia di esecuzioni extra-giudiziarie, sommarie o arbitrarie, Philip Alston,

[201] A. SANTOSUOSSO, op. cit., Paragrafo 2 (pag. 32 – 55) "Il diritto e l'intelligenza artificiale".
[202] EUROPEAN COMMISSION, "*White Paper on Artificial Intelligence - A European approach to excellence and trust*", Brussels, 2020; Paragrafo 5 (pag. 10 – 28) "Un ecosistema di fiducia: Quadro normativo per l'IA", op. cit.
[203] F. BASILE, op. cit., Paragrafo 7 (pag. 33) "Quale futuro ci aspetta?".
[204] U. PAGALLO, op. cit., Paragrafo 4.1 (pag. 621) "Lo stato dell'arte".

*Rapporteur* delle Nazioni Unite, affermò la non differenza tra le tipologie di armi utilizzate se risultano "*conformi al diritto internazionale umanitario*" e nello stesso anno, Christof Heyns, membro del Comitato dei diritti umani delle Nazioni Unite, pose i presupposti per analizzare, con un gruppo di esperti, la legittimità dell'utilizzo di armi letali autonome. Le opinioni a riguardo sono delle più disparate: chi richiede un divieto di utilizzo di ogni arma autonoma (come l'ONG "*Human Rights Watch*") e chi le reputa utili per ridurre le morti tra i civili.

In relazione al diritto penale comune, invece, il dibattito si è soffermato principalmente sui reati informatici, analizzando anche i reati commessi da (e non mediante) i sistemi di intelligenza artificiale.[205]

Nel diritto processuale penale, invece, esistono tre ambiti (in prospettiva *ante delictum* o *post delictum* nelle fasi di *policing*, *profiling* e *sentencing*)[206] in cui i dati generati automaticamente possono porsi in contrapposizione con i principi costituzionali a garanzia dei diritti e libertà personali (come la riservatezza o l'equità stessa del processo penale). Il primo riguarda le investigazioni e i mezzi di ricerca della prova, nella cui fase, data la forte intrusività degli strumenti, potrebbe essere lesa la vulnerabilità e la vita privata dei soggetti destinatari. Il secondo aspetto concerne l'acquisizione probatoria, gravata dall'afflusso dei dati generati fuori processo in via automatizzata con il pericolo di non riuscirne a verificare l'attendibilità. Il terzo, infine, attiene all'utilizzo di modelli computazionali per assistere i soggetti del processo penale nella realizzazione di valutazioni e decisioni sulla base di *database* analizzati da *software* in grado di stabilire correlazioni tra i dati. [207]

Tale intromissione della tecnologia nel mondo del diritto, soprattutto nel penale nel quale sono in gioco i diritti e le libertà dei cittadini, presenta aspetti indubbiamente positivi ma allo stesso tempo rischiosi per la tutela dei diritti fondamentali. Si impone, quindi, un bilanciamento di tutti gli interessi in gioco.[208] Infatti, la consapevolezza delle modalità di funzionamento degli strumenti e delle possibili tutele dei diritti ad essi collegati, permetterà l'utilizzo di algoritmi di intelligenza artificiale in modo razionale ampliando le risorse del sistema e senza necessariamente spaventare o mettere a rischio i terzi. La questione principale su cui focalizzare l'attenzione, quindi, non è la problematica sostituzione dell'essere umano con la macchina ma la corretta modalità di scelta dei dati, il rispetto della trasparenza dell'elaborazione e la liceità degli scopi prefissati. [209]

È necessario, nel diritto processuale, evitare che l'intelligenza artificiale diventi un oracolo insormontabile, con conseguente sfiducia nelle capacità umane che comporta una preferenza verso le competenze della macchina. Gli algoritmi mirano a migliorare le prestazioni del sistema preventivo-repressivo

---

[205] U. PAGALLO, op. cit., Paragrafo 4.1.1 (pag. 621 - 623) "IA criminale".
[206] V. MANES, "*L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*", in DisCrimen Articoli, 2020; Paragrafo 2 (pag. 5 – 6) "L'A.I. e il sistema penale...": https://discrimen.it/wp-content/uploads/Manes-Loracolo-algoritmico-e-la-giustizia-penale.pdf
[207] S. QUATTROCOLO "*Processo penale e rivoluzione digitale: da ossimoro a endiadi?*" in rivista quadrimestrale di Media Laws, Rivista di Diritto dei Media, 3/2020, Milano; Paragrafo 3, (pag. 23 – 26) "Afferrare il nuovo e il mito del robot intelligente": https://www.medialaws.eu/wp-content/uploads/2020/12/RDM-3-2020-Quattrocolo-121-135.pdf
[208] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 2 (pag. 49 – 51) "Intelligenza artificiale: di cosa stiamo parlando?".
[209] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 9 (pag. 70 – 71) "Conclusioni".

garantendo efficacia ed efficienza al fine di tutelare al meglio i beni giuridici. Nonostante ciò, alcuni temono la scomparsa del diritto penale e l'eclissi dei suoi principi fondamentali.[210]

Per il diritto sostanziale è importante sottoporre la nostra attenzione verso le problematiche relative alla personalità giuridica, alla responsabilità penale nel caso di commissione di reati e alla possibilità di istituire nuove fattispecie in linea con l'evoluzione tecnologica.

Analizziamo, dunque, gli aspetti critici e al contempo rivoluzionari del settore penale in cui l'intelligenza artificiale potrà incidere (e incide) in modo rilevante.

### *3.1 Polizia predittiva*

"*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'intelligenza artificiale 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.*" Inizia così il *Concept Paper* di presentazione del Convegno annuale di esperti di Polizia del 2019 (OSCE).[211]

Questa può essere considerata la conferma di come l'impiego di strumenti di intelligenza artificiale da parte delle forze dell'ordine è già realtà e l'evoluzione tecnologica porterà ad un incremento nell'utilizzo.

Gli algoritmi permettono l'analisi combinata dei dati garantendo la realizzazione di correlazioni tra gli stessi, fortemente utile per i corpi di polizia nella prevenzione del crimine. Vengono quindi messe in luce le anomalie e attività sospette che gli individui possono commettere in rete o nel mondo reale.

Si tratta quindi di utilizzare sistemi di intelligenza artificiale in attività di *law enforcement* al fine di prevenire reati.[212] Nella maggior parte dei casi vengono realizzate macchine robotiche con lo scopo di svolgere

---

[210] V. MANES, op. cit., Paragrafo 2 (pag. 5 – 6) "L'A.I. e il sistema penale...".
[211] OSCE, ORGANIZATION FOR SECURITY AND CO-OPERATION IN EUROPE, *"Annual Police Experts Meeting: Artificial Intelligence and Law Enforcement: An Ally or an Adversary?*", Conferenza 23-24 September, Wien, (pag.1): https://dirittopenaleuomo.org/wp-content/uploads/2019/07/19.pdf, in rivista DPU, Diritto Penale e Uomo, Criminal Law and Human Conditions, 23 Settembre 2019: https://dirittopenaleuomo.org/segnalazioni/artificial-intelligence-and-law-enforcement-an-ally-or-an-adversary/
[212] F. BASILE, op. cit., Paragrafo 3 (pag. 8 – 9) "Primo percorso d'indagine - IA e attività di law enforcement".

attività di pattugliamento, sorveglianza, disinnesco di bombe, individuazione di volti o atteggiamenti sospetti. Risulta innegabile la loro utilità al fine di tutelare gli agenti umani e di aumentare l'efficienza delle attività ma allo stesso tempo sussistono diverse problematiche: l'effettività di un controllo umano su tali applicazioni di intelligenza artificiale con ampi margini di autonomia sulla scelta degli obiettivi e sulle modalità di azione; la tutela della *privacy* legata all'utilizzo di un notevole quantitativo di dati, anche sensibili, acquisiti in relazione ad un determinato soggetto, considerando inoltre che tali informazioni possono essere manipolate abusivamente generando un grave pregiudizio alle persone a cui si riferiscono; l'utilizzo di armi autonome non letali o letali in relazione al tasso di fallibilità, all'individuazione del responsabile in caso di lesioni commesse per errore e all'assenza, nelle macchine, di empatia e coscienza umana. Risulta dunque necessario un monitoraggio costante e l'elaborazione di un quadro normativo che regoli l'utilizzo dell'intelligenza artificiale nel rispetto dei diritti fondamentali degli individui.[213]

Analizziamo ora il concetto di polizia predittiva. Essa si sostanzia nell'insieme delle attività e dei metodi statistici volti a predire l'individuazione del luogo, del momento in cui verrà commesso un reato e dell'autore dello stesso, al fine di prevenirlo. Le modalità sono basate su una rielaborazione dei dati (difficilmente realizzabile dal solo essere umano) relative alle notizie di reato, agli spostamenti e alle attività dei sospettati, ai luoghi di rischio criminale, al periodo dell'anno e alle informazioni trovate in rete in base a determinati soggetti o attività pericolose. Rientrano anche l'origine etnica, il livello di scolarizzazione e le condizioni economiche riconducibili a soggetti appartenenti a possibili categorie criminali. [214]

A monte è però necessario sottolineare rapidamente la distinzione tra polizia giudiziaria e polizia di prevenzione. La prima riguarda l'insieme di soggetti che esercitano la funzione di individuazione dei reati, dei responsabili e della raccolta delle prove. Essa viene svolta in rapporto con le Procure della Repubblica e non ha scopi preventivi ma repressivi successivi. La seconda, invece, è chiamata ad occuparsi di ordine e sicurezza pubblica prevedendo ed eliminando gli stati di rischio evitando la realizzazione di reati. L'utilizzo in tale ambito di strumenti di intelligenza artificiale non incombe nei limiti fissati dal codice di procedura penale per l'acquisizione delle prove perché l'individuazione dei criminali viene collocata al centro dell'attività.[215]

I *software* di polizia predittiva si dividono in due categorie: quelli che individuano gli *hotspot* criminali, luoghi che possono risultare come scenario di futuri reati; e quelli che individuano i *crime linking*, analizzando la serialità criminale in determinati soggetti per prevenire la commissione di un prossimo illecito penale. [216]

Nei sistemi del primo tipo rientrano gli algoritmi "*Risk Terrain Modeling"* (RTM), "*PredPol"* e "*X-LAW*". Il primo rielabora dati riguardanti i fattori ambientali e spaziali che sembrerebbero favorire la criminalità (poca luce stradale, vicinanza a locali notturni o a stazioni ferroviarie, a parcheggi e a scuole) al fine di prevenire la commissione dei reati di spaccio di sostanze stupefacenti. Questo ha consentito la

[213] F. BASILE, op. cit., Paragrafo 3.1 (pag. 9 – 10) "RoboCop: dalla fantascienza alla realtà?*"*.
[214] F. BASILE, op. cit., Paragrafo 3.2 (pag. 10 – 11) "Sistemi di intelligenza artificiale e polizia predittiva*"*.
[215] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 4 (pag. 55 – 59) "Polizia di prevenzione e polizia giudiziaria".
[216] F. BASILE, op. cit., Paragrafo 3.2 (pag. 10 – 11) "Sistemi di intelligenza artificiale e polizia predittiva*"*.

possibilità di mappare le aree metropolitane per individuare i c.d. *hotspot* al fine di programmare e attuare interventi di prevenzione dei reati di spaccio. Il secondo, invece, già in uso negli Stati Uniti e nel Regno unito, al fine di prevenire la realizzazione di illeciti (soprattutto furti e rapine), utilizza tre tipologie di dati: il tipo di reato, la data e l'ora dello stesso e il luogo di commissione. Infine, il terzo dispositivo in uso presso la polizia italiana per la prevenzione di determinate tipologie di reati risulta basato su un algoritmo capace di rielaborare i dati estrapolati dalle denunce inoltrate alla Polizia di Stato. Esso genera correlazioni nella commissione di illeciti basandosi sulle somiglianze delle condotte e modalità di realizzazione degli stessi.[217]

Passando ora ai sistemi di *crime linking*, è possibile individuare il *software* "*Keycrime*" in uso in Italia con la duplice funzione di utilizzo per polizia di prevenzione e giudiziaria, "*Precobs*" in Germania e "*Hart - Harm Assessment Risk Tool*" in Inghilterra. Tali tecnologie si basano sul principio che alcuni reati, come ad esempio le rapine, possono essere considerati a ripetizione ravvicinata, in quanto sussiste un alto rischio di nuova simile commissione da parte dagli stessi autori in un lasso di tempo limitato in un'area geografica prossima al primo delitto.[218] Dunque, analizzando dati di varie fonti (telecamere o informazioni relative ai precedenti analoghi reati) gli algoritmi hanno lo scopo e l'abilità di individuare il possibile autore prevedendo anche la sua futura mossa valutando le somiglianze nelle modalità di commissione degli illeciti. Il punto di partenza è legato alla qualità e quantità di informazioni raccolte, alla loro completezza e inserimento tempestivo nel programma. I dati vengono successivamente immagazzinati e confrontati tra loro classificandoli in base all'evento e al possibile autore. In questo modo, in tempo reale vengono analizzate tutte le attività che avvengono in relazione a tali dati inseriti e possono essere individuate automaticamente similitudini e dissonanze tra gli stessi.[219] Questi sistemi hanno, inoltre, la capacità di essere utilizzati anche per la ricostruzione del percorso criminogeno di un determinato soggetto. Il loro limite riguarda la possibilità di essere utilizzati solo a fronte di una serialità delle condotte.[220]

In sede investigativa, dunque, permettono di migliorare l'allocazione efficiente delle risorse di *law enforcement* migliorando le tecniche di *predictive policing* e di *profiling*, al fine di mappare e neutralizzare il rischio di commissione dei reati e al contempo puntano all'individuazione dei responsabili di reati già commessi.[221]

Oltre al *predictive policing*, negli ultimi anni si è affiancata anche il *bigdata policing* nata principalmente negli Stati Uniti. Sono entrambi considerati come algoritmi intelligenti che forniscono agli organi statali strumenti per la prevenzione della criminalità. Secondo P. Jeffrey Brantingham (Professore dell'Università della California a Los Angeles), sono tutti e due suddivisi in tre fasi (inserimento di varie tipologie di dati, analisi degli stessi con metodi algoritmici al fine di prevedere la realizzazione del reato, utilizzo di tali previsioni da parte della polizia per porre in essere attività strategiche sul campo) ed ambedue hanno come

[217] F. BASILE, op. cit., Paragrafo 3.2.1 (pag. 11 – 12) "Sistemi di individuazione degli hotspots*"*.
[218] F. BASILE, op. cit., Paragrafo 3.2.2 (pag. 12 – 13) "Sistemi di crime linking*"*.
[219] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 4 (pag. 55 – 59) "Polizia di prevenzione e polizia giudiziaria".
[220] F. BASILE, op. cit., Paragrafo 3.2.2 (pag. 12 – 13) "Sistemi di crime linking*"*.
[221] V. MANES, op. cit., Paragrafo 2.1 (pag. 6 – 7) "...in sede investigativa".

scopo la riduzione dell'incertezza per far si che la polizia possa allocare in modo ottimale le sue risorse per bloccare l'attività criminale sul nascere.[222] La loro differenza sussiste in relazione alla quantità e qualità dei dati raccolti e utilizzati nel *bigdata policing* (serie ampie di dati apparentemente scollegati che gli strumenti analitici convenzionali non riuscirebbero a gestire).[223]

Nonostante la loro abilità nella prevenzione di alcuni tipi di reati, sussistono comunque delle problematicità come il fatto che il loro utilizzo, non essendo stato ancora regolato normativamente in nessuno Stato, è dunque affidato alla sola prassi e all'iniziativa ed esperienza degli agenti di polizia. Ciò non può essere ammesso in quanto si rischia di ledere la tutela della *privacy* e il divieto di discriminazione. Inoltre, si potrebbero generare circoli viziosi dovuti alla capacità di autoalimentarsi posseduta dai sistemi intelligenti perché, considerando una determinata zona come *hotspot* i controlli in quel luogo si intensificherebbero e dunque i reati rilevati risulterebbero maggiori. Si rischia, per giunta di realizzare una militarizzazione della sorveglianza in alcune aree urbane o verso determinati soggetti in quanto si punta a preventivi interventi attivi della polizia. Infine, sussiste il problema della segretezza degli algoritmi alla base di tali sistemi dovuta alla tutela dei brevetti depositati da aziende private, non potendo, quindi, garantire una piena comprensione del loro funzionamento ledendo, dunque, i requisiti di trasparenza e di verificazione dell'affidabilità dei risultati da essi prodotti.[224] Viene inoltre sottolineato come il focalizzarsi su un determinato soggetto al fine di prevenire i suoi futuri reati, invece di generalizzare il controllo in modo simile su tutti i possibili autori, possa risultare discriminatorio anche nel caso in cui, successivamente, esso risulti veramente il *reo* di un illecito. Ciò però non può essere affermato quando la previsione è frutto di valutazione basata su elementi differenti e bilanciati tra loro e non solo su una probabilità statistica data dalla quantità degli episodi commessi da un soggetto o in uno specifico luogo. Per quanto concerne l'utilizzo di tali sistemi per addossare ad un determinato soggetto la responsabilità penale per reati commessi in passato (attività di polizia giudiziaria) è necessario che tutti gli elementi inseriti nella banca dati rispettino li requisiti insiti nell'art.189 c.p.p. in relazione alle prove atipiche (idoeità ad accertare i fatti e non lesione morale della persona): da tale punto di vista, basandosi su elementi statistici, non può dirsi che questi strumenti accertino alcunché.

Soffermiamoci per un attimo su come la rivoluzione digitale abbia inciso fortemente sulle modalità di investigazione sempre più basate sull'*hacking* (accesso occulto a sistemi di produzione di dati). Lo stesso consiste in azioni intrusive realizzate tramite l'installazione da remoto di *malwares* (*malicious software* con la capacità di accedere a diverse funzioni di apparati digitali in cui vengono inseriti senza che l'utente ne possa essere a conoscenza). Essi consentono di acquisire informazioni scambiate dall'interessato, di attivare da remoto la geo-localizzazione o la registrazione audio-video e di accedere e manipolare *file* presenti

[222] P. J. BRANTINGHAM, "*The Logic of Data Bias and its Impact on Placed-Based Predictive Policing*", in Ohio State Journal of Criminal Law, 2018, p.473.
[223] C. BURCHARD, "*L'intelligenza artificiale come fine del diritto penale? Sulla trasformazione algoritmica della società*", Rivista Italiana di Diritto e Procedura Penale – n.4 – 2019; Paragrafo 3 (pag. 10 – 19) "Le promesse dell'intelligenza artificiale per l'amministrazione della giustizia penale".
[224] F. BASILE, op. cit., Paragrafo 3.2.3 (pag. 13) "Considerazioni conclusive sui sistemi di polizia predittiva*".*

nell'*hardware* infetto del destinatario. L'assenza di un apposito quadro normativo ha spinto la prima giurisprudenza a muoversi entro i margini della disciplina esistente dell'intercettazione delle comunicazioni minimizzandone (erroneamente) l'intrusività. Successivamente alla legge Orlando si è inserita una decretazione d'urgenza che ne ha ampliato l'utilizzo con una regolamentazione approssimativa: d.l. 161/2019 (conv. con mod. in l. 7/2020)[225] e l'ultimo d.l. 28/2020 (conv. con mod. in l. 70/2020)[226]. Tali mezzi collidono con le tutele costituzionali della riservatezza, del domicilio e della corrispondenza oltre al fatto che, dato il costante utilizzo di dispositivi su cui possono essere impiantati tali *malwares*, risulta incongrua l'applicazione della distinzione tra luoghi pubblici, aperti al pubblico e privati. L'articolo 8 della CEDU pone dei limiti anche all'attività investigativa di analisi e profilazione dei dati.

L'appena terminata digressione riguardate l'*hacking* investigativo risulta necessaria in tale sede in quanto le informazioni acquisite in tal modo andranno ad alimentare il bagaglio di informazioni utili per l'apprendimento delle macchine intelligenti utilizzate in sede di polizia predittiva.[227] Ma come già affermato, i dati utili per il processo non vengono acquisiti esclusivamente in tale modalità, ma anche e soprattutto, attraverso qualsiasi supporto digitale. Grazie alla crescita dell'*Internet of Things*, si può trattare di dati generati autonomamente da parte di oggetti di uso quotidiano (assistenti vocali o elettrodomestici intelligenti), senza alcun intervento umano, risultando, quindi, un patrimonio conoscitivo fondamentale per le indagini. Bisogna però necessariamente valutare l'accuratezza del dato generato e raccolto attraverso il solo strumento digitale tramite un procedimento raramente trasparente (data la non conoscenza degli *input* e dei processi elaborativi) risultando non verificabile *ex post*. Si rischia di ricadere in un grave squilibrio conoscitivo tra le parti del processo (già presente, nei casi in cui si necessiti di conoscenze specialistiche, date le risorse illimitate della parte pubblica a discapito della privata).

All'interno del diritto ad un equo processo sancito dall'articolo 111 della Costituzione e dall'articolo 6 della CEDU è possibile riscontrare i criteri di ammissibilità e utilizzabilità della prova al fine di garantire il *fair trial* cercando di evitare l'eccessiva lesione del principio della parità delle armi.[228] È necessario, dunque, che ciascuna delle parti abbia effettiva conoscenza delle argomentazioni della controparte e che fruisca della concreta possibilità di contrastarle.[229]

---

[225] D.l. 161/2019 (conv. con mod. in l. 7/2020), "*Modifiche urgenti alla disciplina delle intercettazioni di conversazioni o comunicazioni*": https://www.gazzettaufficiale.it/eli/id/2019/12/31/19G00169/sg

[226] D.l. 28/2020 (conv. con mod. in l. 70/2020), Testo del decreto-legge 30 aprile 2020, n. 28 (in Gazzetta Ufficiale - Serie generale - n. 111 del 30 aprile 2020), coordinato con la legge di conversione 25 giugno 2020, n. 70 (in questa stessa Gazzetta Ufficiale alla pag. 1), recante "*Misure urgenti per la funzionalità dei sistemi di intercettazioni di conversazioni e comunicazioni, ulteriori misure urgenti in materia di ordinamento penitenziario, nonché disposizioni integrative e di coordinamento in materia di giustizia civile, amministrativa e contabile e misure urgenti per l'introduzione del sistema di allerta Covid-19*": https://www.gazzettaufficiale.it/eli/id/2020/06/29/20A03469/sg

[227] S. QUATTROCOLO "*Processo penale e rivoluzione digitale: da ossimoro a endiadi?"* op. cit., Paragrafo 2 (pag. 124 - 126) "Rivoluzione digitale, investigazione penale e riservatezza".

[228] S. QUATTROCOLO "*Processo penale e rivoluzione digitale: da ossimoro a endiadi?"* op. cit., Paragrafo 3 (pag. 126 – 132) "La prova generata automaticamente e i rischi per la parità delle armi".

[229] CEDU, *Brandstetter v. Austria*, § 68.

Inoltre, a seguito dell'utilizzo di tali sistemi, si rischia di generare un regime repressivo basato sulla sorveglianza di massa ledendo diritti e libertà fondamentali. L'esempio eclatante di questo problema avviene in Cina, dove, per ridurre il tasso di attentati terroristici di cui è accusata la minoranza musulmana degli Uiguri, è stato realizzato un sistema di dati basato sull'intelligenza artificiale per monitorare e tenere sotto controllo il loro spostamenti informando la polizia nel caso di attività sospette (come il raduno di più persone in un unico luogo o il passaggio di taluno in un'area con potenziale intento terroristico). Tale sistema è costituito da telecamere di sorveglianza posizionate in tutto il paese con l'abilità di identificare i soggetti, in base alle loro particolare caratteristiche fisiche e facciali, nell'arco di 7 minuti. Dal 2017 tutti i veicoli vengono tracciati tramite la costellazione "*Beidou*" (sistema geo-localizzazione cinese) con lo scopo di analizzarne i tragitti, grazie ai dispositivi di navigazione satellitare presenti nelle autovetture. Inoltre, tutti i coltelli possiedono un codice *QR* connesso alla carta di identità del proprietario. Inoltre, chiunque utilizzi un cellulare ha l'obbligo di effettuare il riconoscimento facciale: l'immagine verrà archiviata e collegata al documento e alla *SIM* del proprietario. Nel 2019 il Governo ha dichiarato di voler inserire il riconoscimento delle espressioni facciali al fine di comprendere la volontà di realizzare un attacco violento. Tra i dati analizzati dagli algoritmi di controllo rientrano, inoltre, quelli acquisiti tramite l'obbligatorio controllo sanitario che viene effettuato a tutti gli adulti del Paese (DNA, gruppo sanguigno, impronte digitali, registrazioni vocali e scansioni viso).

Com'è possibile immaginare, questi sistemi non vengono utilizzati esclusivamente in Cina, ma anche in Francia, a Londra e negli Stati Uniti. Il tasso di fallimento è passato dal 5% allo 0,1% dal 2010 al 2020 grazie all'esorbitante numero di immagini che vengono caricate ogni giorno sulle piattaforme *social* e grazie alla realizzazione di *chip* maggiormente sofisticati che accelerano il processo di apprendimento della macchina intelligente.

A sottolineare la necessità di una tutela della *privacy* degli individui, Frank La Rue, *ex* relatore speciale delle Nazioni Unite, nel 2013 dichiarò che "*i dati delle comunicazioni sono memorizzabili, accessibili e ricercabili e la loro divulgazione e il loro utilizzo da parte delle autorità statali sono ampiamente non regolamentati. L'analisi di questi dati può essere sia altamente rivelatoria sia invasiva, in particolare quando i dati vengono aggregati. Gli Stati stanno attingendo sempre di più ai dati delle comunicazioni per supportare l'applicazione della legge o indagini per la sicurezza nazionale. Gli Stati stanno anche costringendo le aziende a conservare i dati di comunicazioni per condurre una sorveglianza massiva nel tempo.*" Si necessita, dunque, di una regolamentazione che tuteli gli utenti. È indubbio come tali dati siano fondamentali per le indagini al fine di contrastare alcune tipologie di reati ma allo stesso tempo, rendere gli strumenti di comunicazione di massa meno sicuri causerebbe conseguenze negative per la democrazia dei Paesi.[230] Infatti David Kaye, dell'Alto Commissariato delle Nazioni Unite per i diritti umani, nel rapporto del 2019 affermò che la sorveglianza deve rispettare vari principi al fine di poter essere considerata come necessaria e proporzionata. È interessante soffermarsi sul principio di liceità della stessa, infatti è fondamentale che sussistano norme

[230] A. LONGO E G. SCORZA, op. cit., Paragrafo 3 (pag. 88 - 147) "Vantaggi e rischi non economici dell'AI - I rischi etici, sociali e politici dell'AI - *Sorveglianza e controllo sociale*".

precise, chiare e accessibili al pubblico per consentire all'individuo di agire in modo consapevole e per evitare la discrezionalità dei sorveglianti; risulta inoltre obbligatorio il raggiungimento di un obiettivo legittimo. Bisogna quindi riuscire a dimostrare che quella specifica libertà di espressione generi gravi pericoli per la sicurezza dell'intera Nazione. È infine necessaria la sussistenza di un sistema di controllo indipendente affidato ai giudici per autorizzare le misure di sorveglianza e rimediare in caso di abusi. [231]

Potremmo, dunque, a questo punto, citare Hobbes e il Leviatano fornendo un paragone tra la scienza come potere[232] e i sovrani assolutisti. L'uomo, pienamente libero, rinuncia ai diritti fondamentali che esistono nello Stato di natura al fine di vedere la sua vita maggiormente garantita nello Stato sociale. La paura di una poca tutela lo porta ad assoggettarsi al Leviatano.[233] Le società e gli esseri umani si stanno impegnando animatamente nella realizzazione di un programma informatico-tecnologico di inibizione del crimine attraverso il monitoraggio degli individui, a costo di monitorare anche se stessi e nonostante non tutte le decisioni assunte da sistemi di intelligenza artificiale siano completamente trasparenti. L'essere umano diventa, così, l'esecutore delle strutture di potere alla base di un progetto di sorveglianza che rischia di ledere la tutela della *privacy* e dei diritti inviolabili dell'uomo.

### *3.2 Digitalizzazione della giustizia e giudice robot*

In tema di intelligenza artificiale e giurisdizione non si può non parlare di digitalizzazione del processo penale. Da ultimo, è interessante trattare del progetto, da realizzarsi in un arco temporale di cinque anni più cinque, del CED della Corte di Cassazione per la valorizzazione del patrimonio conoscitivo rappresentato dal *corpus* giurisprudenziale e normativo in possesso dalla Cassazione attraverso la tecnologia del *legal analytics*. Tra le attività da delegare alla macchina rientrerebbero la predizione, l'estrazione di argomenti giuridici, la massimizzazione automatica e la ricerca automatizzata dei documenti. Tramite le tecnologie di intelligenza artificiale il *corpus* normativo verrebbe analizzato da algoritmi che giungerebbero ad esiti di giudizio semplificando la redazione delle massime giurisprudenziali. Nel rispetto di principi etici dettati dal CEPEJ[234] (di cui parleremo a breve) esso ha lo scopo di individuare strumenti razionali di classificazione, organizzazione e utilizzo dei dati rilevanti per la decisione; evitare *bias* cognitivi; avere un quadro completo della giurisprudenza per evitare conflitti di giudicato inconsapevoli.

---

[231] "*Necessary & Proportionate - International Principles on the Application of Human Rights Law to Communications Surveillance – Background and Supporting International Legal Analysis*", Article19, Electronic Frontier Foundation, eff.org, Maggio 2014: https://www.ohchr.org/Documents/Issues/Privacy/ElectronicFrontierFoundation.pdf

[232] L. D'AVACK, "*La rivoluzione tecnologica e la nuova era digitale: principi etici*", Paragrafo 1 (pag. 3 – 6) "Una premessa: il potere della scienza" in RUFFOLO U., "*Intelligenza artificiale – il diritto, i diritti, l'etica*", Collana "Tech and Law" di Giuffrè Francis Lefebvre Edizione Milano 2020; Parte I "Etica", Capitolo 1 (pag. 3 – 29).

[233] T. HOBBES, "*Leviathan or The Matter, Forme and Power of a Common Wealth Ecclesiastical and Civil",* 1651.

[234] COMMISSIONE EUROPEA PER L'EFFICIENZA DELLA GIUSTIZIA (CEPEJ), "*Carta etica per l'uso dell'intelligenza artificiale nei sistemi giudiziari e nel loro ambiente*".

Secondo la Ministra Cartabia, l'intelligenza artificiale è a supporto del lavoro dei giudici, dunque, al seminario europeo "*Digital Justice Ministerial Forum*"[235] del 12 ottobre 2021, la stessa ha sviluppato due argomenti principali: il primato della legge sulla tecnica e l'importanza del dialogo tra gli stessi. È necessario il pieno rispetto dell'indipendenza della magistratura e della garanzia fornita dal controllo umano sulle decisioni allo scopo di permettere che la transizione digitale e le riforme della giustizia vadano di pari passo. La Ministra Cartabia sottolinea la sussistenza di due fasi nell'attività giudiziaria: il conoscere e il decidere. Nella prima, l'intervento dell'intelligenza artificiale nello studio dei fatti, norme e giurisprudenza, può risultare un notevole sostegno al magistrato, data la sua imbattibile capacità di analizzare ed elaborare dati; nella seconda fase, la decisione deve rimanere sempre in possesso al giudice, l'unico capace di cogliere tutte le peculiarità e specificità del caso concreto. L'organo giudicante, nello svolgimento delle sue attività, può essere affiancato ma mai sostituito.

Tale *Ministerial Forum* è stato annunciato nella Comunicazione della Commissione del 2 dicembre 2020 sulla digitalizzazione della giustizia nell'Unione Europea.[236] È importante che il sistema giudiziario stia al passo con la trasformazione tecnologica al fine di giungere ad una sua digitalizzazione nel rispetto dei principi etici. L'obiettivo perseguito dall'Unione Europea è duplice: migliorare i sistemi giudiziari nazionali implementando l'adozione di strumenti digitali e migliorare la cooperazione giudiziaria transfrontaliera tra le autorità competenti dei vari Stati membri. Tutto ciò contempla la digitalizzazione dei servizi pubblici di giustizia, la promozione dell'uso di tecnologie di comunicazione a distanza e l'interconnessione tra banche dati e registri nazionali.[237] Infatti, dal 2013, ogni anno, la Commissione pubblica un quadro per valutare i sistemi giudiziari negli Stati membri anche in relazione alla loro digitalizzazione. Si nota un progresso ma non si è ancora giunti ad un'armonizzazione tra i sistemi dei vari Stati. È infatti possibile assistere ad un continuo utilizzo di fascicoli cartacei, alla mancanza di infrastrutture digitali adeguate, all'insussistenza di un canale di scambio affidabile tra le autorità nazionali e europee[238] e alla lentezza del tasso di adozione di strumenti tecnologici basati sull'intelligenza artificiale tra i vari Stati membri.

L'implementazione di tali sistemi è avvenuta nel campo delle *chatbot* per migliorare l'accesso alla giustizia, della traduzione automatica e dell'anonimizzazione delle decisioni degli organi giurisdizionali. Al contempo, l'opacità delle applicazioni di intelligenza artificiale risulta rischiosa per i diritti fondamentali in

---

[235] EUROPEAN COMMISSION, "*Digital Justice Ministerial Forum*", online, Belgio, del 12 ottobre 2021: https://digital-justice-ministerial-forum-2021.b2match.io/home

[236] COMMISSIONE EUROPEA "*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni – Digitalizzazione della giustizia dell'Unione Europea, Un pacchetto di opportunità*", Bruxelles 2.12.2020, COM(2020) 710 final: https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52020DC0710&from=EN

[237] COMMISSIONE EUROPEA "*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni – Digitalizzazione della giustizia dell'Unione Europea, Un pacchetto di opportunità*", Bruxelles 2.12.2020, COM(2020) 710 final; Paragrafo 1 (pag. 1 – 3) "Introduzione", op. cit.

[238] COMMISSIONE EUROPEA "*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni – Digitalizzazione della giustizia dell'Unione Europea, Un pacchetto di opportunità*", Bruxelles 2.12.2020, COM(2020) 710 final; Paragrafo 2 (pag. 3 – 5) "Sfide per i sistemi giudiziari nell'era digitale", op. cit.

caso di utilizzo in processi decisionali aventi effetti significativi sui diritti delle persone. Tutto ciò, però, non impedisce l'utilizzo dell'intelligenza artificiale come supporto al processo decisionale umano. [239]

A tale scopo è possibile notare come il Progetto UE di strategia in materia di giustizia elettronica (2019-2023)[240] punta alla digitalizzazione delle procedure e alla comunicazione elettronica tra i vari soggetti coinvolti nei procedimenti come elemento essenziale per il funzionamento efficace del sistema giudiziario degli Stati membri grazie alla dematerializzazione dei procedimenti giudiziali e stragiudiziali per offrire un accesso rapido ai tribunali e agevolare il ricorso ai procedimenti stragiudiziali attraverso l'uso di strumenti sicuri di comunicazione transfrontaliera (e-CODEX). Secondo il Piano d'azione 2019-2023 in materia di giustizia elettronica[241], l'intelligenza artificiale potrebbe svolgere un ruolo importante nel settore della giustizia grazie anche all'utilizzo di un *European Case Law Identifier* (ECLI) con lo scopo di rendere più rapida l'individuazione della giurisprudenza maggiormente rilevante agevolando il reperimento delle informazioni.

Passando ad analizzare la situazione nazionale, in Italia, nonostante una notevole parte delle attività venga svolta in modalità digitale, il processo penale telematico risulta essere in una posizione particolarmente arretrata e sperimentale. Non è ancora stato sviluppato un archivio completo delle sentenze dei giudici di merito accessibile agli interessati.[242] È possibile notare segni di ripresa grazie alla diffusione del Portale delle Notizie di Reato.; attualmente si stanno ponendo le basi per un flusso completo di informazioni dalle Procure ai Tribunali, Corti d'Appello e Procure Generali.

Nei vari progetti ministeriali significativi in ambito di digitalizzazione (come il riparto automatico nella stesura degli atti giudiziari di tutte le notizie necessarie tramite i dati dei registri generali delle cancellerie, l'archiviazione elettronica del testo integrale di tutti gli atti inerenti ai maxi-processi, la gestione automatica delle intercettazioni telefoniche e il sistema computerizzato di identificazione delle persone sottoposte alle indagini contenente i dati biometrici) non sempre è presente il tema dell'intelligenza artificiale (visibile esclusivamente nei casi di utilizzo di *big data analysis* e tecniche di *machine learning*) ma risulta comunque interessante mantenere una visione di insieme.[243] È importante, quindi, che gli operatori della giustizia siano formati sull'uso delle applicazioni di intelligenza artificiale anche in campo di decisioni giudiziarie.

Oggi, infatti, è possibile assistere ad un numero crescente di decisioni prese da macchine automatizzate, capaci di influire sulle libertà degli individui (dal settore medico alle decisioni giudiziarie passando per le auto senza guidatore e le pubblicità sul *web*). Tanto è vero che, l'articolo 4 del c.d. GDPR[244] indica come

---

[239] COMMISSIONE EUROPEA "*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni – Digitalizzazione della giustizia dell'Unione Europea, Un pacchetto di opportunità*", Bruxelles 2.12.2020, COM(2020) 710 final; Paragrafo 3 (pag. 6 – 25) "Un pacchetto di strumenti per la digitalizzazione della giustizia", op. cit.

[240] CONSIGLIO DELL'UNIONE EUROPEA, "*Progetto UE di strategia in materia di giustizia elettronica (2019-2023)*", Bruxelles 31 ottobre 2018: https://data.consilium.europa.eu/doc/document/ST-12794-2018-REV-3/it/pdf

[241] CONSIGLIO DELL'UNIONE EUROPEA, "*Piano d'azione 2019-2023 in materia di giustizia elettronica*", Bruxelles 31 ottobre 2018 http://data.consilium.europa.eu/doc/document/ST-11724-2018-REV-4/it/pdf

[242] F. DONATI, op. cit., Paragrafo 9 (pag. 435 - 436) "Considerazioni conclusive".

[243] A. SANTOSUOSSO, op. cit., Paragrafo 4 (pag. 85 - 122) "L'intelligenza artificiale e i giudici".

[244] REGOLAMENTO (UE) 2016/679 DEL PARLAMENTO EUROPEO E DEL CONSIGLIO DEL 27 APRILE 2016 relativo alla "*Protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati*" e che abroga la direttiva

profilazione ogni forma di trattamento automatizzato dei dati "*effettuato per valutare aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica*".[245]

La possibilità di decisioni giudiziarie algoritmiche in materia penale ha spinto il Consiglio d'Europa, nel dicembre 2018 a Strasburgo, ad adottare un atto di *soft law* rivolto indistintamente a pubblici o privati coinvolti nell'utilizzo dell'intelligenza artificiale, tramite la Commissione per l'efficacia della giustizia (CEPEJ, istituita nel 2002 per iniziativa del Comitato dei Ministri del Consiglio d'Europa, con lo scopo di monitorare e misurare la qualità dei sistemi giudiziari dei Paesi membri)[246], la "*Carta etica europea per l'uso dell'intelligenza artificiale nei sistemi di giustizia penale e nei relativi ambienti*"[247] (introdotto precedentemente). Si tratta del primo documento volto a stabilire criteri incentrati ad orientare le modalità di sviluppo e di impiego di sistemi di intelligenza artificiale a supporto delle decisioni giudiziali. Il suo obiettivo è quello di favorire la prevedibilità nell'applicazione della legge e l'uniformità degli orientamenti giurisprudenziali. Nonostante la stessa non sia vincolante (ma sarebbe comunque auspicabile che il legislatore la utilizzi al fine di realizzare una normativa *ad hoc* come ad esempio forme di certificazione dei *software* attestando il rispetto, da parte degli stessi, di determinati principi e garanzie)[248], è interessante notare come raccomandi l'utilizzo di sistemi di intelligenza artificiale per garantire una migliore *cyber*-sicurezza e lotta alla criminalità attraverso l'utilizzo di metodi computazionali che possano rafforzare l'efficacia della giustizia, ma, al contempo, sancendo principi generali che possono essere assimilati a dei limiti al fine di evitare lesioni dei diritti e delle libertà degli individui.[249] Possiamo sintetizzarli in cinque punti.

Il primo riguarda il rispetto dei diritti fondamentali nella fase di progettazione e applicazione dei sistemi di intelligenza artificiale. Gli operatori dovranno dunque formare la macchina al fine di evitare lesioni di diritti degli individui soprattutto legati alla giustizia (diritto di accesso, all'equo processo, al contraddittorio, alla parità delle armi, il principio di legalità e di indipendenza della magistratura).[250]

Si passa poi ad enunciare il principio di non discriminazione inteso come non creare o accentuare discriminazioni tra individui.[251] Tale principio risulta inoltre affermato anche nel GDPR che, nel Considerando

---

95/46/CE (Regolamento generale sulla protezione dei dati): https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32016R0679&from=IT

[245] A. SIMONCINI, "*L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*", in BioLaw Journal – Rivista di BioDiritto, n. 1/2019 - ISSN 2284-4503. (pag. 63-89); Paragrafo 2 (pag. 65 - 67) "Dalla «cibernetica» all'avvento dell'intelligenza «artificiale»".

[246] S. QUATTROCOLO, "*Intelligenza artificiale e giustizia: nella cornice della Carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche*", in *www.lalegislazionepenale.eu*, 18 dicembre 2018; Paragrafo 2 (pag. 3 - 4) "Gli obiettivi della Carta": http://www.lalegislazionepenale.eu/wp-content/uploads/2019/02/Carta-etica-LP-impaginato.pdf

[247] COMMISSIONE EUROPEA PER L'EFFICIENZA DELLA GIUSTIZIA (CEPEJ), "*Carta etica per l'uso dell'intelligenza artificiale nei sistemi giudiziari e nel loro ambiente*".

[248] F. DONATI, op. cit., Paragrafo 8 (pag. 433 - 435) "Norme etiche e regolazione in materia di IA".

[249] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 5 (pag. 59 – 61) "Le indicazioni dell'Unione Europea sui principi dell'i.a.".

[250] S. QUATTROCOLO, "*Intelligenza artificiale e giustizia: nella cornice della Carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche*", op. cit., Paragrafo 3.1 (pag. 4 – 5) "Il rispetto dei diritti fondamentali".

[251] S. QUATTROCOLO, "*Intelligenza artificiale e giustizia: nella cornice della Carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche*", op. cit., Paragrafo 3.2. (pag. 5 - 6) "La non discriminazione".

71 afferma che "*tenendo in considerazione le circostanze ed il contesto specifici in cui i dati personali sono trattati, è opportuno che il titolare del trattamento utilizzi procedure matematiche o statistiche appropriate per la profilazione, metta in atto misure tecniche e organizzative adeguate al fine di garantire, in particolare, che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori e al fine di garantire la sicurezza dei dati personali secondo una modalità che tenga conto dei potenziali rischi esistenti per gli interessi e i diritti dell'interessato e che impedisca tra l'altro effetti discriminatori nei confronti di persone fisiche sulla base della razza o dell'origine etnica, delle opinioni politiche, della religione o delle convinzioni personali, dell'appartenenza sindacale, dello status genetico, dello stato di salute o dell'orientamento sessuale, ovvero che comportano misure aventi tali effetti.*" [252] Tale questione è estendibile a qualsiasi forma di algoritmo predittivo, senza necessariamente limitarsi a quello della profilazione. Il rischio di *bias* nei sistemi di intelligenza artificiale e sempre dietro l'angolo giungendo ad un *output* influenzato da un pregiudizio, nel caso in cui il dato di *input* non sia neutro perché già in partenza sono discriminatorio. Per evitare che le decisioni nascano intrinsecamente distorte, bisogna rettificare i dati in ingresso. Ma, com'è evidente, è più facile a dirsi che a farsi, dato che risulterebbe necessaria la cooperazione anche di chi istruisce le macchine ed è difficile prevedere le fonti da cui le macchine potranno apprendere automaticamente immagazzinando i dati utili per le decisioni future.[253]

Il terzo principio considera la qualità e la sicurezza delle analisi dei dati e delle decisioni giudiziarie, delle fonti certificate utilizzate in un ambiente tecnologico sicuro. Bisogna quindi analizzare peculiarmente la fonte dell'informazione e l'integrità del dato ed inoltre è importante che l'intero processo possa essere verificato *ex post*. [254] I dati possono essere considerati come il petrolio del nostro secolo essendo una risorsa indispensabile per le attività giudiziarie o *extra*-giudiziarie. Maggiore è la loro quantità e qualità, maggiore sarà la possibilità di progettare sistemi di intelligenza artificiale affidabili. [255]

Successivamente viene enunciato il principio di trasparenza, imparzialità e *fairness*. Deve dunque essere garantita l'accessibilità, la comprensibilità e la verificabilità esterna dei processi computazionali realizzati nell'analisi dei dati giudiziari. Tale requisito rischia di entrare in contrasto con la tutela della proprietà intellettuale essendo, i *software* intelligenti, realizzati da società private e quindi soggetti al segreto. In merito alla giustizia penale, però, deve essere garantita la trasparenza delle decisioni essendo connessa alla libertà personale. Si pone, inoltre, la problematica legata al fatto che anche qualora fosse possibile giungere alla spiegazione del funzionamento della macchina, la comprensione della stessa sarebbe riservata esclusivamente a soggetti dotati di competenze tecniche elevate andando quindi ad escludere la maggior parte dei destinatari

[252] REGOLAMENTO (UE) 2016/679 DEL PARLAMENTO EUROPEO E DEL CONSIGLIO DEL 27 APRILE 2016 relativo alla "*Protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati*", op. cit., Considerando 71 (pag.14).
[253] A. SIMONCINI, op. cit., Paragrafo 5.3 (pag. 84 – 86) "Il principio mancante: non discriminazione".
[254] S. QUATTROCOLO, "*Intelligenza artificiale e giustizia: nella cornice della Carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche"*, op. cit., Paragrafo 3.3 (pag. 6 – 7) "Qualità e sicurezza".
[255] F. DONATI, op. cit., Paragrafo 9 (pag. 435 - 436) "Considerazioni conclusive".

della decisione realizzata anche attraverso l'utilizzo di strumenti di intelligenza artificiale.[256] I meccanismi di apprendimento come il *machine learning* utilizzano logiche di natura non deterministica, seguendo criteri di inferenza che non sono comprensibili neanche agli stessi programmatori (*black box*) perché basati su reti di apprendimento neurali o su logiche *fuzzy* che utilizzano anche criteri di parziale verità e parziale falsità. Risulta dunque complesso comprendere la sequenza argomentativa seguita dall'intelligenza artificiale; la stessa deve essere ripercorribile e verificabile in base a deduzioni legate al principio di causalità, con una logica in termini di nesso causa-effetto.[257] Sussiste in capo a ciascuno il diritto, nell'ordinamento europeo e italiano, di conoscere l'esistenza di processi decisionali automatizzati, nel pubblico o nel privato, che lo riguardino e di ricevere informazioni sulla logica utilizzata in modo tale da poter effettivamente comprenderne il procedimento. L'algoritmo, per garantire un effettivo diritto alla conoscenza e alla comprensibilità, deve risultare "razionabile" nel senso di intellegibile secondo criteri logico-razionali. Tali problematiche mettono in crisi la teoria giuridica per cui la forza normativa di una regola consiste nella sua forza persuasiva che, venendo a mancare, causerebbe difficoltà nella comprensione di quale sia il motivo per cui la stessa venga resa obbligatoria.[258] Bisogna quindi intervenire, recependo i valori del diritto costituzionale, durante la produzione degli algoritmi perché una volta realizzati, è troppo tardi per garantire una tutela effettiva. Forse, addirittura, sarebbe fondamentale intervenire quando gli scienziati sono ancora in formazione al fine di trasmettere loro le ragioni della tutela dei diritti fondamentali.[259]

Infine, viene trattato il principio di garanzia del controllo dell'utente al fine di assicurare l'agire informato degli utilizzatori esercitando il loro dominio sulle scelte effettuate riducendo i rischi di standardizzazione delle decisioni automatizzate:[260] il principio di non esclusività della decisione algoritmica automatizzata che produca effetti giuridici che incidono su una persona. Affidando ad un essere umano il compito di controllare, validare o smentire la decisione automatica.[261] Bisogna porre la nostra attenzione sul Regolamento UE 2016/679 del Parlamento Europeo e del Consiglio del 27 aprile 2016 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione dei dati e che abroga la direttiva 95/46 (GDPR)[262] il cui articolo 22 esprime il diritto dell'interessato a non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato che produca effetti giuridici che lo riguardano. Tali decisioni possono essere ammesse solo qualora necessarie per la conclusione di un contratto tra l'interessato e un titolare del trattamento, che sia autorizzata dal diritto dell'Unione o dello Stato Membro e che rispetti la tutela dei diritti e delle libertà del soggetto oltre che basarsi sul suo consenso (art.22.2 GDPR).

---

[256] S. QUATTROCOLO, "*Intelligenza artificiale e giustizia: nella cornice della Carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche"*, op. cit., Paragrafo 3.4 (pag. 7 - 9) "Trasparenza, imparzialità e fairness".
[257] A. SIMONCINI, op. cit., Paragrafo 6 (pag. 86 – 87) "Verso una dottrina della precauzione costituzionale".
[258] A. SIMONCINI, op. cit., Paragrafo 5.1 (pag.77 – 79) "Il principio di conoscibilità".
[259] A. SIMONCINI, op. cit., Paragrafo 7 (pag. 87 – 89) "Verso un diritto costituzionale ibrido".
[260] S. QUATTROCOLO, "*Intelligenza artificiale e giustizia: nella cornice della Carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche"*, op. cit., Paragrafo 3.5 (pag. 9 - 10) "Il controllo dell'utente".
[261] A. SIMONCINI, op. cit., Paragrafo 5.2 (pag. 79 – 84) "Il principio di non esclusività".
[262] REGOLAMENTO (UE) 2016/679 DEL PARLAMENTO EUROPEO E DEL CONSIGLIO DEL 27 APRILE 2016 relativo alla "*Protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati*", op. cit.

In tali casi il titolare del trattamento deve adottare misure appropriate per tutelare i diritti e le libertà dell'interessato (art.22.3 GDPR). L'espressione risulta vaga ed è soggetta a limitazioni di ordine pratico. Il Considerando 71 esprime il diritto ad ottenere una spiegazione della decisione (e quindi a che l'algoritmo sia comprensibile secondo principi di razionabilità ordinari). Tale principio può essere connesso all'articolo 41 della Carta dei Diritti Fondamentali dell'Unione Europea che impone di fornire le ragioni della decisione quando la stessa rischi di ledere un diritto.[263]

Sussistono forti dubbi sull'effettività intrinseca di tale principio dato che, nella pratica, le decisioni prese dagli algoritmi, dipendono da differenti fattori. Chiedere al titolare del diritto leso di dimostrare che la decisione lesiva sia stata presa unicamente sulla base di un algoritmo consisterebbe in una *probatio diabolica*. Forse, quindi, bisognerebbe garantire che il decisore umano sia in grado di esprimere una propria motivazione che giustifichi l'adesione alla valutazione effettuata dall'algoritmo.

A seguito dell'analisi di tali principi bisogna sottolineare come gli algoritmi di intelligenza artificiale vengono per lo più utilizzati in campo di giustizia civile e amministrativa. Per quanto concerne i procedimenti penali il loro impiego potrebbe risultare problematico per quattro ordini di motivi: la limitata capacità di un *robot* intelligente a valutare la reticenza, l'onestà o la falsità di un testimone (mezzo di prova più utilizzato nel procedimento penale); la difficoltà per una macchina di comprendere se determinati indizi possano essere considerati gravi, precisi e concordanti (*ex* art.192 c.p.p.); la non sensibilità posseduta dall'algoritmo[264] al fine di effettuare la valutazione della gravità del reato a cui deve essere connessa la pena (art.133 c.p.) nella quale rientrano anche le analisi in tema di intensità del dolo, grado della colpa, capacità, motivi a delinquere e carattere del *reo*; e la quasi impossibilità per un algoritmo di applicare in modo coerente la regola di giudizio *"dell'oltre ogni ragionevole dubbio*" (*ex* art.533 c.p.p.) dato che la stessa non risponde a logiche binarie o probabilistiche ma esclusivamente basate su fattori umani.[265]

Un'ulteriore problematica risulta essere dovuta al fatto che l'utilizzo di algoritmi di intelligenza artificiale nei sistemi giudiziari venga rimesso principalmente alla volontà dei singoli operatori della giustizia e individui; per il momento non si assiste ad un uso quotidiano e ad un'applicazione su larga scala negli Stati Membri del Consiglio d'Europa. La Carta, elenca le diverse modalità di utilizzazione della macchina catalogandole in base al rischio: incentivare i sistemi di sostegno all'accesso alla giustizia e alla ricerca (algoritmi di *machine learning* per analizzare la giurisprudenza utile per i casi concreti e le *chatbots* per agevolare i singoli individui che si interfacciano con il mondo giuridico); utilizzare con limiti metodologici gli strumenti *on-line* di supporto alla soluzione di controversie o per le investigazioni predittive in ambito penale; attendere ulteriori studi scientifici prima di adoperare gli algoritmi di previsione delle decisioni da parte dei giudici; limitare completamente o gestire con cautela estrema la profilazione automatica degli individui in procedimenti penali in quanto potrebbe causare il rischio di esiti discriminatori e di

[263] A. SIMONCINI, op. cit., Paragrafo 5.2 (pag. 79 – 84) "Il principio di non esclusività".
[264] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 6 (pag. 61 – 63) "La valutazione della prova penale e della responsabilità".
[265] F. BASILE, op. cit., Paragrafo 4 (pag. 14 – 16) "Secondo percorso d'indagine - IA e decisione giudiziaria: la macchina-giudice?".

cristallizzazione dei precedenti in capo ad un soggetto. Si può quindi notare un approccio cauto all'integrazione dell'intelligenza artificiale nei procedimenti giudiziari al fine di bilanciare tutti gli interessi in gioco.[266]

A tale tema risulta fortemente connessa la sostituzione della macchina all'uomo nel processo giurisdizionale. Bisogna tener conto che la decisione giuridica non discende direttamente da un mero confronto di elementi di fatto e di diritto, ma dipende da criteri logici, culturali, sociali e politici peculiari e non facilmente analizzabili.

Citando Reed C. Lawlor, nel 1963: "*Verrà un giorno in cui si sarà in grado di inserire un insieme di dati in una macchina che ha al suo interno precedenti, regole di diritto e regole di ragionamento e in cui la macchina sarà capace di offrire passo dopo passo il ragionamento attraverso il quale si può essere in grado di arrivare a una decisione. Noi potremmo studiarlo e decidere se la macchina ha proposto qualcosa di giusto o di sbagliato. In alcuni casi la macchina non dirà quale potrebbe essere la soluzione, ma vi è una probabilità che la risposta sia corretta, e questa probabilità è del 90%*"[267] In realtà, però, non ci si può basare esclusivamente sul sillogismo aristotelico: la norma vieta il comportamento tenuto da un soggetto e quindi quel soggetto è colpevole. La decisione penale possiede un qualcosa di intrinsecamente creativo che sembra inconciliabile con l'intelligenza artificiale capace di risolvere compiti complessi per quanto riguarda le risorse computazionali necessarie ma incapace di gestire situazioni in cui si necessita di una quantità eterogenea di abilità e competenze.[268] Riprendendo, ma analizzando più precisamente quando accennato dalla Carta etica[269], il giudice quando determina il peso di una prova per affermare la responsabilità dell'imputato, non si basa esclusivamente su giudizi intuitivi dato che la valuta anche in relazione a regole normative e massime di comune esperienza senza limitarsi ad una mera scelta tra vero o falso. Ed è proprio qui che può essere apprezzato il valore aggiunto del giudizio realizzato da un essere umano.

I sistemi di intelligenza artificiale hanno anche la capacità di sostenere l'organo giudicante nella fase di ricerca dei precedenti giurisprudenziali in modo tale da valutare se avviare o meno un processo. L'impego degli strumenti di intelligenza artificiale risulta utile anche nella fase di filtro di ammissibilità delle impugnazioni valutando la ragionevole probabilità di essere accolta. L'obiettivo: rendere più rapido il lavoro dell'organo giudicante sulla base dei principi costituzionali.

Si pone, quindi, il tema dell'applicabilità automatica della legge. Le leggi non sono un algoritmo e la macchina, come accennato precedentemente non può interferire con il potere discrezionale del giudice. L'Illuminismo afferma che la legge di per sé deve essere considerata come sovrana, incondizionata, completa, inequivoca, universale e autosufficiente; su queste basi si fonda la tripartizione dei poteri nello Stato

---

[266] A. SANTOSUOSSO, op. cit., Paragrafo 4 (pag. 85 - 122) "L'intelligenza artificiale e i giudici".
[267] R. C. LAWLOR, 1963, International Association for Artificial Intelligence, "*A Manifesto for Artificial Intelligence in the Law*", slide 12, Professor R. Susskind, 14 giugno 2017: http://www.iaail.org/?q=page/keynote-speeches-icail
[268] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 2 (pag. 49 – 51) "Intelligenza artificiale: di cosa stiamo parlando?".
[269] COMMISSIONE EUROPEA PER L'EFFICIENZA DELLA GIUSTIZIA (CEPEJ), "*Carta etica per l'uso dell'intelligenza artificiale nei sistemi giudiziari e nel loro ambiente*".

costituzionale facendo in modo che essi non si influenzino e non ledano l'applicazione pura della legge. Il giudice nulla è se non la *bouche de la lois*.[270] Al contempo però è necessaria l'interpretazione della stessa al fine di risolvere casi concreti: negarlo significherebbe entrare in un'utopia errata dato che sussistono un'infinità di situazioni che non sempre (anzi, quasi mai) possono essere interamente coperte da leggi senza la necessità di un minimo intervento interpretativo da parte del giudice al fine di colmare lacune[271] o contraddizioni.[272] Per compiere tale opera di integrazione occorre sensibilità psicologica e sociale, equilibrio e buonsenso: è necessario possedere il senso di "Giustizia". Tali sensazioni e requisiti non possono essere contenuti all'interno di un algoritmo e dunque risultano di ostacolo all'applicazione automatica della legge. Allo stesso tempo bisogna sottolineare come una legge debba essere applicata senza eccessivi spazi di discrezionalità del giudice al fine di non ledere gli individui durante l'accertamento dei fatti e a seguito della decisione: le leggi vanno interpretate nel senso di completarle e chiarirle.[273] Nascono quindi i sistemi esperti legali (SEL) connessione tra logica, informatica e diritto, essi riescono a garantire un'applicazione della legge che sia completa, inequivoca e analitica grazie all'integrazione con la giurisprudenza della Corte di Cassazione ma al contempo rischiano di trasformare il sistema giudiziario italiano da *civil law* a *common law*. Alcuni esempi in Italia sono "*Remida*" di *Giuffrè* in materia di rivalutazione monetaria e calcolo di interessi e i SEL della Pretura di Borgomanero che permettono di emettere in modo automatico i decreti penali di condanna.[274]

I nuovi sistemi di intelligenza artificiale tutelano la certezza del diritto grazie alla calcolabilità dello stesso. È necessario che ciò avvenga entro i limiti del diritto costituzionale e dei principi etici. Sarebbe un errore non sfruttare le potenzialità dell'innovazione tecnologica per paura di possibili lesioni ma è necessario un controllo finalizzato al miglioramento della qualità della giustizia. [275] Al contempo, com'è stato fino ad ora esaminato, tale utilizzo solleva numerosi interrogativi in campo alla tutela del singolo soprattutto per quanto concerne la giustizia predittiva di cui ci accingiamo a trattare.[276]

### *3.3 Giustizia Predittiva: Valutazione della pericolosità*

In ipotesi risulterebbe preferibile essere giudicati da un giudice persona fisica o dal c.d. giudice *robot*? Chi avrebbe maggiori capacità di giudicare equamente?

[270] MONTESQUIEU, "*De l'espirit des lois*", 1748.
[271] R. BORRUSO, "*Informatica Giuridica*" voce dell'Enciclopedia del Diritto – I Aggiornamento – Giuffrè, 1994; Sez. II – "Legge, giudice e computer", Paragrafo 22 (pag. 18) "I giudizi di valore delegati dal legislatore al giudice".
[272] R. BORRUSO, op. cit., Paragrafo 18 (pag. 15 – 16) "L'applicabilità automatica della legge".
[273] R. BORRUSO, op. cit., Paragrafo 29 (pag. 21 – 22) "L'interpretazione della legge alla luce della «prova-computer»".
[274] R. BORRUSO, op. cit., Paragrafo 31 (pag. 23) "La nascita dei sistemi esperti legali (SEL)".
[275] F. DONATI, op. cit., Paragrafo 9 (pag. 435 - 436) "Considerazioni conclusive".
[276] F. DONATI, op. cit., Paragrafo 1 (pag. 415 - 418) "Premessa".

Quest'elaborato non mira a fornire una risposta definitiva e univoca; ha, invece, lo scopo di sottolineare il dubbio e far riflettere. Non si tratta del fulcro centrale della tesi ma risulta una problematica interessante su cui soffermare la nostra attenzione.

Per farlo, è necessario porre l'attenzione su un caso avvenuto in negli Stati Uniti in una cittadina del Wisconsin chiamata La Crosse, vicino al Mississippi. La vicenda è nota come sentenza Loomis.[277] Il fatto riguardava una sparatoria notturna avvenuta tra alcuni soggetti all'interno di un'autovettura contro le finestre di un'abitazione dove si sarebbe trovato il *boss* della *gang* criminale antagonista. Da tale evento fu instaurato un processo per direttissima davanti alla Corte di La Crosse nel febbraio 2013. Il signor Eric Loomis, cittadino americano già più volte condannato, era stato accusato di cinque reati per la partecipazione come conducente alla sparatoria ("*drive-by-shooting*") e si era dichiarato colpevole per i due meno gravi ("*eluding an officer"* e "*operating a vehicle without its owner's consent"*). La Corte accolse la sua richiesta di patteggiamento subordinandola alla lettura in aula delle altre imputazioni. Il calcolo della pena da applicare fu realizzato grazie al PSI (*Presentence Investigation Report*) che comprendeva anche la valutazione effettuata dal *software*, basato su applicazioni di intelligenza artificiale, "*COMPAS*" (*Correctional Offender Management Profiling for Alternative Sanctions*), un sistema intelligente capace di prevedere (algoritmo predittivo) il rischio di recidiva degli imputati e appartenente ad una società privata di nome "*Northpointe*" (oggi "*Equivalent*").[278] Il sistema funziona assegnando un punteggio da uno a dieci in tre categorie di rischi: recidiva *pre*-sentenza definitiva, recidiva *post*-sentenza definitiva, recidiva in atti violenti. A Eric Loomis fu attribuito il punteggio massimo in tutti e tre: soggetto ad alto rischio di recidiva[279] (*a high risk of violence, high risk of recidivism, high pretrial risk*). Il risultato fornito dalla macchina veniva accompagnato da un documento contenente le finalità di "*COMPAS*" (identificare i trasgressori e identificare i fattori di rischio) raccomandando di non utilizzarlo al fine di determinare la gravità della pena dato che il sistema fu sviluppato con l'obiettivo di aiutare i direttori di istituti penitenziari nella distribuzione della popolazione carceraria. Nonostante ciò, la Corte della contea di La Crosse, specificando, nella sentenza, che la decisione era stata basata sul sistema "*COMPAS*", condannò Loomis a quattro anni per la prima imputazione e a sette per la seconda.

Il signor Loomis si rivolse alla Corte di Appello (che rinviò la questione alla Corte suprema dello Stato del Wisconsin) lamentando la lesione del diritto costituzionale al giusto processo leso dal giudice della Contea che utilizzò impropriamente "*COMPAS*" al fine di quantificare la pena determinando una sanzione più severa nei suoi confronti,[280] senza neanche permettere all'imputato di conoscere le ragioni della propria sentenza a causa del segreto industriale (*trade secret*) gravante sul codice sorgente di "*COMPAS*" che limitava la

[277] STATE V. LOOMIS, 881 N.W.2d 749 (2016) 754 (USA).
[278] A. LONGO E G. SCORZA, op. cit., Paragrafo 3 (pag. 88 - 147) "Vantaggi e rischi non economici dell'AI - I rischi etici, sociali e politici dell'AI - Algoritmi che discriminano".
[279] A. LONGO E G. SCORZA, op. cit., Paragrafo 7 (pag. 198 - 241) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela - Alcuni ambiti di riflessione giuridica: l'intelligenza artificiale applicata alla giustizia".
[280] G. CONTISSA, G. LASAGNI, G. SARTOR, op. cit., Paragrafo 3 (pag. 622 – 624) "Valutazione automatizzata del rischio: quali limiti di utilizzo?".

possibilità di comprensione piena delle modalità di funzionamento della macchina. [281] Il ricorrente sottolineava, inoltre, che l'analisi svolta dall'algoritmo risultava essere viziata da pregiudizi basati sul genere e sulla razza a seguito di un uso dati statistici generali e non legati al singolo individuo.[282]

La Corte Suprema del Winsconsin respinse le argomentazioni e confermò la sentenza della Corte di La Crosse con una motivazione basata sul fatto che, il *software* "*COMPAS*", se usato correttamente, non viola il diritto dell'imputato al giusto processo, anche perché la Corte della Contea aveva ribadito come la sua valutazione fosse supportata anche da altri fattori indipendenti, quindi il suo uso non era stato determinante nel decidere il caso. L'algoritmo predittivo rimane solo uno strumento disponibile per il giudice che può liberamente basarsi sull'interezza del risultato fornitogli o solo su alcune su parti di esso, nel pieno della sua discrezionalità limitata dalle caratteristiche e risultati emersi dal caso concreto. Inoltre, a causa del *trade secret*, la Corte Suprema del Winsconsin non poté divulgare le informazioni legate alle modalità di funzionamento dell'algoritmo[283] e nonostante ciò, l'imputato e il giudice possedevano le stesse informazioni sul *software*. Si precisa che il magistrato era stato correttamente informato sulla fallibilità dello strumento e sulla sua funzione di mero ausiliario. In più, nonostante il *trade secret*, l'imputato possedeva comunque la facoltà, attraverso l'utilizzo del manuale d'uso di "*COMPAS*" di analizzare i dati di *input* e le valutazioni di *output* al fine di comprendere l'attendibilità del sistema.[284] Dunque il soggetto non ha diritto a conoscere il funzionamento della decisione algoritmica dato che "*COMPAS*" non è un *software* vincolante per il giudice e quindi la sentenza è stata realizzata attraverso le valutazioni che il giudice persona fisica ha espresso nella motivazione tenendo anche conto, insieme agli altri elementi e prove, del ragionamento automatico dato dal sistema.[285]

Nonostante ciò, però, la Corte Suprema del Wisconsin formula un *warning* per il futuro uso di "*COMPAS*" al fine di garantire un utilizzo calibrato degli strumenti di giustizia predittiva,[286] sottolineando le criticità del *trade secret*; la base collettiva e non individuale delle valutazioni; e il pericolo di una sovrastima del rischio di recidiva a carico di talune minoranze etniche. Proprio per questo, si necessita di un controllo umano sulle valutazioni di "*COMPAS*".[287]

Giunto a tale risultato, Loomis tentò il ricorso alla Corte Suprema degli Stati Uniti, per violazione del V e XIV Emendamento, che venne però rigettato. La Corte affermò che l'uso di strumenti di valutazione automatica del rischio può generare lesioni della Costituzione ma che nel caso di specie fu rispettato il principio della non esclusività della decisione algoritmica.

---

[281] A. LONGO E G. SCORZA, op. cit., Paragrafo 7 (pag. 198 - 241) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela - Alcuni ambiti di riflessione giuridica: l'intelligenza artificiale applicata alla giustizia".
[282] F. BASILE, op. cit., Paragrafo 5.3.2.1 (pag. 21 – 22) "In particolare il caso Loomis e il controverso uso di COMPAS in sede di sentencing".
[283] A. SIMONCINI, op. cit., Paragrafo 3.2.1 (pag. 71 – 73) "Caso Compas".
[284] F. DONATI, op. cit., Paragrafo 3 (pag. 421 - 423) "Le discriminazioni algoritmiche".
[285] A. LONGO E G. SCORZA, op. cit., Paragrafo 7 (pag. 198 - 241) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela - Alcuni ambiti di riflessione giuridica: l'intelligenza artificiale applicata alla giustizia".
[286] S. QUATTROCOLO, "*Intelligenza artificiale e giustizia: nella cornice della Carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche*", op. cit., Paragrafo 3.3 (pag. 6 – 7) "Qualità e sicurezza".
[287] F. BASILE, op. cit., Paragrafo 5.3.2.1 (pag. 21 – 22) "In particolare il caso Loomis e il controverso uso di COMPAS in sede di sentencing".

Non sussistono, secondo le Corti, lesioni del diritto al giusto processo.[288] Il motivo di una simile conclusione può forse essere ricercato nel timore di svuotamento delle carceri dovuto al fatto che "*COMPAS*", o sistemi similari, venivano già da tempo utilizzati in sede di processo penale.

Il 23 maggio 2016 "*ProPublica*", organizzazione non governativa, realizzò un'inchiesta intitolata "*Machine bias*" pubblicando due foto (due volti): a sinistra un uomo di colore (Bernard Parker valutato ad alto rischio) e a destra un uomo bianco (Dylan Fugett valutato a basso rischio). Il significato di tale campagna di sensibilizzazione era la discriminazione razziale generata da "*COMPAS*" dovuta al fatto che, a parità di condizioni, veniva concesso un trattamento peggiore a soggetti appartenenti a minoranze etniche perché considerati a più alto rischio di recidiva. Il *report* si basava su dati di settemila persone arrestate tra il 2013 e il 2014 nella contea di Broward, Florida, in California. Due anni dopo furono valutate le previsioni fatte all'epoca dall'algoritmo per verificarne il grado di attendibilità: solo il 20% delle persone che secondo "*COMPAS*" avrebbero commesso nuovi reati violenti, li avevano posti in essere effettivamente e inoltre, il numero dei soggetti di colore (sovrastimati) considerati ad alto rischio era il doppio rispetto ai bianchi (sottostimati).[289] Le previsioni sono dunque inaffidabili e distorte.[290]

A questo punto sorge spontaneo porsi una domanda: come funziona "*COMPAS*"? Quali sono i dati che va ad analizzare per giungere ai suoi risultati?

Come accennavamo, "*COMPAS*" è l'algoritmo predittivo più utilizzato negli Stati Uniti, sviluppato nel 1998 e aggiornato più volte nel corso degli anni al fine di essere al passo con l'evoluzione tecnologica. Secondo il manuale di utilizzo, esso si basa su fattori che empiricamente denotano un'influenza sulla recidiva valutati a seguito di ricerca scientifica svolta presso gli istituti penitenziari o le apposite agenzie. L'algoritmo analizza i risultati di ricerche statistiche sui dati dei condannati, sui fattori di rischio presenti e sulla recidiva,[291] valutando, inoltre, i precedenti giudiziari, un questionario compilato dall'imputato e altri variabili coperti da segreto industriale. [292]

Il questionario, compilato dall'interessato o tramite ricerca nei registri a disposizione delle procure, comprende la risposta a centotrentasette domande riguardanti principalmente: precedenti criminali, illeciti e infrazioni; atti di violenza attuali o passati; contesto sociale e familiare criminale; abuso di sostanze; situazione economica, sociale e professionale; istruzione e formazione; isolamento sociale; instabilità residenziale; opportunità criminali; pensieri pro-criminali; attività svolte nel tempo libero; personalità criminale. [293] I fattori di rischio possono avere impatti differenti sulla possibile recidiva e vengono catalogati in tre gruppi: fattori statici (il genere o l'origine etnica); fattori dinamici, modificabili grazie al trattamento terapeutico (i pensieri

[288] A. SIMONCINI, op. cit., Paragrafo 3.2.1 (pag. 71 – 73) "Caso Compas".
[289] A. LONGO E G. SCORZA, op. cit., Paragrafo 7 (pag. 198 - 241) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela - Alcuni ambiti di riflessione giuridica: l'intelligenza artificiale applicata alla giustizia".
[290] F. DONATI, op. cit., Paragrafo 3 (pag. 421 - 423) "Le discriminazioni algoritmiche".
[291] A. SANTOSUOSSO, op. cit., Paragrafo 4 (pag. 85 - 122) "L'intelligenza artificiale e i giudici".
[292] A. SIMONCINI, op. cit., Paragrafo 3.2.1 (pag. 71 – 73) "Caso Compas".
[293] F. BASILE, op. cit., Paragrafo 5.3.2 (pag. 19 – 21) "COMPAS - Correctional Offender Management Profiling for Alternative Sanctions".

pro-criminali); fattori di rischio acuti, cambiano rapidamente e associati a determinati contesti o condizioni che facilitano le azioni violente (uso di stupefacenti).[294]

Nonostante non venga espressamente menzionata la razza, le critiche (legate alla validità predittiva e all'imparzialità) mosse a tale sistema (come quelle realizzate da "*ProPublica*") vanno a provare proprio una discriminazione di tale tipo.[295]

Inoltre, le valutazioni effettuate dal sistema si basano su tre fattori: età della persona alla sua prima infrazione, età attuale e storia criminale.[296] Il giudice, inserendo i dati a sua disposizione, riceverà una risposta per ciascun imputato in base ad una classificazione delle varie categorie di rischio (da molto basso a molto alto).[297] È importante che il sistema venga costantemente monitorato e che il giudice abbia la capacità di giungere ad una soluzione diversa proprio al fine di evitare gli effetti lesivi dei *bias*.[298]

"*COMPAS*" risulta essere un sistema non trasparente ma nonostante le critiche riportategli ha offerto, in molti casi, buoni risultati al fine di rendere più oggettive, veloci e non discriminatorie le decisioni sulla pericolosità del condannato. Allo stesso tempo, il rischio di discriminazioni, errori o pregiudizi che possono influenzare l'esito del processo decisorio affidato all'intelligenza artificiale non è da sottovalutare dato che possono essere addirittura intensificati dall'utilizzo di tali tecnologie. Il buon funzionamento del sistema può essere leso da problemi di apprendimento, dalla gestione non corretta dei dati nella fase in cui il *software* elabora i propri modelli decisioni che successivamente compromette l'attendibilità degli *output*, oppure da una discriminazione statistica, ossia condizionamenti determinati da rilevazioni precedentemente effettuate.[299]

Il processo decisionale basato sui dati pone tre ordini di problemi: i *bias* nel *set* di dati (la qualità, la rappresentatività e le caratteristiche degli stessi nella raccolta e selezione), i *bias* negli algoritmi applicati o negli obiettivi selezionati e la spiegabilità dei risultati a cui l'intelligenza artificiale è giunta. La macchina replica ciò che vede e apprende dalla società.

I sistemi sono progettati e addestrati da esseri umani, dunque essendo l'algoritmo strutturalmente condizionato dall'insieme dei valori posseduti da chi lo realizza, è possibile che vada ad incorporare nostri pregiudizi automatizzando la discriminazione (*implicit bias*) e generando danni impossibili da individuare e correggere immediatamente. Un circolo vizioso: *bias in, bias out*.

Dunque un algoritmo non dovrebbe calcolare quanto sia pericoloso un imputato ma quale sanzioni altri giudici imporrebbero in un caso analogo, dando così un margine di concretezza all'indagine potendo verificare la pena individuata sulla base di situazioni similari prima ancora che la stessa diventi vincolante.[300]

---

[294] F. BASILE, op. cit., Paragrafo 5.2 (pag. 17 – 18) "La valutazione "attuariale" della pericolosità criminale".
[295] F. BASILE, op. cit., Paragrafo 5.3.2 (pag. 19 – 21) "COMPAS - Correctional Offender Management Profiling for Alternative Sanctions".
[296] A. LONGO E G. SCORZA, op. cit., Paragrafo 7 (pag. 198 - 241) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela - Alcuni ambiti di riflessione giuridica: l'intelligenza artificiale applicata alla giustizia".
[297] A. SIMONCINI, op. cit., Paragrafo 3.2.1 (pag. 71 – 73) "Caso Compas".
[298] A. SANTOSUOSSO, op. cit., Paragrafo 4 (pag. 85 - 122) "L'intelligenza artificiale e i giudici".
[299] F. DONATI, op. cit., Paragrafo 3 (pag. 421 - 423) "Le discriminazioni algoritmiche".
[300] C. BURCHARD, op. cit., Paragrafo 4 (pag. 19 – 26) "La trattazione teorico-penale delle promesse dell'IA".

La problematica aggiuntiva riguarda il fatto di basarsi esclusivamente su rappresentazioni del passato e strettamente connesse a denunce sporte o dati forniti dalla polizia; di conseguenza, se le forze dell'ordine si impegnano a controllare maggiormente una determinata zona o categoria di soggetti, è scontato che i dati negativi legati a quel settore di indagine saranno maggiori generando una discriminazione strutturale:[301] guardando al passato come guida per il futuro proiettando in avanti le disuguaglianze.[302] Si giunge a ciò che viene denominato come "polarizzazione del giudizio", null'altro è se non una stigmatizzazione dei soggetti. La discriminazione, intesa come lesione del principio di uguaglianza, verrebbe così delegata agli algoritmi che andrebbero dunque a "giustificarla" attraverso dati statistici generali legati all'appartenenza ad un gruppo (*group-based generalizations*)[303]e non basati sulle caratteristiche del singolo soggetto: è necessario distinguere per non discriminare.

Proprio per questo, nel 2018, oltre cento organizzazioni per i diritti civili hanno firmato una dichiarazione con lo scopo di sensibilizzare e richiedere uno *stop* all'utilizzo di sistemi di valutazione del rischio in ambito giudiziario.[304]

"*COMPAS*" risulta, dunque, un ottimo strumento di supporto alla giustizia ma sono necessari due requisiti: la garanzia del suo corretto funzionamento al fine di evitare di limitare la libertà a soggetti che non lo meritano e viceversa;[305] e il rispetto della piena autonomia di giudizio in capo al giudice che non deve basare la sua decisione esclusivamente sulle risultanze provenienti dalla macchina.[306]

È dunque necessario valutare l'affidabilità del *software* utilizzato. Grazie al famoso caso Daubert (1993),[307] la Corte Suprema dichiarò la necessità di tale valutazione per l'ammissibilità delle prove scientifiche adottate in processo e vennero stabiliti i criteri Daubert (non analizzati nel caso Loomis). Essi riguardano: la verificabilità empirica della tecnica prospettata dall'esperto; la pubblicazione e la *peer review*; il tasso di errore noto o potenziale nell'applicazione della tecnica; il rispetto di *standard* e controlli; l'opinione espressa dalla comunità scientifica in relazione a detta teoria.[308] Il giudice dovrebbe agire (in contraddittorio) come rasoio di Occam eliminando gli approcci scientifici inaffidabili per quanto concerne l'attendibilità del perito, del metodo scientifico utilizzato e della sua corretta applicazione in relazione al caso concreto.[309] Comunque, la Corte Suprema non si è ancora pronunciata sull'applicabilità di detti criteri per i sistemi di intelligenza artificiale.[310]

---

[301] A. LONGO E G. SCORZA, op. cit., Paragrafo 3 (pag. 88 - 147) "Vantaggi e rischi non economici dell'AI - I rischi etici, sociali e politici dell'AI - Algoritmi che discriminano".
[302] A. M. MAUGERI, op. cit., Paragrafo 3 (pag. 12 - 19) "I rischi connessi all'uso di algoritmi predittivi: discriminazioni e valutazioni generalizzanti".
[303] *Idem.*
[304] A. LONGO E G. SCORZA, op. cit., Paragrafo 3 (pag. 88 - 147) "Vantaggi e rischi non economici dell'AI - I rischi etici, sociali e politici dell'AI - Algoritmi che discriminano".
[305] A. LONGO E G. SCORZA, op. cit., Paragrafo 7 (pag. 198 - 241) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela - Alcuni ambiti di riflessione giuridica: l'intelligenza artificiale applicata alla giustizia".
[306] F. DONATI, op. cit., Paragrafo 7 (pag. 430 - 433) "L'IA come ausilio per il giudicante".
[307] DAUBERT V. MERRELL DOW PHARMACEUTICALS, INC., 113 S.Ct. 2786(1993).
[308] A. M. MAUGERI, op. cit., Paragrafo 5 (pag. 23 - 27) "L'adozione del modello "Daubert" per verificare la scientificità della teoria psico-criminologica e del software".
[309] Sentenza della Cassazione Penale, Sez. IV, n. 43786 del 17 settembre 2010 (Cozzini).
[310] G. CONTISSA, G. LASAGNI, G. SARTOR, op. cit., Paragrafo 3 (pag. 622 – 624) "Valutazione automatizzata del rischio: quali limiti di utilizzo?".

Anche la dottrina italiana ritiene utili i criteri Daubert dichiarando che il giudice deve assolvere un ruolo critico e non di mero spettatore difronte alle prospettazioni scientifiche e individuando alcuni indici sintomatici dell'affidabilità dell'approccio tecnico-scientifico: qualifica professionale e indipendenza del perito; valutazione dello stato complessivo delle conoscenze accreditate; grado di consenso dell'approccio all'interno della comunità scientifica; autorevolezza e indipendenza dell'autore della ricerca scientifica e le sue finalità.[311]

Com'è possibile immaginare, "*COMPAS*" non è l'unico algoritmo predittivo in commercio, programmi analoghi sono stati utilizzati in importanti decisioni come nello Stato del New Jersey dove sono state sostituite le udienze per la concessione della libertà su cauzione con delle valutazioni di rischio ottenute tramite macchine intelligenti[312] e come il *software* "*SAVRY*" (*Structured Assessment of Violence Risk in Youth*) sussistente in un algoritmo di valutazione del rischio di comportamenti violenti nei minori. Nel 2010 nella decisione Malenchik la Corte d'appello dell'Indiana[313] si era basata sull'analisi del rischio effettuata dai sistemi "*LSI-R*" (*Level of Service Inventory-Revised*) e "*SASSI*" (*Substance Abuse Substance Subtle Screening Inventory*) ed anche qui, l'imputato lamentò una lesione del giusto processo a causa di valutazioni discriminatorie effettuate dai *software* nonostante, l'uso da parte del giudice fu valutato legittimo dato che risultò come mero strumento di supporto, affiancato, ai fini della decisione, ad altri elementi indipendenti.[314]

Il secondo interrogativo da porsi riguarda la costituzionalità di tali decisioni prese sulla base di un algoritmo, di natura statistico-probabilistica, non trasparente, che può commettere errori che causano lesioni del principio di uguaglianza e che possiede una grande forza pratica attrattiva rischiando di giungere a decisioni prive di una motivazione effettiva lesiva del diritto di difesa. È costituzionalmente accettabile che un giudice condanni una persona sulla base di valutazioni elaborate dalla macchina di cui i protagonisti del processo non conoscono il funzionamento?

In base a quanto descritto sopra in relazione al caso Loomis e all'uso di tecnologie di intelligenza artificiale nel processo decisionale è facile notare i rischi di lesione del diritto all'equo processo, della presunzione di innocenza e del diritto di difesa dell'imputato.[315] Il condannato verrebbe privato della libertà in base ad un'elaborazione algoritmica e non a seguito di una decisione umana senza, per giunta, poter accede al codice sorgente. È lecito obiettare che, spesso, anche una motivazione realizzata da un giudice persona fisica non risulta trasparente e imparziale.[316]

Effettuando un paragone con l'ordinamento Europeo, è possibile osservare che i principi generali relativi al trattamento dei dati personali (articolo 5 del c.d. GDPR) si applicano anche nell'ambito dei sistemi di intelligenza artificiale.

---

[311] Sentenza della Cassazione Penale, Sez. IV, n. 43786 del 17 settembre 2010 (Cozzini) § 14.
[312] V. MANES, op. cit., Paragrafo 3 (pag. 9 – 10) "L'esperienza d'Oltreoceano".
[313] MALENCHIK V. STATE, 928 N.E.2d 564, 574 (Ind. 2010).
[314] G. CONTISSA, G. LASAGNI, G. SARTOR, op. cit., Paragrafo 3 (pag. 622 – 624) "Valutazione automatizzata del rischio: quali limiti di utilizzo?".
[315] *Idem.*
[316] A. LONGO E G. SCORZA, op. cit., Paragrafo 7 (pag. 198 - 241) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela - Alcuni ambiti di riflessione giuridica: l'intelligenza artificiale applicata alla giustizia".

All'articolo 15 del GDPR viene sancito il diritto dell'interessato a ricevere informazioni sulla logica utilizzata dalla macchina utilizzata nel processo decisionale automatizzato. L'articolo 22 del GDPR, in più, (come spiegato precedentemente) permette la possibilità, per l'interessato, di opporsi alle decisioni basate unicamente (il cui significato letterale del termine risulta dibattuto in dottrina sul se escludere l'applicazione della norma in caso di intervento significativo o marginale dell'essere umano) sul trattamento automatizzato e che abbiano la capacità di produrre effetti nella sua sfera giuridica. Come sottolineato anche nel caso Loomis, è necessario che il giudice non si basi esclusivamente sulla valutazione algoritmica ma che effettui un bilanciamento della stessa alla luce di altri fattori.

La struttura stessa del *software* basato su sistemi di intelligenza artificiale risulta di difficile compatibilità con principi a causa della forte autonomia di analisi e apprendimento (*machine learning*) posseduta dalla macchina (riuscendo a svolgere azioni anche non inizialmente programmate tra le scelte possibili): torniamo al problema più volte toccato del *black box*. È vero anche che spesso le decisioni umane vengono adottate su impulsi complessi da far rientrare in una logica razionale, ma ciò non giustifica l'utilizzo di decisioni algoritmiche insindacabili e non trasparenti.[317] Non si ha la piena conoscenza di come si connettano le varie informazioni presenti nel calcolo e relative al caso concreto, a maggior ragione nel caso in cui l'algoritmo sia coperto dal segreto industriale venendo quindi a mancare il contraddittorio sull'ammissibilità del suo utilizzo e delle risultanze. La *black box decision* dovrebbe essere paragonata ad una decisione priva di motivazione in contrasto con gli articoli 24 e 111 della Costituzione e con l'articolo 6 della CEDU.[318] Un'efficace trasparenza dipende dalla precisione della teoria scientifica alla base, dalla chiarezza del linguaggio usato per tradurlo in formula matematica, dalla possibilità di una revisione *ex post* del passaggio da *input* ad *output*: la fondatezza empirica dell'algoritmo andrebbe valutata in contraddittorio nel rispetto del diritto di difesa.[319]

La trasparenza nella trattazione automatizzata dei dati risulta essere l'unico parametro, insieme alla legalità dell'attività, di legittimità del trattamento (art. 20 GDPR e d.lgs. 51/2018).[320] Tutto ciò spinge gli esperti del settore ad elaborare soluzioni che permettano di spiegare in modo esaustivo il funzionamento del *software,* anche senza ledere il brevetto dei codici sottostanti, oppure garantire delle possibilità di *disclosure* circoscritta al procedimento penale dei segreti commerciali relativi all'algoritmo, in modo tale da rendere sostenibile il loro utilizzo nel processo senza quindi dover giungere ad un inutilizzabilità della prova generata automaticamente perché incompatibile con l'equo processo.[321]

---

[317] F. DONATI, op. cit., Paragrafo 5 (pag. 425 - 428) "Trasparenza della decisione algoritmica e GDPR".
[318] V. MANES, op. cit., Paragrafo 5.4 (pag. 14 - 17) "Giusto processo e tutela del diritto di difesa".
[319] A. M. MAUGERI, op. cit., Paragrafo 4 (pag. 19 - 23) "La mancanza di trasparenza".
[320] D.lgs. 51/2018; "*Attuazione della direttiva (UE) 2016/680 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, nonché alla libera circolazione di tali dati e che abroga la decisione quadro 2008/977/GAI del Consiglio*": https://www.gazzettaufficiale.it/eli/id/2018/05/24/18G00080/sg
[321] S. QUATTROCOLO "*Processo penale e rivoluzione digitale: da ossimoro a endiadi?"* op. cit., Paragrafo 3 (pag. 126 – 132) "La prova generata automaticamente e i rischi per la parità delle armi".

Inoltre, la trasparenza dell'algoritmo avrebbe anche la capacità di ridurre il rischio di discriminazioni (a tutela dall'articolo 3 della Costituzione). La formazione di *bias* è inevitabile dato che risulta impossibile fornire una rappresentazione neutra dell'informazione iniziale; si può dunque neutralizzare tale problema tramite la conoscenza dei criteri utilizzati per la decisione algoritmica al fine di eliminare le discriminazioni individuate comprendendo il ragionamento effettuato dalla macchina. [322]

Tale soluzione non risulta sufficiente quando la discriminazione è insita nei dati statistici raccolti non sussistendo un errore nel funzionamento della macchina ma nelle disuguaglianze sociali presenti nella realtà: si necessita di controlli sull'*output* finale paragonato con i dati di *input* e con la situazione di fatto. Seguire generalizzazioni statistiche predittive riporterebbe in vita l'arcaico e oramai superato tipo criminologico di autore. Si passerebbe ad un diritto penale del profilo d'autore nel quale la pericolosità del soggetto verrebbe desunta solo su base statistica legata a fatti del passato, ledendo il principio dell'individualizzazione del trattamento sanzionatorio e della presunzione di innocenza sanciti agli articoli 13, 25 e 27 della Costituzione.[323]

La trasparenza, nel procedimento penale, si riflette nell'obbligo di motivazione dei provvedimenti decisionali realizzati dall'organo giudicante: il giudice deve dare conto nella sentenza della valutazione di attendibilità operata rispetto a ciascuna prova e del ragionamento logico-giuridico effettuato nell'analisi della situazione di fatto e di diritto sottoposta alla sua attenzione. Non risulta quindi sufficiente, una mera spiegazione algoritmica del funzionamento della macchina basata su sistemi di intelligenza artificiale risulterebbe priva di efficacia in concreto in quanto limitata alla comprensione di pochi esperti tecnici-informatici e oscura per il resto dell'opinione pubblica e dei destinatari della decisione.[324]

La travolgente forza pratica dell'algoritmo, inoltre rischia di condizionare fortemente la motivazione dato che, una volta introdotto un sistema di intelligenza artificiale all'interno del processo decisionale, il giudice tenderà naturalmente ad affidarsi alle risultanze del *computer* visto, erroneamente, come infallibile e sicuro. Il rischio risulta essere un'applicazione automatica di tali decisioni senza sottoporle ad un vaglio da parte del giudice in relazione al caso concreto: si giungerebbe ad una cristallizzazione della giurisprudenza.[325] Il giudice si limiterebbe a validare le risultanze date dall'algoritmo andando di fatto a far rimpiazzare il processo decisionale umano dalla macchina. Quanto evidenziato dall'intelligenza artificiale viene considerato certo e come prova scientifica risultando complessa qualsiasi tipologia di motivazione opposta prospettata dall'organo giudicante. Si vede quindi leso il principio del libero convincimento del giudice che deve essere chiarito e motivato nella sentenza. [326] Il rispetto dell'articolo 111 comma 6 della Costituzione e dell'articolo 6

[322] V. MANES, op. cit., Paragrafo 5.1 (pag. 11 – 13) "Machine bias e principio di eguaglianza".
[323] V. MANES, op. cit., Paragrafo 5.3 (pag. 13 - 14) "Valutazioni statistiche, "diritto penale del fatto" e principio di personalità della responsabilità penale".
[324] A. M. MAUGERI, op. cit., Paragrafo 4 (pag. 19 - 23) "La mancanza di trasparenza".
[325] F. DONATI, op. cit., Paragrafo 7 (pag. 430 - 433) "L'IA come ausilio per il giudicante".
[326] A. M. MAUGERI, op. cit., Paragrafo 7 (pag. 29 - 34) "Il giudizio individualizzato e il controllo del giudice sull'output".

CEDU tendono a limitare l'utilizzo dell'intelligenza artificiale nei casi in cui non sia possibile spiegare *ex post* ogni passaggio della decisione giudiziaria del caso specifico. [327]

Considerare le decisioni provenienti da sistemi di apprendimento automatico come delle totali *black box* a causa dell'incontrollabilità del meccanismo dell'algoritmo che di conseguenza comporta l'inspiegabilità della sentenza può risultare fuorviante perché in realtà ogni singola operazione svolta dalla macchina è di per sé spiegabile ma essendo in numero elevato e basandosi su interazioni non lineari non esiste un unico percorso tra *input* e *output* e dunque non è facile da ricostruire cosa accada nel sistema da parte di un essere umano. Se volessimo essere pignoli, potremmo, inoltre, affermare che una sorta di effetto *black box* sussiste anche nelle decisioni prese da un giudice persona fisica (della cui mente non è possibile conoscere a pieno il funzionamento) dato che non sempre risulta chiaro l'*iter* logico-giuridico da lui seguito (basato anche sull'intuizione, il senso di giustizia e l'esperienza che di per sé sono complessi da esplicare) ed inoltre è possibile notare casi di segreti industriali che limitano la motivazione e la conoscenza del giudice o casi di consulenze tecniche legate a contributi scientifici estremamente complessi e di difficile comprensione.[328] Quindi le stesse critiche mosse all'intelligenza artificiale sulla sua oscurità e assenza di motivazione, possono in realtà essere riscontrate anche nelle decisioni di giudici umani anch'esse influenzabili da pregiudizi. Nonostante tali similitudini le due decisioni sono percepite in modo opposto: gli errori causati da giudici persone fisiche o le loro decisioni non sempre chiare, vengono sì criticati (e censurati in sede di impugnazione) ma al contempo accettati, mentre l'essere umano ha difficoltà a sopportare l'errore realizzato da una macchina, come se ci aspettassimo più competenza dai sistemi informatici che da noi stessi.[329] Bisogna solo ammettere l'esistenza di *black box* negli umani come nelle macchine e a comprendere come muoversi nel sistema.[330]

Un'altra critica legata ai sistemi di apprendimento automatico riguarda la lesione dell'articolo 101 della Costituzione in quanto il giudice, che dovrebbe essere soggetto esclusivamente alla legge, si ritroverebbe a decidere senza giudicare affidandosi senza remore ad un algoritmo intelligente. La certezza del diritto diventerebbe illusoria e astratta, scollegata dalle esigenze della vicenda concreta.[331] Inoltre l'articolo 102 della Costituzione affida l'esercizio della funzione giurisdizionale ai magistrati e l'articolo 111 secondo comma della Costituzione impone che i processi vengano svolti difronte ad un giudice terzo e imparziale. Dunque l'impiego dell'intelligenza artificiale in sostituzione al giudice sarebbe contraria ai nostri principi costituzionali attinenti alla giurisdizione, al diritto di difesa delle parti, alla qualità della decisione e all'obbligo di motivazione.[332]

---

[327] S. QUATTROCOLO "*Processo penale e rivoluzione digitale: da ossimoro a endiadi?"* op. cit., Paragrafo 3 (pag. 126 – 132) "La prova generata automaticamente e i rischi per la parità delle armi".
[328] A. SANTOSUOSSO, op. cit., Paragrafo 4 (pag. 85 - 122) "L'intelligenza artificiale e i giudici".
[329] G. CONTISSA, G. LASAGNI, G. SARTOR, op. cit., Paragrafo 5 (pag. 630 – 631) "Tutta colpa dell'intelligenza artificiale? Black box "umane" e processo penale".
[330] A. SANTOSUOSSO, op. cit., Paragrafo 4 (pag. 85 - 122) "L'intelligenza artificiale e i giudici".
[331] G. UBERTIS, op. cit., Paragrafo 7 (pag. 82 – 83) "… in sede giurisdizionale".
[332] F. DONATI, op. cit., Paragrafo 6 (pag. 428 - 430) "Limiti all'applicazione di strumenti di giustizia predittiva per la soluzione di controversie giudiziarie".

L'idea di fondo, seguita per legittimare l'utilizzo di sistemi di intelligenza artificiale a sostegno delle decisioni del giudice, è legata all'efficienza delle decisioni e alla limitazione dei pregiudizi personali del giudice persona fisica al fine di fornire maggiore obiettività e neutralità nell'applicazione del diritto.[333] Ma il caso Loomis ha dimostrato come possano essere sollevate perplessità in relazione alla loro effettiva validità ed imparzialità al punto di poter giungere ad un elevato grado di discriminazione. [334] L'aspetto positivo degli algoritmi è che escludono le intuizioni soggettive e l'arbitrarietà del processo rappresentando un modello matematico codificato in modo da poter essere gestito da una persona fisica. È necessaria la rivedibilità delle decisioni realizzate dalla macchina per garantire un equo processo. Secondo tale teoria, fare affidamento esclusivamente sull'esperienza del giudice non risulterebbe etico generando sentenze oscure e irrazionali: bisognerebbe selezionare le possibili opzioni nella scelta sanzionatoria al fine di individuare la migliore per ragioni rieducative e di limitazione della recidiva, focalizzandosi sul trattamento che meglio risponde alle reali necessità del reo.[335]

Fino ad ora, in Europa, gli algoritmi predittivi di pericolosità sociale non sono entrati nelle aule penali[336] grazie anche ai limiti imposti dal GDPR e al divieto sancito dall'articolo 220 comma 2 c.p.p. nel quale vengono dichiarate inammissibili le perizie per stabilire l'abitualità o la professionalità nel reato, la tendenza a delinquere, il carattere e la personalità dell'imputato.[337]

In Italia il desiderio di applicazione dell'intelligenza artificiale al sistema penale è legato a diverse motivazioni: la crisi della certezza del diritto a seguito della crescente incidenza del precedente e degli studi sulla calcolabilità del diritto che spingono verso una giustizia predittiva affidata alle macchine; la notevole espansione delle misure di prevenzione basate principalmente sulla valutazione della pericolosità del soggetto su dati indiziari e tendenzialmente predittiva; la sfiducia nei confronti del giudice-umano che genera un interesse verso macchine intelligenti che possano sostituire il giudice rimanendo più razionali e meno fallibili liberando la giustizia dall'emotività e passioni degli esseri umani. La sfiducia verso la magistratura porta verso la fiducia nelle macchine e nei processi automatizzati e neutrali. Infine, l'algoritmo può essere letto dal giudice come strumento di deresponsabilizzazione rispetto alle decisioni che gravano su di lui.[338]

La sorveglianza umana è un presidio necessario per garantire l'affidabilità dell'intelligenza artificiale anche alla luce dei principi del giusto processo.[339] Un'idea per il suo utilizzo in sede di commisurazione della pena, potrebbe essere quella di limitarlo esclusivamente per situazioni in *bonam partem*: la pena deve necessariamente rispettare il fine rieducativo, deve essere disegnata su misura dell'interessato ed inoltre è connessa al grado di colpevolezza espressa nel fatto commesso. La possibilità di intervenire con azioni di

[333] C. BURCHARD, op. cit., Paragrafo 3 (pag. 10 – 19) "Le promesse dell'intelligenza artificiale per l'amministrazione della giustizia penale".
[334] F. BASILE, op. cit, Paragrafo 5.4.1 (pag. 22 – 24) "Considerazioni conclusive".
[335] A. M. MAUGERI, op. cit., Paragrafo 2 (pag. 6 - 12) "I vantaggi dell'uso di algoritmi predittivi contro l'incertezza del giudizio di pericolosità sociale".
[336] C. PARODI E V. SELLAROLI, op. cit., Paragrafo 8 (pag. 66 – 70) "Le valutazioni di pericolosità".
[337] F. BASILE, op. cit., Paragrafo 5.4.1 (pag. 22 – 24) "Considerazioni conclusive".
[338] V. MANES, op. cit., Paragrafo 4 (pag. 10 – 11) "Le prospettive di applicazione dell'AI nel sistema penale italiano".
[339] A. M. MAUGERI, op. cit., Paragrafo 7 (pag. 29 - 34) "Il giudizio individualizzato e il controllo del giudice sull'output".

prevenzione speciale viene concessa dall'ordinamento italiano solo al fine di mitigare l'intervento punitivo nel caso di limitata capacità a delinquere rispettando dunque il principio di individualizzazione della pena insito nel principio sancito dall'articolo 27 della Costituzione andando ad individuare il trattamento più adatto in considerazione delle caratteristiche del soggetto (come nel caso dei *sex-offender*).[340] Sarebbe possibile inoltre seguire un diverso approccio andando a considerare l'algoritmo solo come strumento volto alla limitazione della fallibilità delle decisione giudiziale nella collaborazione uomo-macchina, imponendo la costante verifica dell'*output* prodotto dall'intelligenza artificiale alla luce di altri elementi di prova. Usare la macchina come strumento di verifica della scelta operata dal giudice imponendo un obbligo di motivazione rafforzata nei casi in cui la decisione prospettata dal sistema intelligente e quella ipotizzata dalla persona fisica non coincidano (seguendo la giurisprudenza della sentenza Cozzini[341]) valutando inoltre il rispetto dei canoni di verificabilità della prova scientifica raggiunto dall'algoritmo. È importante sottolineare come la decisione del giudice necessiti di restare centrale.[342]

L'intelligenza artificiale dovrebbe essere concepita come base per un futuro diritto penale aperto alla tecnologia e all'uomo nell'ottica della ragionevolezza.[343] Per raggiungere tale obiettivo è, però, necessario che gli operatori del diritto vengano formati in modo tale da avere una piena consapevolezza delle capacità e limiti dell'intelligenza artificiale imparando ad interagire correttamente con essa. Per essere certi che la macchina non sia di per sé lesiva dei diritti fondamentali, sarebbe utile istituire un meccanismo di certificazione *ex ante* che ne analizzi il sistema a seguito di un controllo svolto da autorità pubbliche in modo tale che le imprese che le producono siano costrette a produrre informazioni sul loro funzionamento e qualità, allo stesso tempo garantendo tutela per i segreti industriali (tale approccio già sussiste in ambito medico e di aviazione). È importante imporre che la decisione presa dall'intelligenza artificiale non possa, da sola, fondare la decisione e che sia sempre accompagnata da altri elementi probatori. Inoltre sarebbe interessante introdurre un secondo controllo sui risultati raggiunti dall'intelligenza artificiale tramite un secondo algoritmo sviluppato e progettato in modo diverso.[344]

[340] A. M. MAUGERI, op. cit., Paragrafo 8 (pag. 34 - 37) "Le prospettive relative all'uso dell'algoritmo in bonam partem".
[341] Sentenza della Cassazione Penale, Sez. IV, n. 43786 del 17 settembre 2010 (Cozzini).
[342] V. MANES, op. cit., 6 (pag. 19 - 22) "Riflessioni conclusive: l'algoritmo come supporto alla decisione giudiziale".
[343] C. BURCHARD, op. cit., Paragrafo 6 (pag. 32 – 34) "Prospettive".
[344] G. CONTISSA, G. LASAGNI, G. SARTOR, op. cit., Paragrafo 6 (pag. 631 – 634) "Verso un ricorso (davvero) effettivo: alcune proposte".

*3.4 Intelligenza Artificiale come strumento, vittima o autore del reato*

L'intelligenza artificiale può, anche, risultare come l'oggetto del giudizio. In tali casi, è necessario adattare le categorie giuridiche tradizionali al mondo tecnologico in cui le decisioni vengono assunte da algoritmi che spesso adottano scelte non prevedibili neanche dal programmatore.

Fino a pochi anni fa, le macchine erano integralmente progettate dal programmatore e da lui preimpostate in tutte le loro azioni: non si avevano margini di discrezionalità in capo al dispositivo e tutto poteva essere previsto *ex ante* e controllato dall'essere umano. Non si poneva, dunque, il dubbio sul considerare o meno il *robot* come titolare di diritti e di doveri o come autore di un reato; era sufficiente il rispetto del modello di imputazione della responsabilità indiretta dell'uomo. Tale principio permette di muovere un rimprovero, almeno a titolo di colpa, verso l'essere umano. La macchina si atteggiava come mero strumento nelle mani del *reo*.[345]

Oggi, l'intelligenza artificiale ha permesso un *quid pluris* ai *robot*, finalmente in grado di agire sulla base di scelte discrezionali non sempre prevedibili *ex ante* dal programmatore grazie ai meccanismi di *machine learning*. Ed è qui che si focalizza il fulcro di questo elaborato: analizzare e individuare la responsabilità derivante da lesioni ai beni giuridici causate da sistemi di intelligenza artificiale, siano essi *driveless cars*, droni o sistemi di diagnostica autonomi.[346]

Tragici eventi sono accaduti negli ultimi anni in cui, ad esempio, nel 2016, un drone ha ucciso un sospetto cecchino nella città di Dallas e nel 2018 un'auto senza conducente ha investito un pedone in Arizona.[347]

A seguito di tali eventi sorgono in ognuno di noi domande al limite degli esperimenti sociologici: un autoveicolo intelligente, lungo un percorso trova legate quattro persone incapaci di spostarsi, per deviarle finirebbe in un burrone mettendo a rischio la vita del conducente e del passeggero. Come bisognerebbe programmare l'algoritmo? Quale scelta sarebbe la migliore? Chi risponderebbe del reato posto in essere: il programmatore, il produttore o l'utilizzatore? Potrebbe rispondere l'intelligenza artificiale come autore del reato in base alle sue caratteristiche? [348]

Per superare (o quantomeno provare a spiegare) tali interrogativi da un punto di vista giuridico penalista occorre soffermarci, come suggerisce il titolo di questo sotto-paragrafo, sulle tre ipotesi di coinvolgimento di un sistema di intelligenza artificiale nella commissione di un reato: strumento, vittima e autore dello stesso.

---

[345] R. BORSARI "*Intelligenza Artificiale e responsabilità penale: prime considerazioni"* in rivista quadrimestrale di Media Laws, Rivista di Diritto dei Media, 3/2019, Milano; Paragrafo 3 (pag. 264) "Le entità intelligenti come strumento del reato commesso dall'uomo": https://www.medialaws.eu/wp-content/uploads/2019/11/borsari.pdf

[346] F. DONATI, op. cit., Paragrafo 1 (pag. 415 - 418) "Premessa".

[347] F. BASILE, op. cit, Paragrafo 6 (pag. 24 – 25) "Quarto percorso d'indagine - IA e reato: possibili ipotesi di coinvolgimento – come *strumento*, come *autore*, o come *vittima* – di un sistema di IA nella commissione di un reato".

[348] U. PAGALLO, op. cit., Paragrafo 2 (pag.617 - 619) "Le sfide etiche dell'IA".

Nel caso di strumento del reato, ci ritroveremmo in una situazione nella quale l'intelligenza artificiale, date le sue infinite capacità, viene utilizzata dall'uomo al fine di commettere il reato. Ciò avviene soprattutto nei crimini informatici, economici e ambientali, nel traffico di sostanze stupefacenti e di esseri umani, nei reati di diffamazione, nella realizzazione di *fakenews* tramite *bot* e nelle le lesioni alla *privacy* o alla proprietà intellettuale.

Due esempi riguardano il bagarinaggio *on-line* e la manipolazione del mercato (art. 185 T.U.F).

Nel primo reato la condotta riguarda il caso in cui un sito metta in vendita dei biglietti per un evento e in un brevissimo lasso di tempo ne venga acquistato un gran quantitativo da pochi soggetti attraverso i *bot* ad una velocità inaccessibile a qualsiasi essere umano per poi rivenderli in un mercato parallelo a prezzi più alti comportando, inoltre, fenomeni di elusione ed evasione fiscale.

Nel secondo esempio, le manipolazioni vengono commesse tramite programmi informatici che eseguono transazioni finanziarie decise a seguito di comparazioni realizzate da un algoritmo chiamato "*HFT*" (*High Frequency Traders*) capace di realizzare un elevato numero di operazioni in frazioni di secondo. Ciò causa oscillazioni, di rilevanza penale, dei prezzi sui mercati finanziari senza mutamenti del valore sostanziale del titolo. [349]

Come analizzeremo nel secondo capitolo, nel caso di intelligenza artificiale utilizzata quale strumento del reato, la responsabilità dovrebbe essere imputata a titolo di dolo all'utilizzatore dello strumento se intenzionalmente lo sfrutta al fine di commettere illeciti e a titolo di colpa all'utilizzatore o programmatore, qualora il reato si sia realizzato a causa di un malfunzionamento della macchina intelligente che si sarebbe potuto evitare grazie ad un suo comportamento attivo evitando la realizzazione dell'evento.[350] Si giunge a tale conclusione in base al fatto che, in queste situazioni, la macchina non si pone come l'autore dell'illecito grazie ad un suo autonomo comportamento causativo dell'evento. È l'essere umano ad aver ideato, voluto e realizzato l'evento utilizzando come strumento l'intelligenza artificiale.

Per quanto riguarda la situazione in cui l'intelligenza artificiale sia considerata come vittima dell'reato, è possibile analizzare fenomeni in cui i *robot*, con sembianze umane o animali, utilizzati nei casi di *doll therapy* o *pet therapy*, a sostengo di autistici, disabili o malati di *Alzheimer* vengano distrutti o danneggiati. Essendo dotati di capacità cognitiva e realizzati allo scopo di instaurare una relazione con il soggetto, ci si domanda se tale condotta integri un semplice danneggiamento o se si sia al limite del maltrattamento di animali (art. 544 *ter* c.p.) se non, addirittura, maltrattamento contro familiari e conviventi (art. 572 c.p.).

Un altro esempio concerne i casi del cosiddetto stupro robotico riguardante gli atti sessuali con *robot* aventi le sembianze di minori. Sussistono, al contempo, dubbi in relazione al concetto del consenso (inteso in termini diversi rispetto all'essere umano) all'atto sessuale da parte della macchina intelligente. Come vedremo nel terzo capitolo, risulta interessante analizzare il rapporto con l'articolo 600 *quater* 1 c.p. in base al quale viene punita la pornografia virtuale.

---

[349] F. BASILE, op. cit., Paragrafo 6.2 (pag. 25 – 27) "Il sistema di IA quale *strumento* di commissione del reato".
[350] G. UBERTIS, op. cit., Paragrafo 5 (pag. 80 – 81) "Profili di diritto sostanziale".

Considerare l'intelligenza artificiale come vittima del reato si affianca alla tematica, accennata nel secondo paragrafo di questo capitolo ma che svilupperemo maggiormente nel prossimo, del considerare l'intelligenza artificiale non solo come cosa inanimata ma anche come soggetto equiparato ad una persona che subisce un reato. Si arriverebbe ad un'umanizzazione della macchina limitata dalla concezione che i sistemi di intelligenza artificiale non hanno e non avranno sentimenti umani.[351] Proprio per questo la dottrina tende ad interrogarsi sulla possibilità di introdurre nuove figure di reato e di categorie giuridiche calibrate secondo i criteri delle macchine intelligenti tramite un intervento legislativo.[352]

In ultima ipotesi, se l'intelligenza artificiale coinvolta nella commissione del reato fosse di ultima generazione, con autonoma capacità di apprendimento e di decisione, si potrebbe considerare essa stessa come autore del reato e non come mero strumento.

Si pone quindi il problema: un sistema di intelligenza artificiale può rispondere penalmente?

Rispondere a questa domanda non è cosa da poco ed è necessario svolgere un'indagine d'insieme alla luce di tutti gli aspetti legati al concetto di responsabilità trasmigrato sul sistema di intelligenza artificiale. È difficile escludere del tutto la macchina intelligente dal meccanismo di attribuzione della responsabilità, proprio perché il *robot* ha la capacità di incidere nell'ambiente circostante in modo autonomo. Per imputare la responsabilità al programmatore, fornitore o utilizzatore, è necessario dimostrare il grado di controllo che questi avevano sul dispositivo: maggiore è l'intelligenza della macchina, minore è possibilità di garantire tale controllo, anche perché, il *software*, apprendendo autonomamente, non sarebbe *ex ante* prevedibile e dunque potrebbe giungere a decisioni non immaginate o volute dall'essere umano. Non siamo più difronte ad un mero strumento nelle mani dell'uomo ma a macchine che hanno la capacità di muoversi a prescindere dall'intervento del programmatore o utilizzatore senza la necessità di comandi precedentemente impartiti.[353]

In alcuni casi l'attività della macchina risulta essere prevedibile dal programmatore, potendo dunque applicarsi lo schema della responsabilità indiretta o vicaria, facendo rispondere per colpa il programmatore, il costruttore o l'utilizzatore dell'intelligenza artificiale.

In altri, invece, il grado di autonomia decisionale, in capo all'elaboratore, risulta elevato al punto tale da poter modificare gli algoritmi originari in modo non ipotizzabile dal programmatore. In questi casi si potrebbe addossare la responsabilità (oltre alle ipotesi che considerano responsabile la macchina) anche al programmatore, costruttore o utilizzatore per gli eventi cagionati dal *robot* esclusivamente quando gli stessi abbiano eluso un dovere giuridico di monitoraggio dello strumento in base allo schema dell'articolo 40 cpv. c.p.: è infatti necessario che lo stesso possieda i poteri giuridici o naturalistici di impedire l'evento.

Ad esempio, nel 2014 il *robot* "*Vital"* prodotto da *Aging Analytics* venne nominato come membro del consiglio di amministrazione di *Deep Knowledge*, società giapponese. La scelta fu legata alle sue capacità di

[351] F. BASILE, op. cit., Paragrafo 6.4 (pag. 31 – 33) "Il sistema di IA quale vittima del reato".
[352] F. BASILE, op. cit., Paragrafo 6 (pag. 24 – 25) "Quarto percorso d'indagine - IA e reato: possibili ipotesi di coinvolgimento – come *strumento*, come *autore*, o come *vittima* – di un sistema di IA nella commissione di un reato".
[353] F. BASILE, op. cit., Paragrafo 6.3 (pag. 27 – 28) "Il sistema di IA quale autore del reato: machina delinquere potest?".

individuare tendenze di mercato non ovvie per gli esseri umani in relazione ad investimenti di successo. Nel caso in cui "*Vital"* commetta una condotta di distrazione che conduca alla bancarotta fraudolenta della società (*ex* art. 216 L.F.): chi risponde? I membri del consiglio di amministrazione potranno essere perseguiti per negligenza oppure potrebbe essere perseguita la società in quanto tale e il *robot* intelligente come rappresentante della stessa. La situazione è attualmente confusa non essendoci una normativa idonea a regolamentare tali situazioni.[354]

L'essere umano, nel caso di responsabilizzazione della macchina si sentirebbe di fatto deresponsabilizzato. Si avrebbe quello che può ben essere definito come scarico di responsabilità morale dell'agente umano sul *robot* intelligente come nel caso di droni da combattimento militare: la direzione dell'azione rimane all'essere umano ma non è lui il reale agente (considerando che lo stesso si trova distante dallo scenario di guerra e quindi non ne risulta condizionato) innestando una complessa catena di comando a cui attribuire la responsabilità. Si giungerebbe ad un'alienazione dell'azione con effetti deresponsabilizzanti: il drone non innesca un'interazione con l'avversario limitandogli la possibilità di negoziazione e di difesa. Si avrebbe dunque una riduzione delle tutele garantite dal diritto penale ai beni giuridici. Più si riduce il grado di responsabilizzazione dell'essere umano più si pone il tema della possibilità di responsabilizzare penalmente e direttamente i sistemi di intelligenza artificiale.[355]

Si può quindi responsabilizzare un non-umano?

Platone, nella sua ultima opera "*Le Leggi*"[356], attribuisce la responsabilità penale anche alle cose e agli animali; tale concezione prosegue fino agli inizi dell'Illuminismo. È possibile notarlo anche nell'attuale ordinamento giuridico italiano che, tramite il d.lgs. 231 del 2001[357], configura margini di responsabilità a carico degli enti (persone giuridiche e non fisiche). La questione si fonda sul quesito del se sia possibile o meno considerare i sistemi di intelligenza artificiale come persone o quantomeno equipararli ad esse. [358]

Saranno questi i temi oggetto del prossimo capitolo, incentrato sulla responsabilità penale per i reati commessi dall'intelligenza artificiale avendo riguardo a questioni come la personalità giuridica della macchina, la colpevolezza, il fine rieducativo della pena, la responsabilità oggettiva e il controllo significativo esercitato dall'essere umano. Lo scopo è comprendere se *machina delinquere potest o non potest*.

---

[354] U. PAGALLO, op. cit., Paragrafo 4.1.1 (pag. 621 - 623) "IA criminale".
[355] F. BASILE, op. cit., Paragrafo 6.3.1 (pag. 28 – 29) "Tra deresponsabilizzazione dell'uomo e responsabilizzazione della macchina"
[356] PLATONE, "*Le Leggi*" IV secolo a.C.
[357] 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.
[358] F. BASILE, op. cit., Paragrafo 6.3.2 (pag. 29) "Vacilla il confine tra *machina* e persona?".

# *CAPITOLO 2:*

# *L'INTELLIGENZA ARTIFICIALE E LA RESPONSABILITÀ 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*

Abbiamo terminato il primo capitolo analizzando gli ambiti del diritto penale in cui il sistema di intelligenza artificiale può essere impiegato. È stato studiato come ausilio del giudice e della polizia giudiziaria, abbiamo osservato il suo impiego nella commissione degli illeciti, come strumento, come vittima e come autore del reato. Adesso dobbiamo volgere la nostra 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 lui commessi e dunque sarà interessante 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, costituzionalmente inaccettabili.

Procediamo con ordine.

È possibile considerare l'intelligenza artificiale come un agente soggetto al diritto? È possibile imputare alla macchina diritti e doveri?

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.[359]

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.[360] 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. [361]

"*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)*."[362]

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.[363] 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 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

[359] Random House Kernerman Webster's College Dictionary, 2010: https://www.thefreedictionary.com/agent
[360] A. SANTOSUOSSO, op. cit., Paragrafo 7.9 (pag. 193 – 195) "Diritti, storicità, artificialità – L'agente per il diritto".
[361] *Idem.*
[362] L. FLORIDI, "*Philosophy of computing and information. 5 Questions*", (pag. 95) Automatic Press/VIP, 2008.
[363] H. KELSEN,"*General Theory of Law and State*", (pag. 93 – 95) Harvard University Press, Cambridge, 1945.

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.[364]

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.[365]

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[366] 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. Nonostante le loro differenze, entrambe sono considerate come entità titolari di diritti e doveri. Le società sono enti artificiali che 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.[367]

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

[364] A. SANTOSUOSSO, op. cit., Paragrafo 7.8 (pag. 190 – 192) "Diritti, storicità, artificialità – Creature di Dio e artefatti giuridici".
[365] 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".
[366] E. PALMERINI, op. cit., Paragrafo 3.2. "Che cos'è un robot? Base tecnologica e questioni giuridiche".
[367] A. SANTOSUOSSO, op. cit., Paragrafo 7.9 (pag. 193 – 195) "Diritti, storicità, artificialità – L'agente per il diritto".

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 appena affermato, 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, 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.

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.[368]

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 gli 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 desiderabilità sociale dell'innovazione tecnologica e che dunque i danneggiati dovrebbero pur sempre essere tutelati. [369] 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

[368] E. PALMERINI, op. cit., Paragrafo 6.1 "Le proposte di schemi alternativi alle comuni regole di responsabilità".
[369] E. PALMERINI, op. cit., Paragrafo 6.2 "Alcuni rilievi critici".

da coloro che sono intervenuti nella sua creazione, in modo tale da permettere alla macchina intelligente di rispondere delle obbligazioni.[370] 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.[371] Il riconoscimento della personalità elettronica al sistema autonomo potrebbe generare un crescente antropomorfismo nei confronti degli agenti artificiali rischiando di comportare una deresponsabilizzazione degli operatori. L'autonomia dei *robot* non viene considerata come sufficiente al fine del riconoscimento della soggettività giuridica.[372]

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.[373] 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.[374]

Analizziamo di seguito le varie teorie.

### *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 artificiatale, 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*.[375]

---

[370] E. PALMERINI, op. cit., Paragrafo 6.1 "Le proposte di schemi alternativi alle comuni regole di responsabilità".
[371] E. PALMERINI, op. cit., Paragrafo 6.2 "Alcuni rilievi critici".
[372] 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".
[373] E. PALMERINI, op. cit., Paragrafo 6.1 "Le proposte di schemi alternativi alle comuni regole di responsabilità".
[374] E. PALMERINI, op. cit., Paragrafo 6.2 "Alcuni rilievi critici".
[375] 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".

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. [376] In tal modo, si genera una separazione dei patrimoni e dunque, uno schermo per i produttori e investitori evitandogli un'eccessiva esposizione al rischio.

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.[377] 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.[378] 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.[379]

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

---

[376] 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'".

[377] 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".

[378] 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".

[379] 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".

di diritto liberi e capaci di compiere scelte volontarie e dunque non considerabili come penalmente rimproverabili o punibili per le condotte da loro realizzate.[380]

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.[381] 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.[382]

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 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.[383]

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

[380] B. PANATTONI, op. cit., Paragrafo 5.1 "Forme di responsabilità diretta a carico dell'agente artificiale".
[381] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?".
[382] 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".
[383] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?".

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.[384]

### *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.[385]

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, 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. [386]

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.[387]

---

[384] 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".

[385] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'".

[386] 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".

[387] 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".

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.[388]

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 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. [389] 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[390] 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

---

[388] 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".

[389] M. B. MAGRO, "*Biorobotica, robotica e diritto penale*", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali".

[390] 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".

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 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. [391]

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. [392]

### *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 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

[391] 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".
[392] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'".

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. [393]

### *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 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.

---

[393] 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".

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. [394]

### *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 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. [395]

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

[394] 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".

[395] S. QUINTARELLI, op. cit., Capitolo 6.2 (pag. 110 – 113) "Personalità giuridica. La convivenza tra uomini e software".

giuridico, la società esiste in quanto persona giuridica a differenza dell'artefatto intelligente che non è dotato di personalità.[396] Nel caso in cui dovessimo considerare una soggettività giuridica in capo all'intelligenza artificiale, l'elemento umano verrebbe marginalizzato al punto tale da scomparire del tutto; questione impensabile nelle società in cui la presenza di tale elemento è necessaria per la loro sopravvivenza. [397]

Infatti è possibile notare similitudini tra il brocardo latino "*societas delinquere non potest"* (oramai superato) e il brocardo *"machina delinquere non potest*" il cui superamento questa tesi ha il compito di analizzare e di valutare.

L'obiezione tradizionale alla possibile responsabilità penale degli enti era basata sul fatto che le società, mancando di corpo e anima non possono integrare l'elemento soggettivo del reato e non possono neanche porre in essere un'azione propria. Inoltre, sussiste un'impossibilità di applicare la pena carceraria alla società anche se, ad oggi, esistono diverse tipologie di sanzioni con una notevole capacità afflittiva. Infine, si riscontrava il rischio di una violazione del principio della responsabilità penale personale sancito dall'articolo 27 della Costituzione. Infatti, per evitare di ricadere in tale violazione, il decreto legislativo 231 del 2001[398] prevede esclusivamente sanzioni amministrative.[399] Nonostante tale previsione, la Corte Suprema di Cassazione[400], riprendendo i criteri Engel espressi dalla Corte Europea dei Diritti dell'Uomo[401] (la qualificazione giuridica della misura in causa nel diritto nazionale, la natura della misura e la natura e il grado di severità della sanzione), per evitare violazioni del *ne bis in idem*, ha affermato che, qualunque sia la denominazione usata dal legislatore, tali sanzioni hanno natura sostanzialmente penale attribuendo inoltre, capacità penale alle società non più considerate come un mero schermo per coprire le persone fisiche che la compongono. [402]

Parte della dottrina, tra cui Hallevy, pone l'accento sul fatto che le società nonostante non abbiano un corpo e un'anima vengono considerate responsabili penalmente perché, in un bilanciamento di interessi è prevalsa la necessità di repressione di determinati reati commessi dalle *corporation* portando al superamento del brocardo *societas delinquere non potest*. Quindi, ciò che frena il superamento del medesimo brocardo declinato per i sistemi di intelligenza artificiale, secondo tale dottrina, è un pregiudizio antropocentrico. [403]

---

[396] 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".
[397] S. QUINTARELLI, op. cit., Capitolo 6.2 (pag. 110 – 113) "Personalità giuridica. La convivenza tra uomini e software".
[398] 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.
[399] A. SANTOSUOSSO, op. cit., Paragrafo 7.10 (pag. 195 – 197) "Diritti, storicità, artificialità – Società commerciali e pene: chi fa schermo a chi".
[400] Sentenza della Cassazione Penale, Sezione III del 30 gennaio 2006, n. 3615, (c.d Caso Jolly Mediterraneo s.r.l.).
[401] Sentenza della Corte Europea dei Diritti dell'Uomo, Grande Camera, caso Engel e altri c. Paesi Bassi, 8 giugno 1976.
[402] A. SANTOSUOSSO, op. cit., Paragrafo 7.10 (pag. 195 – 197) "Diritti, storicità, artificialità – Società commerciali e pene: chi fa schermo a chi".
[403] 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".

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. [404]

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. [405] 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 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.[406]

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[407], ed è loro che il decreto legislativo 231/2001[408] vuole colpire andando a considerare una responsabilità penale degli enti. [409]

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*.[410] 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

[404] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'".
[405] 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".
[406] B. PANATTONI, op. cit., Paragrafo 2 "Dall'automazione tecnologica all'autonomia artificiale".
[407] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?".
[408] 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.
[409] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'".
[410] 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".

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,[411] 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.[412] Ad essere imputato è un fatto proprio dell'ente e la colpevolezza rientra in una colpa di organizzazione propria della società[413] (l'ente è responsabile solo se non si è organizzato per evitare il reato);[414] 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. [415]

## ***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. [416]

Nonostante ciò, è osservabile 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*".[417] 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

---

[411] 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.

[412] 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".

[413] 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".

[414] 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".

[415] B. PANATTONI, op. cit., Paragrafo 5.1 "Forme di responsabilità diretta a carico dell'agente artificiale".

[416] I. SALVADORI, op. cit., Paragrafo 4, "Gli agenti artificiali come autori di un reato".

[417] 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'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.[418]

Bisogna, però, valutare con attenzione se l'evoluzione robotica possa generare attriti con le norme esistenti. È importante analizzare un quadro di regole chiaro per garantire certezza rispetto ai doveri e alle responsabilità degli attori del progresso tecnologico al fine, inoltre, di indirizzare tale sviluppo al rispetto dei diritti fondamentali.[419] Viene quindi coinvolta anche la branca del diritto penale che apparentemente sembrerebbe la più lontana in relazione all'impatto dell'intelligenza artificiale nel mondo del diritto. Il diritto penale è infatti fortemente connesso all'essere umano: solo un umano può commettere un reato in modo colpevole dato che solo esso risulta in possesso di beni giuridici e solo lui può lederli per scopi personali. E dunque, solo un essere umano può essere punito per tali azioni. Da qui nasce il brocardo *machina delinquere non potest*: il *robot* non può essere direttamente responsabile per la commissione di reati, al massimo si potrà parlare di responsabilità indiretta dell'individuo persona fisica considerando la macchina come mero strumento nelle mani dell'essere umano. [420]

L'evoluzione dell'autonomia dei sistemi di intelligenza artificiale mette però in crisi tale modello tradizionale di responsabilità indiretta.[421] Infatti, nel caso di verificazione di un illecito, questo non potrebbe essere attribuito all'utente/creatore tranne il caso in cui la macchina sia stata progettata appositamente per realizzare uno scopo illecito oppure nel caso in cui si tratti di un evento dovuto ad un errore di funzionamento. Per questo, parte della dottrina ritiene di dover riconoscere capacità penale alle macchine intelligenti in modo tale da ricondurre un fatto penalmente rilevante ad una loro azione causalmente punibile. [422]

Infatti, più i *robot* sono autonomi, meno possono essere considerati come semplici strumenti del reato: ciò pone dubbi sulla sufficienza delle attuali norme giuridiche e sulla necessità di istituire nuovi principi e regole in tema di responsabilità penale dei vari attori per le azioni imputabili alla macchina.

Tale dilemma si pone, non quando il *robot* sia comandato dall'umano, ma quando il sistema abbia la capacità di muoversi autonomamente in base ad un algoritmo di *machine learning* andando a generare, inoltre, problematiche in relazione all'articolo 27 della Costituzione secondo cui la responsabilità penale è personale. Come analizzeremo nel prossimo paragrafo, nascono quindi dubbi sulla tipologia di responsabilità da considerare: oggettiva dell'essere umano, diretta della macchina oppure indiretta e da controllo dell'essere umano.[423]

Dunque l'intelligenza artificiale 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.

---

[418] E. PALMERINI, op. cit., Paragrafo 3 "Una miriade di problemi, un approccio unitario?".
[419] *Idem.*
[420] 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".
[421] R. BORSARI, op. cit., Paragrafo 1 (pag. 262 - 263) "Evoluzione dell'Intelligenza Artificiale e diritto penale".
[422] I. SALVADORI, op. cit., Paragrafo 4, "Gli agenti artificiali come autori di un reato".
[423] A. LONGO E G. SCORZA, op. cit., Capitolo 7 (pag. 198 – 240) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela".

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.[424] 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*".[425]

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).[426]

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 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. [427]

Il principio dell'articolo 27 della Costituzione secondo cui la responsabilità penale è personale pone un limite quasi invalicabile alla possibilità di considerare il sistema basato sull'intelligenza artificiale come responsabile per un reato commesso. La personalità di cui parla l'articolo necessita di essere riferita ad un

[424] A. SIMONCINI, op. cit., Paragrafo 3 (pag. 67 - 69) "L'impatto della rivoluzione cibernetica su diritto costituzionale: intelligenza artificiale, autonomia e libertà".
[425] RISOLUZIONE DEL PARLAMENTO EUROPEO DEL 16 FEBBRAIO 2017 recante "*Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*" (2015/2103(INL)), op cit.
[426] A. SIMONCINI, op. cit., Paragrafo 3.1 (pag. 69 – 71) "Il soggetto "catturato" dallo strumento".
[427] I. SALVADORI, op. cit., Paragrafo 4, "Gli agenti artificiali come autori di un reato"

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. [428]

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 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.[429]

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

[428] *Idem*.
[429] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'".

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.[430]

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 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

[430] 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".

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[431]): dunque, grazie a tale correttivo, è possibile affermare che è causa dell'evento ogni azione che, tenendo conto delle circostanze del caso concreto, non 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

[431] 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

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).[432]

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 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

[432] 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".

umano risulta tendenzialmente libero di esprimersi e di muoversi senza condizionamenti limitativi del suo comportamento posti a priori. [433]

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.

Su queste basi si fonda la prospettiva di una possibile responsabilità penale in capo alle macchine intelligenti. [434]

Procediamo con ordine.

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[435], 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

[433] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'".
[434] F. BASILE, op. cit., Paragrafo 6.3.3 (pag. 30 – 31) "Una colpevolezza "disumana"?".
[435] Sentenza della Corte Costituzionale n. 364, del 23 marzo 1988: https://www.cortecostituzionale.it/actionSchedaPronuncia.do?anno=1988&numero=364

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.[436]

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 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). [437]

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*."

[436] 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".
[437] 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".

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. [438]

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 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).[439]

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. [440]

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.[441]

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 e non consapevoli che però si riflettono sul sistema

[438] A. SANTOSUOSSO, op. cit., Paragrafo 7.11 (pag. 198 – 202) "Diritti, storicità, artificialità – Creature artificiali e proprietà ontologiche".
[439] S. QUINTARELLI, op. cit., Capitolo 6.2 (pag. 110 – 113) "Personalità giuridica. La convivenza tra uomini e software".
[440] A. SANTOSUOSSO, op. cit., Paragrafo 7.11 (pag. 198 – 202) "Diritti, storicità, artificialità – Creature artificiali e proprietà ontologiche".
[441] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'".

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.[442] 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.[443]

Questo viene confermato anche dal *test* della stanza cinese ideato da Searle, filosofo statunitense, per confutare la tesi di Turing: 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 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

[442] 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".
[443] M. B. MAGRO, "*Biorobotica, robotica e diritto penale*", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali".

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. [444]

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) 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.[445]

In base a questo approccio un sistema intelligente potrebbe ottenere soggettività giuridica e capacità di agire superata una certa soglia di coscienza emotiva.[446]

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. [447]

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

---

[444] M. B. MAGRO, "*Biorobotica, robotica e diritto penale*", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali".

[445] 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".

[446] S. QUINTARELLI, op. cit., Capitolo 6.2 (pag. 110 – 113) "Personalità giuridica. La convivenza tra uomini e software".

[447] 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".

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 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. [448]

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 altrui; i *robot* non

[448] 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".

sono dotati di capacità di scelta pura: mancando di colpevolezza non possono neanche essere considerati rimproverabili e dunque soggette a pena.[449]

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*".[450] 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. [451]

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 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

[449] 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".
[450] 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".
[451] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'".

azione od omissione. Per effettuare un'interpretazione conforme a Costituzione, è necessario seguire il principio espresso dalla Sentenza 364/1988[452] (esposto all'inizio di questo sotto-paragrafo).[453]

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. [454]

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 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.[455]

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. [456] Essi si dividono in *data bias* (relativi a problematiche concernenti i dati di *input*) e *automation bias* (riguardanti le modalità di acquisizione automatica delle

---

[452] Sentenza della Corte Costituzionale n. 364, del 23 marzo 1988: https://www.cortecostituzionale.it/actionSchedaPronuncia.do?anno=1988&numero=364

[453] 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".

[454] 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".

[455] 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".

[456] M.C. CARROZZA, C. ODDO, S. ORVIETO, A. DI MININ, G. MONTEMAGNI, op. cit., Paragrafo 2 (pag. 2 – 6) "Il dato".

informazioni e la loro analisi, a causa, anche, di discriminazioni congenite della società) che generano errori di valutazione comportando azioni omissive o commissive. [457] 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. [458] 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* 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.[459]

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

[457] 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".
[458] A. LONGO E G. SCORZA, op. cit., Capitolo 1 (pag. 14 – 68) "Un'introduzione all'intelligenza artificiale: le basi storiche e tecnologiche".
[459] 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".

psicologica del *reo*, se non sussiste la piena capacità di intendere e di volere non si può parlare di imputabilità.[460]

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 alla piena maturità psichica e alla piena sanità mentale del soggetto.[461] 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.[462]

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.

---

[460] 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".

[461] 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".

[462] 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. [463]

Secondo M. Minsky, che come abbiamo accennato nel primo capitolo è considerabile come il padre dell'intelligenza artificiale, la volontà è esclusivamente illusoria[464] 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*. [465]

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*. [466] Ma, al contempo, 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. [467]

---

[463] G. GALUPPI, op. cit., Paragrafo 3 "La prospettiva delle varie "scuole" giuridiche sull'imputabilità e sul libero arbitrio".
[464] G. GALUPPI, op. cit., Paragrafo 4 "Libero arbitrio e neuropsicologia moderna".
[465] E. PALMERINI, op. cit., Paragrafo 6 "I problemi della sicurezza e della responsabilità nel mercato emergente della robotica".
[466] 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".
[467] 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".

*2.3 Fine rieducativo della pena*

Accanto, al quesito "*machina delinquere potest*?", bisognerà porsi anche il connesso quesito: "(*quomodo*) *machina puniri potest?"* [468]

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)[469]. 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 perché non è detto che l'obiettivo venga raggiunto (e spesso, purtroppo, il sistema carcerario non riesce a garantirlo).[470]

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

[468] F. BASILE, op. cit., Paragrafo 6.3.4 (pag. 31 – 32) "Quali pene per i sistemi di IA?".
[469] 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
[470] 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".

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).[471]

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*.[472]

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:[473] 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. [474] 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.[475] 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)[476] 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[477] (ma questo significherebbe agire direttamente su di essa, senza la necessità di uno strumento

---

[471] 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

[472] S. QUINTARELLI, op. cit., Capitolo 6.2 (pag. 110 – 113) "Personalità giuridica. La convivenza tra uomini e software".

[473] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?".

[474] F. BASILE, op. cit., Paragrafo 6.3.4 (pag. 31 – 32) "Quali pene per i sistemi di IA?".

[475] 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".

[476] C. PIERGALLINI, op. cit., Paragrafo 4.1 "Machina artificialis delinquere et puniri potest? Grundlinien di un (improbabile) diritto penale 'robotico'".

[477] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?".

indiretto di pressione come la sanzione penale).[478] 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 della pena alla commissione del reato.[479] 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. [480]

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. [481]

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.[482]

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.

---

[478] 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".
[479] R. BORSARI, op. cit., Paragrafo 5 (pag. 266 - 268) "Superamento dell'assioma del machina delinquere (et puniri) non potest?".
[480] 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".
[481] 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".
[482] A. LONGO E G. SCORZA, op. cit., Capitolo 7 (pag. 198 – 240) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a 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.[483] 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). [484]

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.*[485]

Ma procediamo con ordine.

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

[483] M. D'AGOSTINO, M. PAGANINI E R. DI BELLA, op. cit., Paragrafo 3 (pag. 14 - 36) "Qualificazione giuridica dell'Intelligenza Artificiale".
[484] 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*".
[485] A. LONGO E G. SCORZA, op. cit., Capitolo 7 (pag. 198 – 240) "Quali soluzioni: i diritti a rischio e gli strumenti giuridici a tutela".

artificiale in cui la condotta è opera dell'azione condivisa tra umano e *robot*.[486] 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.

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*)?[487]

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.[488]

Analizziamone il contenuto e le critiche.

Il primo modello, anche detto "*the perpetration-via-another liability model",* 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

[486] V. MANES, op. cit., Paragrafo 1.1 (pag. 2 – 5) "L'imputazione della responsabilità penale per azioni (autonome) delle macchine"
[487] M. B. MAGRO, "*Biorobotica, robotica e diritto penale*", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali".
[488] G. HALLEVY, op. cit., Paragrafo 1 (pag. 171 – 175) "Introduction: The Legal Problem".

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.

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. [489]

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. [490]

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 avrebbe tenuto. Si rimprovera al soggetto il fatto-reato che lo stesso avrebbe dovuto prevedere al fine di evitarlo.

---

[489] G. HALLEVY, op. cit., Paragrafo 3 (pag. 177 – 194) "Three models of the criminal liability of artificial intelligence entities".
[490] I. SALVADORI, op. cit., Paragrafo 5.1, "La responsabilità a titolo doloso".

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.[491]

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.). [492] 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 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.[493]

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

[491] G. HALLEVY, op. cit., Paragrafo 3 (pag. 177 – 194) "Three models of the criminal liability of artificial intelligence entities".
[492] I. SALVADORI, op. cit., Paragrafo 5.2, "La responsabilità colposa".
[493] 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".

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:[494] 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.

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.

Analizziamo i vari elementi.

### *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

[494] G. HALLEVY, op. cit., Paragrafo 3 (pag. 177 – 194) "Three models of the criminal liability of artificial intelligence entities".

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. [495]

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. [496]

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.[497]

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

[495] 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".
[496] 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".
[497] B. PANATTONI, op. cit., Paragrafo 5 "Le sfide poste dall'«emergence»: il responsibility gap".

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*.

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[498] 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. [499]

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 dall'ambiente esterno immettendo dentro di sé *input* raccolti senza alcun

[498] 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

[499] B. PANATTONI, op. cit., Paragrafo 5.2 "Responsabilità colposa degli operatori per i comportamenti emergenti dei sistemi di IA".

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*. [500]

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. [501] L'utilizzatore dunque, non può conoscere in anticipo il loro comportamento ma tale imprevedibilità è però prevedibile dal programmatore e dunque, da lui limitabile.[502] 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. [503] 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.[504]

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 decisionale e generando soluzioni inaspettate al problema delegatogli, sfuggono, quindi, al controllo umano. Dunque, sussistono problematiche in relazione alla ripartizione della responsabilità.[505]

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.

---

[500] E. PALMERINI, op. cit., Paragrafo 6 "I problemi della sicurezza e della responsabilità nel mercato emergente della robotica".
[501] C. PIERGALLINI, op. cit., Paragrafo 4 "Machine learning e diritto penale".
[502] M. B. MAGRO, "*Biorobotica, robotica e diritto penale*", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali".
[503] C. PIERGALLINI, op. cit., Paragrafo 4 "Machine learning e diritto penale".
[504] 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'*"".
[505] B. PANATTONI, op. cit., Paragrafo 2 "Dall'automazione tecnologica all'autonomia artificiale".

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.[506]

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 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.[507] 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

[506] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".
[507] S. QUINTARELLI, op. cit., Capitolo 4 (pag. 72 – 92) "L'Etica dell'Intelligenza artificiale".

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.[508]

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 nell'ambito della riconnessione delle azioni della macchina alla *suitas* dell'essere umano da ritenere responsabile.[509]

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.[510] 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. [511]

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[512] 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. [513]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

[508] S. QUINTARELLI, op. cit., Capitolo 6 (pag. 106 – 117) "Le leggi in gioco con l'intelligenza artificiale".
[509] B. PANATTONI, op. cit., Paragrafo 3.3 "Modelli di responsabilità e diversi livelli di autonomia: il "problema del controllo"".
[510] G.UBERTIS, op. cit., Paragrafo 8 (pag. 83 – 84) "Necessità di un controllo significativo".
[511] S. QUINTARELLI, op. cit., Capitolo 4 (pag. 72 – 92) "L'Etica dell'Intelligenza artificiale".
[512] 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
[513] B. PANATTONI, op. cit., Paragrafo 3.3 "Modelli di responsabilità e diversi livelli di autonomia: il "problema del controllo"".

cognizione da parte dell'utente del 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.[514]

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 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

[514] S. QUINTARELLI, op. cit., Capitolo 4 (pag. 72 – 92) "L'Etica dell'Intelligenza artificiale".

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.[515] 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. [516]

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.

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. [517]

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.[518] I programmatori dovranno realizzare i sistemi di intelligenza

---

[515] C. PIERGALLINI, op. cit., Paragrafo 4 "Machine learning e diritto penale".
[516] 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".
[517] S. QUINTARELLI, op. cit., Capitolo 4 (pag. 72 – 92) "L'Etica dell'Intelligenza artificiale"
[518] *Idem*.

artificiale (soprattutto per quelli "ad alto rischio" secondo la Proposta di Regolamento europeo del 21 aprile 2021)[519] che garantiscano una sufficiente trasparenza permettendo agli utenti di utilizzarli in modo appropriato al fine di permettere una supervisione ed un effettivo controllo umano.[520]

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.[521]

La figura del controllore umano opera come tutela per impedire che il malfunzionamento della macchina arrechi danni evitabili, altrimenti sarà lui a 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.[522] 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.[523]

### *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

[519] 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
[520] B. PANATTONI, op. cit., Paragrafo 3.3 "Modelli di responsabilità e diversi livelli di autonomia: il "problema del controllo"".
[521] E. PALMERINI, op. cit., Paragrafo 6.3 "Scenari presenti e futuri".
[522] C. PIERGALLINI, op. cit., Paragrafo 3 "Danno da prodotto e intelligenza artificiale sotto il controllo umano".
[523] S. QUINTARELLI, op. cit., Capitolo 6 (pag. 106 – 117) "Le leggi in gioco con l'intelligenza artificiale".

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). [524]

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 accertare se l'azione doverosa che si è omesso di compiere avrebbe impedito la realizzazione del fatto concreto da parte dell'autore.[525]

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.[526]

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.

---

[524] 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".
[525] 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".
[526] 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*[527] esprime il principio secondo cui il giudice deve valutare in concreto quali fossero gli strumenti a disposizione dei soggetti per prevenire il reato: 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.[528]

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

---

[527] 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

[528] 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".

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),[529] 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[530] 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: 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. [531]

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[532] 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.[533] Allo stesso tempo, il

---

[529]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

[530] 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

[531] B. PANATTONI, op. cit., Paragrafo 6 "Responsabilità diffuse eLegal Protection by Design".

[532] 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.

[533] 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

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.[534]

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. [535]

La concezione etica del rischio graduata in base alla tollerabilità sociale dovrà indirizzare le future scelte politiche al fine di governare tali situazioni. [536]

### *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.

---

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).

[534] B. PANATTONI, op. cit., Paragrafo 6 "Responsabilità diffuse eLegal Protection by Design".

[535] I. SALVADORI, op. cit., Paragrafo 6, "Danno da agente artificiale e responsabilità plurisoggettiva".

[536] C. PIERGALLINI, op. cit., Paragrafo 2 "Danno da prodotto e Intelligenza Artificiale".

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. [537]

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. [538] 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 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[539] ponga gli stessi dubbi che si posero nel considerare il software come bene mobile. [540]

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 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

---

[537] 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".

[538] 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.

[539] PONZANELLI, "*Responsabilità per danno da computer: alcune considerazioni comparative*", in Resp. Civ. prev., 1991, 650.

[540] 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).

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. [541]

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. [542] Allo stesso tempo, parte della dottrina[543] 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.[544]

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 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[545] del 21 aprile 2021. [546]

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.[547] 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

---

[541] B. PANATTONI, op. cit., Paragrafo 4 "La responsabilità penale per danno da prodotto".
[542] B. PANATTONI, op. cit., Paragrafo 5.4 "Possibili soluzioni al responsibility gap".
[543] 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..
[544] B. PANATTONI, op. cit., Paragrafo 5.4 "Possibili soluzioni al responsibility gap".
[545] 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
[546] B. PANATTONI, op. cit., Paragrafo 6 "Responsabilità diffuse eLegal Protection by Design".
[547] I. ASIMOV, "*I, Robot*", Mondadori, 1950.

da parte del programmatore che sfociano in una sua colpa eventuale dovuta ad un'astratta prevedibilità del rischio che non viene da lui limitato. [548]

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).[549]

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 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. [550]

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.[551]

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

---

[548] M.B. MAGRO, "*Biorobotica, robotica e diritto penale*", op. cit., Paragrafo 3 (pag. 7 – 15) "La robotica, i droni e le Intelligenze Artificiali".

[549] 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

[550] 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

[551] I. SALVADORI, op. cit., Paragrafo 9 "Considerazioni finali".

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.[552]

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. [553]

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 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".

[552] B. PANATTONI, op. cit., Paragrafo 4 "La responsabilità penale per danno da prodotto".
[553] B. PANATTONI, op. cit., Paragrafo 5.4 "Possibili soluzioni al responsibility gap".

# ***CAPITOLO 3:***

# ***L'INTELLIGENZA ARTIFICIALE E CRIMINALITÀ***

---

**SOMMARIO**: 1. Le *self-driving cars*: lesioni e omicidio colposo stradale – 1.1 Le auto a guida autonoma – 1.2 Gli errori commessi dalle auto a guida autonoma e i soggetti responsabili – 1.3 L'omicidio stradale e le lesioni in relazione all'uso dei veicoli autonomi: profili penalistici – 2. L'intelligenza artificiale e la responsabilità dell'operatore sanitario – 2.1 I *cyborg* e il potenziamento umano – *2.2 I robot* chirurghi – 2.3 La responsabilità medica in relazione all'utilizzo dei sistemi intelligenti: profili penalistici – 3. *Cybercrime* e i reati informatici "in senso ampio" – 3.1 Le *chatbot* e i reati di odio e discriminazione – 3.2 Le *fake-news* e gli algoritmi intelligenti – 3.3 Il fenomeno del *deep-fake* e la tutela dell'immagine – 3.4 La pedopornografia e il *revenge porn* ai tempi del *deep-fake* – 3.5 Lo stupro e l'abuso sessuale di "minori robotici"

---

## ***1. Le self-driving cars: lesioni e omicidio colposo stradale***

Per quanto riguarda l'ambito penale, il settore dell'intelligenza artificiale risulta relativamente nuovo ma destinato ad assumere sempre maggiore rilevanza dato lo sviluppo di tali tecnologie. Si sono già verificati casi relativi alle azioni lesive realizzate dall'intelligenza artificiale come gli incidenti stradali dovuti alle macchine a guida autonoma o ai messaggi di odio diffusi dai *social bots*. Ciò porta ad individuare un gruppo distinto di reati: gli *Artificial Intelligence Crimes*. Gli illeciti nei quali può incidere l'intelligenza artificiale possono interessare diversi settori del diritto penale. Invero, i gruppi di reati maggiormente rilevanti riguardano: i reati relativi ai mercati finanziari (grazie agli agenti algoritmici di *trading*); i traffici internazionali di stupefacenti o di altri prodotti illeciti (utilizzando droni a guida autonoma per consegnare i carichi di droga); le frodi e i reati informatici in senso stretto compresa la violazione della protezione dei dati personali (grazie a tecnologie di intelligenza artificiale per lo sviluppo di nuovi *malware* o di strumenti di *social engineering*); i reati contro la persona soprattutto per quanto riguarda forme di veicolazione dei contenuti illeciti (attraverso *deep-fakes* o *social bots*) e reati di lesioni e omicidio (tramite macchine a guida autonoma o *robot* medici).

Le aree penali di interesse sono di due tipologie: la realizzazione dei reati commessi attraverso o contro i sistemi di intelligenza artificiale (avrà ad oggetto principalmente condotte dolose); i reati commessi

direttamente dagli agenti artificiali (rimproverabili a titolo di colpa agli operatori che si occupano della programmazione, sviluppo, produzione e utilizzo della macchina).[554]

Per quanto riguarda i reati commessi attraverso o contro i sistemi di intelligenza artificiale, si tratta di fattispecie in cui l'evoluzione tecnologica e l'impiego della macchina ha comportato il ricorso a nuove modalità di commissione dei reati già tipizzati o ne ha creati di nuovi. Esistono infatti proposte dottrinali volte a criminalizzare le intelligenze artificiali realizzate al fine di commettere reati. Bisogna però valutare se i *cybercrimes* già esistenti siano applicabili alle fattispecie concrete e colmare i possibili vuoti di tutela. Ci troviamo nella stessa fase attraversata nel momento dell'introduzione dei *computer crimes.*[555]

In relazione, invece, al caso dei reati commessi direttamente dall'intelligenza artificiale come autore dell'illecito bisogna valutare possibili responsabilità in capo ai soggetti dietro l'intelligenza artificiale. Bisogna individuare regimi di imputazione della responsabilità in tali casi. La Proposta di Regolamento trattata nel primo capitolo stabilisce obblighi giuridici proporzionati al grado di rischio che connota i diversi sistemi di intelligenza artificiale in modo tale da anticipare la tutela. Tali obblighi dovranno poi armonizzarsi con le normative riguardanti la responsabilità penale per danno da prodotto. I possibili rischi relativi all'intelligenza artificiale riguardano il caso in cui il funzionamento dei *robot* generi offese a beni giuridici protetti facenti capo a terzi che interagiscono con la macchina oppure il caso in cui l'autonomia del sistema generi comportamenti emergenti. Il secondo non è un rischio legato direttamente all'autonomia di funzionamento della macchina ma connesso alle abilità che i programmatori decidono di attribuire alla macchina e che gli permettono di esercitare. Dunque i reati potranno essere connessi a difetti presenti nei prodotti, ad azioni negligenti dei programmatori o a sviluppi di *outputs* inaspettati che la macchina è in grado di realizzare grazie alle sue capacità di autoapprendimento. Come nel caso del *Random Darknet Shopper* (sistema algoritmico svizzero realizzato nel 2014 con il compito di comprare casualmente dal *deep web* oggetti per una mostra a Zurigo, ma nello svolgimento del compito acquistò inoltre droghe illegali andando quindi a costituire un reato).[556]

Per questioni di razionalità, non sarà possibile in tale elaborato analizzare tutti i possibili reati in cui può fare ingresso l'intelligenza artificiale: si tratta di ambiti estremamente interessanti e variegati che meriterebbero, in una sede a parte, un maggiore approfondimento. Di conseguenza, tale tesi mirerà ad analizzare esclusivamente gli ambiti penali relativi ai veicoli a guida autonoma, alla colpa medica in caso di utilizzo di sistemi di intelligenza artificiale e, infine, i *cybercrime,* in senso stretto (che prevedono la macchina come elemento essenziale del fatto tipico) ma soprattutto in senso ampio (che prevedono il sistema informatico

[554] 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 3 "Gli albori di una nuova disciplina: il «diritto penale dell'IA»".

[555] B. PANATTONI, "*Intelligenza artificiale: le sfide per il diritto penale nel passaggio dall'automazione tecnologica all'autonomia artificiale*"op. cit., Paragrafo 3.1 "Reati commessi attraverso o contro i sistemi di IA".

[556] B. PANATTONI, "*Intelligenza artificiale: le sfide per il diritto penale nel passaggio dall'automazione tecnologica all'autonomia artificiale*", op. cit., Paragrafo 3.2 "Sistema di IA quale «autore» del reato e responsabilità degli operatori".

quale elemento eventuale), in cui l'utilizzo di algoritmi di intelligenza artificiale può generare forti lesioni nella sfera personale degli individui.

Nel trattare queste fattispecie, il nodo centrale sarà l'effettività di una possibile responsabilità dell'operatore per omesso impedimento dell'evento lesivo.[557]

Il tema, per quanto possa sembrare lontano e intangibile, è in realtà sempre più vicino: la tecnologia avanza e l'intelligenza artificiale, entrata oramai da decenni nelle nostre vite quotidiane, si sta facendo sempre più spazio, diventando quasi irrinunciabile. I *robot* e i *computer* stanno sostituendo le attività umane, le forme di delegazione ai sistemi intelligenti stanno diventando sempre più incisive ed impattanti per le vite degli individui. Ovviamente, i benefici che porta con sé sono notevoli: l'accessibilità alle fonti di informazioni, la rapidità nelle ricerche su *Internet* grazie ad algoritmi in grado di predire le scelte dell'utente, la riduzione dei tempi di analisi dei dati statistici tramite banche dati analizzabili da algoritmi in pochi decimi di secondo, e l'interconnessione tra i vari sistemi domestici che permette una migliore esperienza di utilizzo grazie all'*Internet of Thing*.

L'elenco dei vantaggi è, dunque, molto lungo, ma, purtroppo lo è anche quello dei rischi. Infatti, le nuove tecnologie possono incidere sui diritti fondamentali degli individui generando lesioni ai loro beni giuridici tutelati. Ciò è dovuto, come abbiamo già sottolineato, al fatto che le macchine percepiscono il mondo esterno in modo diverso dall'essere umano, possedendo esclusivamente una comprensione superficiale in quanto agiscono in modo probabilistico.

Tale capitolo ha il compito di analizzare il concetto di *dual-use technology* applicato all'intelligenza artificiale; vale a dire la capacità di essere utilizzata sia a fin di bene che per la realizzazione di illeciti. Dunque ci soffermeremo sull'analisi delle fattispecie di reato connesse all'utilizzo di automobili a guida autonoma, di *robot* chirurghi, di *chatbot* e sulla realizzazione di illeciti tramite i sistemi di intelligenza artificiale come il fenomeno del *deep-fake*.

### *1.1 Le auto a guida autonoma*

Il mondo del lavoro e delle telecomunicazioni è radicalmente cambiato grazie all'avvento della tecnologia. Ma non ci si fermerà qui: infatti, presto potrebbe cambiare anche il nostro tradizionale modo di guidare grazie all'ingresso dell'intelligenza artificiale nel mondo automobilistico. Per ora il fenomeno è ancora in via di sviluppo e le poche funzioni di guida assistita già immesse sul mercato sono destinate ai modelli più costosi di autovetture.[558]

[557] *Idem.*
[558] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".

Prima di passare ad un'analisi prettamente penalistica della materia, analizziamo, preliminarmente e senza pretesa di completezza tecnica, il mondo dell'automazione dei veicoli.

Essa può coinvolgere una o più parti dello stesso e delle sue funzioni di guida ed in base al grado di automazione sarà possibile osservare un coinvolgimento maggiore o minore dell'essere umano (guidatore esperto, controllore attento ed infine, mero passeggero distratto). La complessità della categorizzazione viene ridotta grazie all'adozione, nel settore automobilistico, di *standard* di automazione dei singoli veicoli.[559] Il metodo più diffuso è quello realizzato dalla *SAE International* (l'ente internazionale che regola l'industria aereospaziale e automobilistica) nel 2014 e rivisto nel 2018: una scala di sei livelli per valutare il grado di automazione di ogni autovettura denominata *standard J3016*.

Al livello zero (*No Driving Automation*), rientrano le auto tradizionali senza alcuna tipologia di guida assistita: infatti, tutte le funzioni sono a carico dell'essere umano. Al primo livello (*Driver Assistance*) è possibile trovare ausili come il *cruise control,* l'*ESC* (il controllo elettronico di stabilità), il supporto dinamico di frenata, i sistemi di mantenimento della corsia di marcia e il parcheggio assistito ma il guidatore umano rimane fortemente attivo. Al secondo livello (*Partial Driving Automation*) si ha, invece, una parziale automazione in base alla quale l'utente può delegare alla macchina il controllo di azioni limitate (come la *Tesla S* e la *Mercedes Classe E*).[560] Tali due livelli sono accumunati da una costante presenza di un controllore umano che deve risultare un guidatore attivo e attento dato che il veicolo autonomo non va a sostituirlo ma solo ad assisterlo: proprio per questo sono autorizzate a circolare in base alla normativa[561] vigente.[562] Dal terzo livello (*Conditional Driving Automation*) si giunge ad una guida autonoma limitata, in base alla quale, la macchina può essere investita del pieno controllo delle funzioni di sicurezza in alcune situazioni particolari.[563]A tale livello, la macchina è in grado di auto-condursi continuando a mantenere i comandi di guida manuale e la figura del guidatore che deve necessariamente intervenire in caso di emergenza. Siamo approdati nel mondo delle *self-driving cars* vere e proprie, anche se ancora non pienamente autonome (definibili infatti come vetture semi-autonome).[564] Al quarto livello (*High Driving Automation*) si arriva ad un'automazione elevata in base alla quale la macchina può svolgere tutte le funzioni di sicurezza monitorando

---

[559] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" (pag. 325 – 353) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018. In Diritto penale Contemporaneo, Rivista trimestrale 2/2019. Paragrafo 2 (pag. 327 – 328) "Dalle auto semi-autonome a quelle totalmente *self-driving*: i "livelli" di automazione".

[560] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".

[561] *Codice della Strada*, D.Lgs. n. 285 del 30 aprile 1992 ed entrato in vigore il 1 gennaio 1993; articolo 46 "Nozione di veicolo": https://www.normattiva.it/atto/caricaDettaglioAtto?atto.dataPubblicazioneGazzetta=1992-05-18&atto.codiceRedazionale=092G0306&atto.articolo.numero=0&atto.articolo.sottoArticolo=1&atto.articolo.sottoArticolo1=10&qId=8c6b76a4-7cdb-4062-834b-fcd8f66378c0&tabID=0.03217102139088101 6&title=lbl.dettaglioAtto

[562] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 2 (pag. 327 – 328) "Dalle auto semi-autonome a quelle totalmente *self-driving*: i "livelli" di automazione".

[563] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".

[564] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 2 (pag. 327 – 328) "Dalle auto semi-autonome a quelle totalmente *self-driving*: i "livelli" di automazione".

le condizioni della strada[565] (essendo capace di affrontare la grande maggioranza degli scenari possibili) [566] senza la necessità di un controllo umano costante (tranne in situazioni meteorologiche e di traffico estreme). [567] Nonostante tale abilità, il veicolo possiede ancora i comandi manuali permettendo all'essere umano di intervenire con il volante e i pedali in situazioni particolari.[568] L'ultimo livello (*Full Driving Automation*) permette un'automazione completa in cui l'utente ha il solo compito di impostare la destinazione, potendosi poi anche distrarre dato che il veicolo viene interamente guidato dall'agente artificiale. [569] A tale livello (l'unico considerabile come interamente autonomo) scompaiono del tutto i comandi per la guida manuale dato che la vettura ha l'abilità di affrontare ogni scenario possibile: dunque, l'utente non sarà più un conducente ma un semplice passeggero del veicolo. [570]

In commercio sono disponibili esclusivamente i veicoli fino al terzo livello i cui sistemi di guida semi-autonomi (tra i più famosi*, Apple*, *Uber e Waymo,* impresa del gruppo *Alphabet* facente capo a *Google*) si basano su un *software* e dei sensori (solitamente *laser* e *radar* chiamati *LiDAR*) di apprendimento automatico in modo tale da permettere all'autovettura di comprendere a pieno il mondo circostante potendosi così muovere autonomamente. L'utilizzo dei *laser* non risulta necessario, infatti *Tesla* e *Nissan* preferiscono affidarsi esclusivamente a fotocamere e sensori ad ultrasuoni che forniscono una visibilità a 360° intorno all'auto per un raggio di 250 metri rilevando oggetti anche attraverso la pioggia e la nebbia. Un'ulteriore differenza caratterizza il sistema della *Cadillac*, dal 2019 disponibile sul mercato, che permette una guida di terzo livello a mani libere. Permettendo una guida autonoma solo sulle autostrade che vengono mappate con precisione quasi assoluta. Ciò non toglie l'obbligo del conducente di osservare la strada e l'operato dell'autovettura.

Nonostante alcune marginali differenze, tutti i sistemi utilizzano l'intelligenza artificiale il cui algoritmo viene addestrato inizialmente in un ambiente di simulazione (che permette al sistema di testare le sue capacità in contesti particolari); successivamente, per ridurre i c.d. *blind-spot* (situazioni non replicabili dal simulatore), su strade comuni e affiancando il sistema ad un pilota umano che insegni al veicolo come comportarsi in situazioni reali.[571] In questo modo si insegna al pilota automatico come controllare la velocità, il cambio di corsia, la distanza dagli oggetti, il percorso più rapido e come garantire la completa sicurezza di viaggio. [572]

Bisogna però sottolineare che, nonostante siano in commercio solo i sistemi fino al terzo livello di automazione, esiste un unico servizio di quarto livello autorizzato: quello di *Waymo*. Si tratta di *taxi* disponibili

---

[565] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".
[566] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 2 (pag. 327 – 328) "Dalle auto semi-autonome a quelle totalmente *self-driving*: i "livelli" di automazione".
[567] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".
[568] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 2 (pag. 327 – 328) "Dalle auto semi-autonome a quelle totalmente *self-driving*: i "livelli" di automazione".
[569] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".
[570] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 2 (pag. 327 – 328) "Dalle auto semi-autonome a quelle totalmente *self-driving*: i "livelli" di automazione".
[571] N. RUGGIERO, "*Auto a guida autonoma, ecco le frontiere (tra reti neurali e 5G)*", in *Agendadigitale.eu*, 20 marzo 2019: https://www.agendadigitale.eu/infrastrutture/auto-a-guida-autonoma-ecco-le-frontiere-attuali-tra-reti-neurali-e-5g/
[572] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".

solo a Phoenix (Arizona) su strade selezionate e solo per soggetti (*tester* addestrati allo scopo) prescelti. Quindi, è possibile affermare che, ad oggi, si sia giunti esclusivamente al quarto livello di automazione concesso solo in circostanze specifiche in percorsi predefiniti e con un utente umano pronto a riprendere il controllo: ma tali veicoli sono ancora in fase di sperimentazione. [573] Nonostante ciò, secondo *PWC*, dal 2025 saranno disponibili le auto completamente autonome di quinto livello e raggiungeranno i 12 milioni di unità vendute nel 2030.[574]

L'evoluzione sarà rapida e questo è anche dovuto alla volontà di raggiungere funzionalità straordinarie grazie ai sistemi di intelligenza artificiale. Infatti, essi, oltre al "semplice" guidare, grazie ai numerosi sensori interni (posizionati sul cruscotto) ed esterni all'auto, possono anche valutare attentamente le condizioni psicofisiche del conducente (come stanchezza, malessere e distrazione) e le condizioni tecniche del veicolo (come l'usura o i guasti).[575] Tanto è vero che le *concept car* di *BMW* e *Mercedes* sono in grado di prevedere i movimenti e i comandi dell'utente prevedendo il suo livello di *stress* e dunque, la sua idoneità alla guida potendo prendere il controllo in caso di sonno o malessere e di restituire i comandi di guida all'utente in scenari di circolazione stradale non noti al sistema. Tutto ciò, grazie alla combinazione delle informazioni che riescono ad acquisire sui movimenti del corpo del conducente (come mani, occhi e volto) e ai comandi vocali con cui lo stesso può impartire ordini.[576]

Inoltre, grazie al riconoscimento facciale, tali sistemi permettono di identificare il conducente e modificare le caratteristiche del veicolo (come la musica, il riscaldamento, il percorso preimpostato e la distanza dei sedili). Per di più, essendo connesso anche al calendario digitale dell'utente, il veicolo ha la capacità di seguire gli appuntamenti dello stesso e di portarlo alla destinazione senza che sia lui a indicarla. Una volta giunta a destinazione e una volta che l'utente sarà sceso dalla vettura, inserendo la modalità parcheggio, sarà essa stessa a cercarlo e ad immettercisi comunicando tramite telefono con il proprietario, il quale potrà a sua volta richiamarla a sé tramite l'applicazione apposita.[577]

Queste ed altre funzionalità permettono di raggiungere vantaggi formidabili come l'impossibilità di commissione di errori da parte dell'essere umano[578] (infatti, il 90% dei casi di incidenti sono dovuti a comportamenti inappropriati del conducente connessi a volontarie violazioni del Codice della strada oppure a distrazioni)[579] aumentando così la sicurezza alla guida e riducendo il numero di incidenti.[580] Infatti, le auto a guida autonoma non sono soggette a distrazioni o a violazioni normative; esse sono appositamente progettate per rispettare il Codice della strada e per avere abilità superiori rispetto agli esseri umani, potendosi accorgere

---

[573] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".
[574] L. PRINCIPALI E D. SALERNO, "*5G e mobilità, ecco le innovazioni per le auto (in arrivo anche in Italia)*", in *Agendadigitale.eu*, 21 novembre 2018: https://www.agendadigitale.eu/infrastrutture/5g-e-mobilita-ecco-le-innovazioni-per-le-auto-in-arrivo-anche-in-italia/
[575] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".
[576] N. RUGGIERO, op. cit.
[577] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".
[578] *Idem.*
[579] L. BUTTI, "*Le auto guideranno da sole, ma con quali responsabilità?"*, in Il Bo Live, Università di Padova, 9 novembre 2018: https://ilbolive.unipd.it/it/news/auto-guideranno-sole-quali-responsabilita
[580] A. LONGO E G. SCORZA, op. cit., Capitolo 5 (pag. 168 – 183) "L'AI nella quotidianità: a casa, in automobile".

di situazioni che l'utente non riuscirebbe a percepire con la stessa rapidità e prontezza.[581] Inoltre, con l'aumentare dell'autonomia dei veicoli, verrebbero meno i rischi dovuti alla guida sotto l'influenza dell'*alcol* o di sostanze stupefacenti sanzionata dagli articoli 186, 186 *bis* e 187 del Codice della strada. Infatti, per le vetture di quinto livello, è ipotizzabile l'inserimento di un comando di *drunk driving* che permetterebbe al veicolo di trasportare l'utente sotto l'effetto di sostanze fino alla destinazione desiderata senza che sia lui a guidare rendendo di fatto inapplicati gli articoli citati senza rischio per la tutela di possibili vittime. Problemi però rimarrebbero nel caso di veicoli semi-autonomi in cui è comunque necessario un costante controllo del guidatore sulle operazioni realizzate dal veicolo: in tal caso, si rientrerebbe comunque nelle fattispecie di reato dato che il soggetto è comunque considerato come un conducente. Sicuramente i rischi di lesioni verrebbero ridotti ma non si uscirebbe da un'incriminazione penale. È auspicabile dunque l'introduzione di specifiche normative o l'estensione dell'applicazione degli articoli già vigenti.[582]

Quindi, le auto completamente autonome, come affermato precedentemente, trasformano il conducente in un mero passeggero[583] permettendo un'estensione del servizio anche a soggetti privi di patente o inabili alla guida (*drunk driving*) e si ridurrà lo *stress* connesso alla guida.[584] Ed inoltre, grazie alla loro capacità di condividere informazioni anche con altri veicoli autonomi a loro connessi si potrà generare un bacino di conoscenza collettiva sfruttabile da tutti i sistemi interconnessi per accrescere il loro numero di situazioni note su cui agire aumentando la propria esperienza di guida.[585] Infatti, tramite tale cooperazione tra i sistemi si riuscirebbe ad incrementare l'efficienza della circolazione (che diventerebbe più fluida grazie alle capacità sensoriali e a tempi di reazione straordinari) che genererebbe una riduzione dei tempi di trasporto e un risparmio di carburante a beneficio, inoltre, della tutela dell'ambiente.[586]

Fino ad ora, tuttavia, i *test* mostrano che le autovetture a guida autonoma realizzate guidano con un'eccessiva prudenza, rallentando difronte ad ostacoli minimi e generando un viaggio con troppi cambi di velocità.[587] Questo è in parte dovuto alla complessità del mondo reale: i nostri sensi riescono a percepire e a gestire quotidianamente le sfide che la realtà ci pone; i sistemi intelligenti meno. Infatti, in ambienti protetti, le difficoltà riscontrate dalle auto a guida autonoma sono nettamente inferiori. Basti pensare ai veicoli dotati di intelligenza artificiale che vengono adottati in miniera per trasportare le estrazioni da un luogo all'altro della stessa grazie a sistemi *GPS* e a sensori che percepiscono l'ambiente circostante. Inoltre, è interessante raccontare della prima sfida, su lunga distanza (150 *km*) per veicoli interamente autonomi, svoltasi nel deserto del Mojave nel 2004 e finanziata dal Dipartimento della difesa degli Stati Uniti. Alla prima edizione nessuno riuscì a tagliare il traguardo mentre nella successiva, ben cinque autovetture (su ventitré) interamente

[581] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 3.1 (pag. 329 – 331) "I benefici sociali".
[582] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 6 (pag. 341 – 343) "*Autonomous driving* e reati stradali in senso stretto".
[583] I. SALVADORI, op. cit., Paragrafo 3, "Dall'automazione all'autonomia: classificazione degli agenti artificiali".
[584] L. BUTTI, op. cit.
[585] S. QUINTARELLI, op. cit., Capitolo. 6 (pag. 106 – 117) "Le leggi in gioco con l'intelligenza artificiale".
[586] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 3.1 (pag. 329 – 331) "I benefici sociali".
[587] A. LONGO E G. SCORZA, op. cit., Capitolo 5 (pag. 168 – 183) "L'AI nella quotidianità: a casa, in automobile".

autonome riuscirono a raggiungerlo; a vincere fu il veicolo della *Stanford University* realizzato da Sebastian Thrun (ora capo di *Waymo*).

In questi esempi, però, non ci troviamo in presenza di pedoni che possono attraversare la carreggiata o di ostacoli nel percorso ed inoltre, le modalità di movimento sono definite a priori. La strada per giungere ad una completa automazione dei veicoli risulta sicuramente tortuosa. Infatti, per realizzare autovetture di quinto livello si necessita di una straordinaria capacità computazionale, dato che, per permettere una guida completamente autonoma, i dati, che vengono immagazzinati dai sensori del veicolo e che devono essere analizzati in tempo reale per permettere un viaggio sicuro, superano i 5 *TB* (*terabyte*) l'ora. Inoltre, devono essere prese in considerazione decisioni non previste dal programmatore, grazie al meccanismo del *machine learning* (fondamentale in un ambiente così dinamico e versatile come le strade urbane). [588]

Ciò comporta sicuramente dei rischi, dovuti principalmente alla sicurezza stradale.

Infatti, per quanto sia vero che i sinistri diminuirebbero drasticamente a seguito dell'eliminazione del fattore umano, essi non potrebbero del tutto essere eliminati, soprattutto in relazione a quelli derivanti dall'esterno e dunque a prescindere dal tipo di vettura utilizzata. Per di più, un margine di rischio dovuto di per sé alla guida autonoma risulta inevitabile e può essere previsto a livello statistico generale ma mai concreto in relazione ai singoli casi. [589] In aggiunta, finché tali veicoli autonomi viaggeranno insieme ai veicoli tradizionali, i rischi di collisione resteranno, perché dovuti, soprattutto, al comportamento umano che, oltre al ragionamento razionale associa quello intuitivo-passionale che non può essere previsto dalla macchina intelligente. In una fase di completa automazione del traffico, invece, si dovrebbe giungere ad una maggiore sicurezza stradale. [590]

In aggiunta a tali problematiche, si inseriscono i pericoli legati alla connessione ad *Internet* di tali veicoli autonomi. Essi, essendo interamente connessi, risultano infatti soggetti a possibili attacchi criminali che, entrando nel sistema del veicolo possono fargli commettere crimini (come, ad esempio, furti, rapine, omicidi o attentati). [591] Tali rischi sono imprevedibili fino alla loro realizzazione in concreto e consistono nell'*hacking* dei sistemi di *cybersecurity* delle *self-driving cars.* [592]

Tale ambito criminoso risulta delicato: l'*hacker* va a sfruttare le pecche del programma di connessione *wireless* del veicolo allo *smartphone* del conducente o ad altri veicoli o dispositivi, agendo da remoto. In tal modo, l'*hacker* prende il possesso di tutte le funzioni del veicolo potendolo dirigere a distanza. Queste azioni criminali mettono a rischio la sicurezza stradale e del guidatore (prigioniero di un veicolo nelle mani di un *hacker*).[593]

---

[588] S. QUINTARELLI, op. cit., Capitolo 2 (pag. 33 – 54) "Intelligenza artificiale. Applicazioni e tecnica".
[589] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 3.2 (pag. 331 – 334) "I profili di rischio, tra percezione sociale e logiche di precauzione".
[590] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 3.1 (pag. 329 – 331) "I benefici sociali".
[591] A. LONGO E G. SCORZA, op. cit., Capitolo 5 (pag. 168 – 183) "L'AI nella quotidianità: a casa, in automobile".
[592] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 3.2 ((pag. 331 – 334) "I profili di rischio, tra percezione sociale e logiche di precauzione".
[593] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 7 (pag. 343 – 346) "Nuove fenomenologie criminali e nuove esigenze di tutela: in particolare, il nodo della *cybersecurity".*

Per quanto concerne la *privacy* infatti, potranno essere analizzati i luoghi visitati, le abitudini di guida, le infrazioni commesse, gli stati fisici di sonnolenza o alterazione e le registrazioni delle conversazioni avvenute in auto. Ciò esporrebbe gli utenti ad aggressioni criminali (o ad un eccessivo controllo da parte dalle forze dell'ordine nel caso di utilizzo di *trojan* autorizzati).[594]

L'interconnessione tra veicoli autonomi di cui si faceva cenno precedentemente, permetterebbe ai criminali di organizzare attentati "*cyberterroristici*" su larga scala. Ovviamente, tale intromissione abusiva, da un punto di vista penalistico, integrerebbe il reato di accesso abusivo a sistema informatico (*ex* art. 615 *ter* c.p.), infatti alcuni[595] hanno sostenuto la necessità di inserire nel nostro ordinamento una fattispecie *ad hoc* a seguito della gravità degli effetti che possono derivare dall'accesso abusivo ad un sistema di intelligenza artificiale. Sicuramente, sarà necessario implementare gli *standard* di *cybersecurity* specifici per tali ipotesi.[596]

Il concetto di rischio non è assoluto e va connesso alla tollerabilità accettata dalla società e dunque è sottoposto ad una valutazione politica del bilanciamento rischi-benefici dati dall'introduzione dell'intelligenza artificiale nel campo automobilistico. Tale valutazione non segue esclusive logiche scientifico-oggettive ma si basa soprattutto su pregiudizi (talvolta irrazionali) diffusi in un determinato periodo storico in quella specifica società presa in considerazione. A riprova della difficoltà di valutazione politica sull'introduzione totale di sistemi di guida completamente automatizzati, è possibile analizzare varie ricerche statistiche che dimostrano la diffidenza di gran parte degli individui nell'affidare la propria incolumità ad algoritmi famosi per la loro oscurità (un paradosso se sottolineiamo le volte in cui gli stessi soggetti hanno deliberatamente scelto di affidarsi ad autisti pericolosi a causa del loro alto grado di distrazione o di errore). Di fatto, inserendo algoritmi, connessi al criterio etico del minor danno, si verrebbe esposti, in caso di necessità, al rischio di essere sacrificati dal veicolo, perché considerati il minor danno collaterale, in base alla situazione concreta. Problemi simili si porrebbero se l'algoritmo non si basasse su tale criterio ma fosse regolabile dall'utente in base ad un'etica differenziata sulla volontà del proprietario. [597]

Nonostante i rischi e le difficoltà, l'industria automobilistica crede nell'impresa impossibile: arrivare alle stesse prestazioni (se non addirittura superiori) che avrebbe il miglior automobilista umano. Siamo ancora lontani dalla totale autonomia ma grazie alla quantità di dati che viene raccolta costantemente per le sperimentazioni di tali sistemi si sta giungendo ad un'accelerazione dell'evoluzione tecnologica. Serviranno corposi investimenti e risorse per sviluppare i sistemi necessari per lo sviluppo dell'intelligenza artificiale necessaria per la guida interamente autonoma.[598]

---

[594] A. LONGO E G. SCORZA, op. cit., Capitolo 5 (pag. 168 – 183) "L'AI nella quotidianità: a casa, in automobile".
[595] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. (pag. 345).
[596] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 7 (pag. 343 – 346) "Nuove fenomenologie criminali e nuove esigenze di tutela: in particolare, il nodo della *cybersecurity*".
[597] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 3.2 (pag. 331 – 334) "I profili di rischio, tra percezione sociale e logiche di precauzione".
[598] N. RUGGIERO, op. cit.

Fin ora abbiamo parlato di autovetture, ma l'intelligenza artificiale può essere adottata anche per aeroplani o imbarcazioni. Infatti, il primo pilota automatico è stato inventato nel 1914 permettendo di mantenere l'aereo alla quota e nella direzione stabilite.

In aria, a causa delle perturbazioni e dei cambi di rotta, risulta sicuramente complesso realizzare un sistema di volo che assista i piloti delegando all'intelligenza artificiale alcune attività di controllo o di volo: il comandante umano, infatti, dovrà costantemente controllare che il volo proceda per il meglio. In ambito militare, quasi la totalità della crociera è automatizzata, mentre per gli aerei di linea non si è giunti a questo livello per semplici questioni di responsabilità anche se la tecnologia sta procedendo rapidamente al punto tale che nel gennaio 2020 un aereo civile è atterrato otto volte completamente guidato dal pilota automatico. Bisogna però affermare che in aria, per quanto le condizioni meteorologiche e l'altezza rendano la situazione estrema, non sussistono eventi improvvisi (se non in rare situazioni) e gli altri veicoli vengono segnalati con grande anticipo dai sensori di controllo.[599]

Superata questa breve parentesi connessa ai sistemi di volo automatizzati, è importante domandarci se sussista una regolamentazione dei sistemi di guida autonoma. Infatti, i *self-driving cars* si trovano in un'area grigia del diritto. La loro circolazione in strada può avvenire sulla base di norme introdotte appositamente che l'autorizzano in via sperimentale imponendo requisiti di sicurezza, disciplinando i profili di responsabilità e gli aspetti assicurativi.[600]

Per ora, la maggior parte della sperimentazione su strada dei veicoli autonomi è avvenuta negli Stati Uniti a partire dal 2014 (soprattutto in Nevada) predisponendo norme apposite per la loro circolazione, inizialmente con la necessaria presenza di un soggetto esperto pronto ad intervenire. Nel 2018 la California ha permesso la sperimentazione su strada di veicoli completamente autonomi.[601]

In Europa, la Convenzione di Vienna sul traffico stradale del 1968[602] esclude la circolazione di veicoli autonomi su strade pubbliche perché impone, all'articolo 8, un costante controllo del guidatore: "*Every moving vehicle or combination of vehicles shall have a driver* […] *Every driver shall at all times be able to control his vehicle* […]". Dunque, la completa automazione non è permessa. Risulta però accettata entro una certa soglia, grazie ad una modifica del testo avvenuta nel 2016, ove sia assicurata la presenza a bordo di una persona che monitori costantemente l'operato del veicolo e che possa assumerne immediatamente il controllo disattivando la guida autonoma.[603]

Nel luglio 2022 dovrebbero entrare in vigore le modifiche poste alla Convenzione stessa in base alle quali sarà forse ammesso per il conducente affidarsi al veicolo di terzo livello senza l'obbligo di mantenere le mani sul volante; il guidatore dovrà però restare costantemente vigile e in grado di riprendere il comando non

[599] S. QUINTARELLI, op. cit., Capitolo 2 (pag. 33 – 54) "Intelligenza artificiale. Applicazioni e tecnica".
[600] E. PALMERINI, op. cit., Paragrafo 4.1 "Dalla qualificazione generica alla lacuna normativa".
[601] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 1 (pag. 326 – 327) "Introduzione".
[602] *Convenzione di Vienna sulla circolazione stradale*, Conclusa a Vienna l'8 novembre 1968, Approvata dall'Assemblea federale il 15 dicembre 1978: https://fedlex.data.admin.ch/filestore/fedlex.data.admin.ch/eli/cc/1993/402_402_402/20200519/it/pdf-a/fedlex-data-admin-ch-eli-cc-1993-402_402_402-20200519-it-pdf-a.pdf
[603] E. PALMERINI, op. cit., Paragrafo 4.1 "Dalla qualificazione generica alla lacuna normativa".

appena il veicolo lo richieda. Infatti, le modifiche sottoposte a discussione internazionale riguardano l'articolo 34 *bis* della Convenzione di Vienna che dovrebbe stabilire che il veicolo dotato di sistema a guida autonoma, conforme alle regolamentazioni tecniche nazionali dei Paesi firmatari e alla legislazione nazionale sul funzionamento, soddisferebbe il requisito di cui all'articolo 8 del testo internazionale.

Qualora tale modifica dovesse realmente avvenire, sarà sicuramente necessario intervenire sul Codice della strada italiano al fine di adeguarlo alla nuova normativa.[604]

In Italia, infatti, solo in tempi recenti sono stati avviati *test* su strada, grazie al decreto ministeriale (c.d. *Smart Road*) del 28 febbraio 2018 del Ministero delle Infrastrutture e dei Trasporti, pubblicato in Gazzetta ufficiale il 18 aprile 2018 (n.90) attinente alle "*Modalità attuative e strumenti operativi della sperimentazione su strada delle soluzioni di Smart Road e di guida connessa e automatica*".[605] Tale decreto prevede interventi volti alla diffusione della connessione per la gestione intelligente del traffico ed inoltre, permette l'autorizzazione per la sperimentazione su strada dei veicoli a guida autonoma.[606] A seguito di questa normativa, le aziende, le università e gli enti di ricerca possono richiedere un'autorizzazione ministeriale per la sperimentazione dei veicoli autonomi. Al di fuori di tale autorizzazione, la circolazione di *self-driving cars* viene vietata dal Codice della strada che impone una guida condotta da un essere umano che se ne assuma la responsabilità.[607]

Si giungerà, quindi, ad una graduale assimilazione delle vetture autonome alle macchine tradizionali al fine di permetterne un loro uso consentito sulle strade comuni sottoponendole allo stesso regime assicurativo e di responsabilità (con possibili modifiche per adattarli il più possibile all'autonomia del veicolo).[608] Dunque, assisteremo ad una progressiva integrazione della disciplina giuridica: inizialmente grazie ad un affiancamento tra le vetture tradizionali e le semi-autonome con l'obbligatoria presenza vigile del conducente per poi spostarsi lentamente verso una scomparsa del guidatore. Tale processo seguirà l'evoluzione tecnologica e le sue stesse difficoltà di inserimento e progressione.[609]

---

[604] M. CAPRINO, "*Distrarsi alla guida diventa legale a luglio: via libera agli Adas di livello 3*", del 10 febbraio 2022, in il Sole 24 Ore: https://www.ilsole24ore.com/art/guida-autonoma-convenzione-vienna-apre-sistema-adas-AEm3x1BB

[605] D.M. Infrastrutture e trasporti 28/02/18, G.U. 18/04/18 n. 90 sulle "*Modalità attuative e strumenti operativi della sperimentazione su strada delle soluzioni di Smart Road e di guida connessa e automatica*" (18A02619): http://www.prefettura.it/FILES/AllegatiPag/1173/Decreto_Smart_Road_registrato_Prot._n._70_del_28_febbraio_2018.pdf

[606] L. PRINCIPALI E D. SALERNO, op. cit.

[607] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 1 (pag. 326 – 327) "Introduzione".

[608] E. PALMERINI, op. cit., Paragrafo 4.1 "Dalla qualificazione generica alla lacuna normativa".

[609] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 3 (pag. 328 – 329) "Le prospettive evolutive della disciplina giuridica della guida autonoma, tra bilanciamento e rischio consentito".

*1.2 Gli errori commessi dalle auto a guida autonoma e i soggetti responsabili*

Abbiamo parlato dei vantaggi che le macchine a guida autonoma possono garantire alla società, ma come ogni sperimentazione, purtroppo sussistono dei rischi e in passato si sono verificati incidenti. Infatti, il 14 febbraio 2016 una *Google Car* ha colpito uno *autobus* in California, il 18 marzo 2018 un'auto autonoma di *Uber* ha investito un pedone che attraversava la strada in Arizona e il 19 settembre 2019 un veicolo semi-autonomo di *Tesla* non ha rilevato una linea continua andando nella corsia di marcia opposta per sorpassare.[610]

Nel primo caso citato, una delle auto a guida autonoma di *Google* si è scontrata contro un autobus a Mountain View viaggiando a 3 *km/h* senza (fortunatamente) causare feriti. L'incidente è avvenuto su El Camino Real (un viale a sei corsie molto trafficato) mentre l'auto stava cercando si superare alcuni sacchi di sabbia posti sulla strada colpendo, con la sua parte anteriore sinistra, il lato destro dell'*autobus*. L'auto a guida autonoma, una volta individuati i sacchi di sabbia, ha frenato, ha fatto passare gli altri veicoli per poi avvicinarsi lentamente verso il centro della corsia al fine di superare l'ostacolo, entrando sfortunatamente in collisione con l'autobus. L'autovettura aveva rilevato l'*autobus* in avvicinamento ma aveva previsto che lo stesso avrebbe rallentato e che l'avrebbe fatta passare; proprio per questo, l'utente del veicolo *Google*, posto nel sedile anteriore in modo tale da intervenire in ogni momento (come imposto dalla legge statale), non aveva il controllo della vettura quando è avvenuta la collisione. Quindi, il *software* non ha evitato l'incidente e il conducente umano non è intervenuto. Con tutte le difficoltà di cui si è detto circa l'individuazione del responsabile: il *software*, l'utente non intervenuto o l'autista dell'*autobus*. Proprio per questo è necessaria una normativa anche in tema di responsabilità civile e di assicurazioni. Comunque, sono passati anni da questo incidente e *Google* ha perfezionato il suo *software* basandosi sul fatto che non sempre i veicoli (soprattutto di grandi dimensioni) rallentano in situazioni come queste.[611]

Nel secondo caso, nelle strade di Tempe a Phoenix (Arizona), un *SUV Volvo* del 2017 con la tecnologia *Autopilot* di *Uber* a velocità di 70 *km/h* ha causato la morte di un pedone (Elaine Herzberg, una donna di 49 anni) che attraversava fuori dalle strisce pedonali identificandolo come un falso positivo e dunque non rallentando prima dell'impatto. Il veicolo era in modalità autonoma con alla guida Rafaela Vasquez che al momento dell'incidente non stava toccando il volante. [612] Il pedone era sbucato all'improvviso ma il veicolo aveva rilevato comunque l'ostacolo considerandolo, però, come non pericoloso per posizione e movimento scartando l'ipotesi che si trattasse di una persona. [613]

---

[610] S. QUINTARELLI, op. cit., Capitolo. 6 (pag. 106 – 117) "Le leggi in gioco con l'intelligenza artificiale".
[611] N. BOWLES, "*Google self-driving car collides with bus in California, accident report says*", The Guardian, 1/03/2016: https://www.theguardian.com/technology/2016/feb/29/google-self-driving-car-accident-california
[612] M. ROVELLI, "*Uber, la polizia scagiona l'auto a guida autonoma: «Incidente inevitabile*", Corriere della sera – Tecnologia/New Economy, 20 marzo 2018: https://www.corriere.it/tecnologia/economia-digitale/cards/uber-polizia-scagiona-l-auto-guida-autonoma-incidente-inevitabile/caso-uber-impatto-inevitabile_principale.shtml
[613] S. QUINTARELLI, op. cit., Capitolo 2 (pag. 33 – 54) "Intelligenza artificiale. Applicazioni e tecnica".

Non si tratta del primo incidente mortale avvenuto con tali tecnologie, infatti, il 7 maggio 2016 una *Tesla* Model S in cui era stato inserito il sistema di *Autopilot*, si è schiantata contro un *camion* nelle autostrade della Florida causando la morte del passeggero Joshua Brown di 40 anni, membro della *Navy Seals* della Marina Militare Americana ed esperto in guerra elettronica.[614] L'incidente è dovuto a circostanze particolari: il *camion* ha svoltato a sinistra ponendosi difronte alla *Tesla* la quale ha confuso il rimorchio bianco del *camion* con il cielo a causa della giornata particolarmente luminosa, quindi non ha rallentato. L'*Autopilot*, è un *optional* che, se inserito, mantiene costante la velocità del veicolo e controlla eventuali ostacoli rallentando o accelerando di conseguenza, potendo, inoltre, sorpassare. Tutto sotto il controllo dell'utente umano. I sistemi di *Autopilot* non possono sostituire a pieno l'essere umano ma solo sostenerlo.[615]

Di chi è, dunque, la responsabilità? Del programmatore, di *Uber* o Tesla, dell'utente dell'auto, del pedone o dello Stato dell'Arizona o della Florida che hanno permesso di testare l'auto? [616] Nel caso di *Tesla* e di *Uber*, le auto non sono state considerate responsabili dato che nessuna di esse è ancora interamente autonoma: necessitando di un soggetto a bordo che possa intervenire in caso di emergenza.[617] Infatti, essendo ancora nella fase *beta* di sperimentazione, il conducente non può mai togliere le mani dal volante.[618]

Ovviamente la guida autonoma riduce fortemente il rischio di incidenti per distrazione del conducente ma non azzera le probabilità di collisioni. [619]

Nel terzo caso è ancora *Tesla* ad essere protagonista. Nella contea di Osceola in Florida, una donna, la ventinovenne Samantha Jensen, a bordo della sua *Tesla Model 3* del 2019, è rimasta uccisa in un incidente sulla Osceola Polk Line Road vicino a Shady Oak Drive. La vettura, su cui era impostato *Autopilot,* ha virato nella corsia opposta nel tentativo di superare il traffico entrando, però, in collisione con un *camion* che andava in direzione contraria.[620]

Questi fatti dimostrano che le autovetture in modalità *Autopilot* tendono ad entrare in collisione con veicoli o altri oggetti posti in posizioni anomale a causa del fatto che i sistemi intelligenti non riescono a gestire l'imprevisto come un essere umano non applicando correttamente le conoscenze acquisite in situazioni simili.[621]

---

[614] M. ROVELLI, op. cit.
[615] D. SPARISCI, "*Morto per un errore del pilota automatico. Cos'è l'Autopilot di Tesla e cosa è successo*" Corriere della sera – Tecnologia/New Economy, 1 luglio 2016: https://www.corriere.it/tecnologia/provati-per-voi/cards/morto-un-errore-pilota-automatico-cos-l-autopilot-tesla-cosa-successo/morte-marine_principale.shtml
[616] 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*".
[617] M. ROVELLI, op. cit.
[618] D. SPARISCI, "*Morto per un errore del pilota automatico. Cos'è l'Autopilot di Tesla e cosa è successo*" Corriere della sera – Tecnologia/New Economy, 1 luglio 2016: https://www.corriere.it/tecnologia/provati-per-voi/cards/morto-un-errore-pilota-automatico-cos-l-autopilot-tesla-cosa-successo/morte-marine_principale.shtml
[619] M. ROVELLI, op. cit.
[620] B. VOLZ, "*Family calls for investigation after Tesla driver dies in Osceola crash*", ClickOrlando.com, 26 settembre 2019: https://www.clickorlando.com/news/2019/09/26/family-calls-for-investigation-after-tesla-driver-dies-in-osceola-crash/
[621] A. LONGO E G. SCORZA, op. cit., Paragrafo 3 (pag. 88 - 147) "Vantaggi e rischi non economici dell'AI - I rischi etici, sociali e politici dell'AI - Algoritmi che discriminano".

I sistemi di guida autonoma non si limitano al settore della mobilità di terra. Noti i due disastri aerei che hanno causato la morte di tutte le 346 persone a bordo di due *Boeing 737 Max* (il volo *Lion Air 610* il 29 ottobre 2018 e il volo *Ethiopian Airlines 302* il 10 marzo 2019). I disastri sono stati attribuiti al sistema di intelligenza artificiale *Maneuvering Characteristics Augmentation System* (*MCAS*) che non ha reagito correttamente in circostanze particolari portando l'aereo automaticamente in una posizione di picchiata e che ha impedito ai piloti di intervenire come avrebbero dovuto. A causa dei difetti di progettazione del veicolo, nel 2019 è stato emanato da parte delle autorità aeronautiche del mondo, un provvedimento di interdizione al volo del *Boeing 737 Max* e nel 2020 l'azienda ha sospeso la sua produzione. Il suo possibile ritorno in servizio dipende dalle modifiche che verranno effettuate al sistema *MCAS.* [622]

Questi sono solo alcuni esempi di errori commessi dai sistemi di intelligenza artificiale che però impattano gravemente sulle vite degli esseri umani. Ciò comporta una certa diffidenza della società in relazione a tecnologie di questo tipo. Infatti, negli anni sono stati realizzati diversi studi sociologico-statistici per comprendere come la società vorrebbe che i sistemi di intelligenza artificiale si comportino in situazioni critiche. In tal modo si potrebbe ricostruire un'etica del sistema che, al netto degli errori di malfunzionamento, riuscirebbe a gestire situazioni moralmente complesse in modalità condivise dagli individui.

Non sono estranei al tema anche questioni di ordine morale che riguardano i casi in cui l'agire dell'intelligenza artificiale impone la scelta tra due beni giuridici di eguale valore. Si pensi allo scenario in cui la *self-driving car* sia costretta a scegliere tra la vita del passeggero e quella del pedone che attraversa la strada. Sul punto è stato fatto uno studio dalla rivista *Science*[623] nel 2016 in cui sono stati intervistati 2.3 milioni di persone di cui una grande fetta ha dichiarato che avrebbe preferito un'auto autonoma che protegga i pedoni anche a danno del passeggero anche se non l'avrebbe mai acquistata. I risultati hanno mostrato l'esistenza di alcuni principi morali condivisi da tutto il campione (salvare gli umani a danno degli animali e preferire i gruppi e i più giovani). Al contempo, sono notevoli le variazioni nelle preferenze morali in base ai Paesi di provenienza degli intervistati, ciò è legato alle culture più o meno individualiste, più o meno propense verso la tutela dei più vulnerabili, dei minori, delle donne, degli anziani e degli animali. Inoltre, i Paesi più poveri sono più tolleranti verso le infrazioni del Codice della strada (anche da parte del pedone) rispetto ai paesi fortemente industrializzati.

Quindi non esiste una morale universale in relazione alle decisioni che dovrebbero prendere le auto a guida autonoma. Nonostante ciò, alle macchine verrà affidata la responsabilità di decidere chi debba vivere o morire.[624]

---

[622] W. LANGEWIESCHE, "*What Really Brought Down the Boeing 737 Max?*", 18 settembre 2019, aggiornato il 2 luglio 2021, in The New York Times Magazine: https://www.nytimes.com/2019/09/18/magazine/boeing-737-max-crashes.html

[623] J.F. BONNEFON, A. SHARIFF E I. RAHWAN, "*The social dilemma of autonomous vehicles*" in Science, Vol. 352, Issue 6293 (pag. 1573 – 1576), 24 giugno 2016: https://www.science.org/doi/10.1126/science.aaf2654

[624] F. SUMAN, "*Dilemmi morali per le auto a guida autonoma*", in Il Bo Live, Università di Padova, 7 novembre 2018: https://ilbolive.unipd.it/it/news/dilemmi-morali-auto-guida-autonoma

Oltre a tali problematiche legate in parte ai difetti e *bias* di sistema e alla negligenza umana si aggiungono gli attacchi criminali alla sicurezza dei *software:*[625] c.d. *adversarial attacks.* Tali attacchi si basano sul porre in errore la macchina ingannando i sensori (ad esempio modificando i segnali stradali che sembrano dare la precedenza invece di imporre uno *stop*, oppure ponendo degli adesivi sulla carreggiata che spingano il veicolo ad invertire il senso di marcia in luoghi pericolosi). I rimedi si basano su una riprogrammazione minima del sistema in base agli *adversarial attacks* che vengono riconosciuti in modo tale da evitare che l'autovettura sbagli nuovamente: ma tale soluzione è esclusivamente a posteriori, inaccettabile nei casi in cui vengano messe a rischio vite umane. Infatti, è necessario allenare l'algoritmo in modo tale che possa reagire anticipatamente a tali attacchi grazie a simulazioni che vengono realizzate dai programmatori. Ed inoltre è fondamentale inserire sistemi di sicurezza che tengano conto di tali rischi anche imponendo che la decisione finale, in caso di situazioni complesse, spetti sempre all'essere umano. [626]

Dunque, è interessante analizzare chi debba rispondere per i danni arrecati dal veicolo a guida autonoma.

Quando il veicolo autonomo entra in collisione con un soggetto o un oggetto, in realtà, il *software*, che lo guida, sta funzionando perfettamente, andando ad applicare i modelli statistici e matematici più efficienti e servendosi di un bacino di conoscenza condiviso. Allora perché commette l'errore? La risposta risulta banale: i calcoli che effettua non sempre riescono a prevedere ogni specifica situazione, ciò dovuto in parte, anche, all'esperienza che il dispositivo possiede per alimentare il suo interno meccanismo di *machine learning*. Ma come sappiamo, tale sistema impara sbagliando: cosa succede se, però, ci va di mezzo un essere umano? Non è possibile estendere le norme a tutela dei consumatori dato che, come detto poco fa, in questo caso non sussiste alcun malfunzionamento del sistema. Ciò che è accaduto è una probabilità minima di comportamento della macchina difforme da quello auspicabile: ad oggi non possediamo una normativa specifica.[627] Il bilanciamento di interessi deve quindi riguardare da un lato, la riduzione delle vittime della strada e dei rischi di distrazione degli esseri umani che spinge verso una totale introduzione delle auto con guida autonoma e dall'altro il timore, dovuto alla completa estromissione del controllo umano dall'autonomia della macchina, di imprevisti che il sistema intelligente non riesce a prevedere ed evitare.

### *1.3 L'omicidio stradale e le lesioni in relazione all'uso dei veicoli autonomi: profili penalistici*

Come accennato poco fa, a seguito dell'immissione nel mercato di veicoli sempre più autonomi, si prospetteranno problematiche in relazione alla responsabilità penale in caso di incidente stradale.

[625] S. QUINTARELLI, op. cit., Capitolo 2 (pag. 33 – 54) "Intelligenza artificiale. Applicazioni e tecnica".
[626] A. LONGO E G. SCORZA, op. cit., Paragrafo 3 (pag. 88 - 147) "Vantaggi e rischi non economici dell'AI - I rischi etici, sociali e politici dell'AI - Algoritmi che discriminano".
[627] S. QUINTARELLI, op. cit., Capitolo. 6 (pag. 106 – 117) "Le leggi in gioco con l'intelligenza artificiale".

Tradizionalmente il procedimento penale in caso di sinistri stradali riguarda esclusivamente gli automobilisti coinvolti e i possibili pedoni danneggiati, in futuro potrà coinvolgere anche il produttore del veicolo nel caso in cui sussistano alcune condizioni: il mancato rispetto delle linee guida internazionali nella fase di sviluppo del *software*, la decisione irragionevole del veicolo in base al suo livello di informazione ed esperienza al momento del fatto, immissione sul mercato di tale tecnologia in assenza di un miglioramento degli *standard* di sicurezza e la violazione del principio di non discriminazione da parte del sistema.[628]

Per quanto riguarda il tradizionale omicidio stradale e le lesioni personali gravi o gravissime, introdotto nel nostro codice penale dalla legge 41/2016[629], dobbiamo far riferimento all'articolo 589 *bis* c.p. (omicidio stradale) il quale si riferisce a chiunque cagioni per colpa la morte di una persona con la violazione del Codice della strada e all'articolo 590 *bis* c.p. (lesioni personali stradali gravi o gravissime) che punisce chiunque cagioni per colpa ad altri una lesione personale in violazione del Codice della strada. Tale normativa, nonostante la *ratio* (tutelare gli eccessivi morti della strada) fosse condivisibile, è stata fortemente criticata in quanto troppo legata ad esigenze mediatiche e dunque fortemente imprecisa.[630] Gli articoli in questione vanno ad aumentare le rispettive pene dell'omicidio colposo e delle lesioni personali gravi o gravissime nel caso in cui sussistano violazioni del Codice della strada. Gli autori critici dubitano della necessità dell'introduzione di tali nuovi articoli in relazione al fatto che già sussistevano le aggravanti stradali connesse agli articoli 589 e 590 c.p. (abrogate a seguito della legge 41/2016).[631] Secondo la Cassazione, infatti, tali fattispecie costituiscono reati autonomi e non ipotesi aggravate dei reati di omicidio colposo e lesioni colpose.[632]

Si tratta di un reato di evento, proprio (circoscritto al solo guidatore di veicolo a motore) e procedibile d'ufficio, con aggravanti ad effetto speciale non sottostanti al giudizio di bilanciamento (art. 590 *quater* c.p.).[633] Infatti, l'articolo 589 *bis* c.p. prevede la reclusione da otto a dodici anni, se il soggetto si pone alla guida di un veicolo a motore in stato di ebrezza il cui valore sia superiore a 1.5 oppure in uno stato di alterazione dovuto all'assunzione di sostanze stupefacenti ed infine se si tratta di un soggetto che guida professionalmente ed ha un tasso alcolemico tra lo 0.8 e l'1.5; l'articolo in questione prevede, inoltre, la reclusione da cinque a dieci anni se il soggetto si pone alla guida con un tasso alcolemico con valore compreso tra 0.8 a 1.5 oppure realizzi violazioni gravi del Codice della strada (superare i limiti di velocità, attraversare un incrocio con il rosso, guidare contromano, effettuare un'inversione di marcia in prossimità di un incrocio

---

[628] L. BUTTI, op. cit.

[629] Legge 23 marzo 2016, n. 41, "*Introduzione del reato di omicidio stradale e del reato di lesioni personali stradali, nonché disposizioni di coordinamento al decreto legislativo 30 aprile 1992, n. 285, e al decreto legislativo 28 agosto 2000, n. 274 (16G00048)*", entrata in vigore il 25/03/2016 (GU. N.70 del 24/03/2016): https://www.gazzettaufficiale.it/eli/id/2016/03/24/16G00048/sg

[630] A. MASSARO, "*Omicidio stradale e lesioni personali stradali gravi o gravissime: da un diritto penale frammentario a un diritto penale frammentato*", in Diritto penale Contemporaneo, scritto in relazione al Convegno "Omicidio e lesioni stradali – tra esigenza di giustizia, di deterrenza e di obiettività ricostruttiva della dinamica dell'incidente stradale" svoltosi il 1 aprile 2016 presso la Sala Unità d'Italia della Corte di Appello di Roma, Paragrafo 1 (pag. 2 – 4) "L'irresistibile tentazione di critica demolitoria «a prima lettura»".

[631] A. MASSARO, op. cit., Paragrafo 2 (pag. 4 – 6) "Da un diritto penale «frammentario» a un diritto penale «frammentato»: la tendenza alla parcellizzazione (anche) del nucleo duro del diritto penale".

[632] Sentenza della Cassazione penale, Sez. IV del 14 giugno 2017, n. 29721.

[633] A. MASSARO, op. cit., Paragrafo 4 (pag. 7 – 8) "I «nuovi» delitti di omicidio e di lesioni stradali: le fattispecie base".

o di una curva).[634] Viene inoltre previsto il concorso di reati qualora a causa dell'azione, si generi la morte di più persone oppure la morte e le lesioni di più persone: in tal caso, si applicherà la pena del reato più grave aumentata fino al triplo con il limite dei diciotto anni.[635]

Tra le circostanze, l'articolo 589 *ter* c.p. prevede l'aggravante della fuga. In caso di fuga e di omissione di soccorso a seguito della commissione dell'illecito, si prevede un aumento della pena da un terzo a due terzi comunque maggiore di cinque anni nel caso di omicidio stradale e di tre anni nel caso di lesioni stradali.[636]

Per quanto concerne le attenuanti, invece, al settimo comma di entrambi gli articoli 589 *bis* e 590 *bis* c.p. è prevista una diminuzione della pena fino alla metà nel caso in cui l'evento non sia esclusiva conseguenza dell'azione od omissione del colpevole. Si rientra dunque nello schema normativo (più ampio degli articoli 40 e 41 c.p.) della *conditio sine qua non* in base al quale la condotta umana è necessaria ma non sufficiente nella spiegazione causale dell'evento dovuto ad una serie causale di condotte apparentemente indipendenti. È necessario che la condotta rientri in una delle condizioni che hanno fatto verificare l'evento ma che non sia l'unica ad averlo realizzato. Dunque, per applicare tali circostanze attenuanti speciali, è necessario che sussista un "*contributo concorrente fornito dalla vittima nella determinazione dell'evento, ma anche in ogni altra ipotesi che sia dipesa dalla condotta di altri conducenti e da altri fattori esterni da individuarsi di volta in volta*".[637] Quindi, tra le concause vi rientrano la condotta (non necessariamente risultante in un fatto illecito) di un altro conducente oppure la condotta (dolosa o colposa) della vittima. In tale secondo caso bisogna coordinare gli articoli in questione con l'articolo 62 comma 5 c.p. (fatto doloso della vittima): infatti, mentre per gli articoli 589 *bis* comma 7 e 590 *bis* comma 7 c.p. la condotta della persona offesa, che ha concorso a generare l'evento lesivo insieme alla condotta del *reo*, rileva sia se colposa sia se dolosa (in quanto non viene specificato alcunché dalle norme), per l'articolo 62 comma 5 c.p. rileva esclusivamente il fatto doloso commesso dalla persona offesa. Tra le norme sussiste, dunque, un rapporto di specialità reciproca: l'articolo 62 comma 5 c.p. fa riferimento al fatto doloso della vittima senza definire l'evento penalmente rilevante; gli articoli 589 *bis* e 590 *bis* c.p. al settimo comma non fanno riferimento alla tipologia di concausa ma evidenziano l'evento illecito come la morte o lesioni derivanti da violazioni del codice della strada.[638]

In tale complessa articolazione normativa si va inoltre ad aggiungere il principio dell'affidamento sulla buona condotta altrui: nel caso di reati stradali, questo viene attenuato in quanto si risponde anche qualora il fatto altrui sia (soltanto) astrattamente prevedibile, dato che, la diligenza alla guida deve essere massima. Ma così facendo si rischia di svuotare di significato tale principio. In questo modo, il guidatore viene comunque ritenuto responsabile del fatto materialmente cagionato anche quando non era concretamente prevedibile l'altrui comportamento: si avrebbe un generico dovere di prevedere tutto, dovuto ad una scelta di politica criminale volta alla tutela delle vittime della strada. Infatti, spesso, il consapevole ed imprevedibile

[634] A. MASSARO, op. cit., Paragrafo 4.1 (pag. 8 – 10) "La nutrita schiera delle circostanze aggravanti".
[635] A. MASSARO, op. cit., Paragrafo 4.1.2 (pag. 10 – 12) "Possibili spazi per la configurabilità del reato complesso?".
[636] A. MASSARO, op. cit., Paragrafo 6 (pag. 21 – 22) "L'aggravante della fuga".
[637] Sentenza della Cassazione penale, Sez. IV del 26 marzo 2019, n. 13103.
[638] A. MASSARO, op. cit., Paragrafo 4.2 (pag. 12 – 16) "L'attenuante della «concausa»".

comportamento della persona offesa che si pone in pericolo non risulta sufficiente per liberare dalla responsabilità il guidatore (al massimo, come affermato precedentemente, può rilevare come attenuante).[639]

Per quanto concerne l'elemento soggettivo sussistono dei dubbi riguardanti la necessità del dolo o della colpa. Il legislatore fa rientrare tali reati tra i delitti colposi mentre la giurisprudenza apre la strada ad una responsabilità per dolo al limite tra il dolo eventuale e la colpa cosciente. Soprattutto per quanto riguarda la guida in stato di ebrezza o il superamento dei limiti di velocità. Infatti, taluno considera rientrante nel dolo la condotta di chi si ponga consapevolmente ubriaco alla guida in quanto si avrebbe l'accettazione del rischio di provocare la morte o le lesioni a qualcuno; anche se non è detto che il soggetto stia effettivamente accettando tale rischio per il semplice fatto di porsi alla guida in tale stato, anzi, spesso il conducente è sicuro di poter guidare bene (rientrando dunque in una colpa cosciente). Altri considerano l'elemento soggettivo del reato, senza ombra di dubbio, come colpa anche perché, nel caso in cui ci fosse il dolo, si rientrerebbe nell'omicidio comune. Nonostante la norma faccia chiaramente riferimento alla colpa specifica, taluni ritengono sufficiente una colpa generica consistente in imperizia, imprudenza o negligenza (facendo riferimento al fatto che l'articolo 140 del Codice della strada prevede un obbligo generico di guida prudente e attenta alla sicurezza degli altri utenti). Una terza via considera come reato solo gli eventi lesivi o mortali dipendenti dalla violazione del Codice della strada per condotte che siano indistintamente dolose o colpose: ciò comporterebbe, però, problematiche in relazione all'effettività della pena.[640]

Premessi questi brevi cenni alla disciplina dell'omicidio stradale, ci si deve porre la domanda sul chi risponderà qualora il fatto lesivo o mortale venga cagionato da un veicolo a guida autonoma. Ad oggi, come esplicato precedentemente, la circolazione di tali prototipi (esclusivamente semi-autonomi e con la presenza obbligatoria di un guidatore umano che possa intervenire in ogni momento) è permessa solo per le sperimentazioni autorizzate dal Ministero e dunque la valutazione che effettueremo è proiettata verso un futuro, forse prossimo, in cui sarà consentita liberamente la circolazione di questi veicoli.

Bisogna preliminarmente distinguere tra veicoli semi-autonomi (di terzo livello) e completamente autonomi (di quarto e quinto livello).

Nel primo caso infatti, è sempre individuabile un conducente persona fisica (pronta ad intervenire in caso di necessità) che può essere considerato responsabile per i danni realizzati. Dunque, per tale categoria di soggetti risulta semplice sottoporli alla disciplina dei reati stradali comuni (art. 589 *bis* e 590 *bis* c.p.). [641] Infatti, fino a tale livello di automazione, nonostante sussista un margine di autonomia, viene necessariamente richiesto un controllo umano gravante sul conducente al fine di evitare danni a terzi. È quindi possibile considerarlo come garante (più marcata nei primi due livelli, rispetto al terzo dove l'utente deve intervenire

[639] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 4 (pag. 334 – 337) "Reato colposo d'evento e vetture semi-autonome: il *control dilemma*".
[640] A. MASSARO, op. cit., Paragrafo 5 (pag. 16 – 21) "Come si stava senza l'omicidio stradale? Dalla colpa al dolo: andata e ritorno".
[641] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 4 (pag. 334 – 337) "Reato colposo d'evento e vetture semi-autonome: il *control dilemma*".

solo in caso di emergenza).[642] Il conducente risulterà quindi gravato da un obbligo di controllo sul veicolo semi-autonomo seguendo il dettato dagli articoli 140 e 141 del Codice della strada che impongono al guidatore di comportarsi in modo da non recare pericolo per gli altri grazie ad un controllo sul proprio veicolo anche realizzando manovre straordinarie e necessarie per la sicurezza. In tal modo il diritto penale riesce a tutelare anche queste situazioni innovative. A margine, però, uno spunto di riflessione andrebbe posto sull'effettiva possibilità di intervento da parte del conducente sul proprio veicolo semi-autonomo: sussistono, infatti, situazioni in cui la vettura realizza quasi autonomamente operazioni che l'essere umano difficilmente riesce a controllare (come azioni autonome del veicolo sul volante).Viene quindi da chiedersi se considerare l'utente come effettivo conducente non sia in realtà una *fictio iuris*.

Per quanto concerne i veicoli autonomi di quarto livello, in cui la figura conducente (potenziale e solo in caso di necessità) continua ad essere presente anche se in via fortemente eccezionale, potrebbe ipotizzarsi uno spazio di non punibilità per il guidatore.

In tal caso, l'evento realizzato dal veicolo semi-autonomo imputabile al conducente passerebbe dall'essere una condotta attiva ad una omissiva dato che il soggetto si sarebbe astenuto dal compiere un atto idoneo ad impedire la realizzazione della lesione. A tale grado di automazione, a differenza dei veicoli tradizionali in cui il conducente guida effettivamente (anche nel caso di vetture di terzo livello in cui essa non fa altro che assistere il guidatore durante il tragitto), il soggetto si limita a sorvegliare l'attività dell'autovettura semi-autonoma. Ciò diversamente da quanto avviene nella tradizionale fattispecie di colpa stradale consistente in un evento imprudente dovuto ad una condotta attiva (dato che, anche qualora consista in una mancanza, in realtà si inserisce nella condotta attiva di guidare): infatti, ogni condotta colposa contiene in sé un'omissione consistente nel non aver posto in essere la condotta cautelare doverosa e nel caso di guida tradizionale questa consiste nel non aver guidato nel rispetto del Codice della strada (che di per sé, nel suo complesso, costituisce una condotta attiva). Passando ad un grado più alto di automazione (il quarto livello) l'attività di guidare si trasforma in una sorveglianza priva di interventi attivi (salvo in alcuni casi): qualora il conducente pecchi di sorvegliare e il veicolo provochi un incidente, si ricadrà in una condotta omissiva di non idonea sorveglianza. Dunque, si passa da una piena condotta attiva nei veicoli tradizionali ad una condotta alternativamente commissiva (qualora il conducente intervenga erroneamente nella guida) o omissiva nei veicoli semi-autonomi.

Nonostante ciò, bisogna anche sottolineare che il conducente del veicolo (qualunque esso sia) ha un generale obbligo di prudenza e dunque, anche nel caso dei veicoli semi-autonomi, si potrebbe far rientrare l'evento lesivo all'interno del paradigma degli articoli 589 *bis* e 590 *bis* c.p. in quanto, tramite quel dovere generale si fa ricadere la responsabilità per colpa per qualunque evento si verifichi. Nel mondo delle auto semi-autonome, l'obbligo di prudenza del conducente non deve solo prevedere le imprevedibili azioni degli altri utenti della strada ma anche quelle del veicolo stesso.

[642] C. PIERGALLINI, op. cit., Paragrafo 3 "Danno da prodotto e intelligenza artificiale sotto il controllo umano".

Nonostante tali difficoltà di prevenzione totale del rischio, la *ratio* dei reati stradali, come accennato precedentemente, sta nella necessaria tutela delle vittime degli incidenti e dunque, tale protezione voluta da scelte politiche difficilmente si concilia con una limitazione della punibilità per i conducenti delle auto semi-autonome (anche a causa della diffidenza verso la sicurezza delle auto *self-driving*). Dunque, qualora dovessero essere poste in commercio (e non solo ai fini della sperimentazione) le auto semi-autonome, sicuramente, un approccio basato sulla responsabilità del conducente per le azioni commesse con il suo veicolo (anche se in concreto incapace di controllare gli eventi realizzatisi) può risultare più rassicurante per tutti coloro che temono un vuoto normativo di tutela delle vittime ma al contempo sanzionerebbe penalmente come "capro espiatorio" il guidatore, non sempre effettivamente capace di impedire l'evento dannoso generatosi.

Nasce così il c.d. *control dilemma* dovuto al fatto che nonostante l'elevata automazione del veicolo, il conducente non verrebbe liberato dalla responsabilità degli eventi lesivi anche nel caso in cui non abbia le capacità per evitarli (anche a causa dell'essere, in futuro, disabituati a guidare grazie a tali veicoli autonomi che si sostituiscono in concreto ad un guidatore vero e proprio). Si rischia così di entrare in contrasto con il principio di colpevolezza. [643]

In fine, in relazione alle *fully self-driving cars,* ossia i veicoli di quinto livello di automazione, l'inquadramento di forme di responsabilità per l'addebito penale risulta ancora più complesso dato che, l'assenza di comandi manuali fa venir meno la figura del conducente che si trasforma in mero passeggero. Egli potrà dunque astrattamente rispondere a titolo di colpa solo per la mancata manutenzione del veicolo, quando potrà essere provato che l'incidente lesivo si sia verificato per tale ragione (e solo nel caso in cui gravi su di lui un obbligo giuridico di garantire la piena funzionalità del veicolo). Dunque tutte le lesioni generate non per ragioni di non manutenzione del veicolo ma per errori o malfunzionamenti del *software* di guida autonoma non potranno essere imputati al conducente.

Si rischia così di generare un vuoto di tutele che deve necessariamente essere colmato.

Abbiamo già escluso, nel secondo capitolo, la possibilità di far rispondere (ad oggi) direttamente il sistema dotato di intelligenza artificiale e quindi bisognerà individuare un soggetto persona fisica che possa rispondere per colpa degli eventi lesivi realizzatisi: i programmatori e i costruttori del veicolo autonomo.

Si potrebbe così individuare una responsabilità per danno da prodotto difettoso (se il danno sia in concreto connesso ad un malfunzionamento del *software* di intelligenza artificiale) nonostante, in relazione all'intelligenza artificiale, sussistano notevoli problematicità connesse alla complessità della fase di produzione della stessa a causa dell'elevato numero degli scenari che i programmatori sono chiamati a valutare. Infatti, per un unico sistema, è possibile trovare il lavoro svolto in simbiosi da più soggetti coordinati tra loro in porzioni indistinguibili e inseparabili che genereranno poi il prodotto finale. [644]

---

[643] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 4 (pag. 334 – 337) "Reato colposo d'evento e vetture semi-autonome: il *control dilemma*".

[644] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 5 (pag. 338 – 341) "L'addebito colposo del sinistro stradale nelle auto completamente autonome, fra danno da prodotto, imprevedibilità tecnologica e "vuoto" di responsabilità".

Si crea, così, una distinzione tra responsabilità per tipo e per modo di produzione. La prima consiste nella qualità del prodotto come risultato dell'attività produttiva andando a valutare la pericolosità dello stesso, non sempre evitabile in base a misure cautelari. La seconda, invece, riguarda il processo produttivo realizzato al fine di evitare la concretizzazione di rischi di danno del prodotto. Altra distinzione importante riguarda i difetti di costruzione (inerenti all'intera serie prodotta e connessi a difetti di ideazione, di tecnica o di materiali), i difetti di fabbricazione (che colpiscono alcuni elementi della serie prodotta e che possono essere statisticamente calcolati dato che sono dovuti ad errori della macchina o dell'essere umano), i difetti di informazione (nel caso in cui il prodotto perfettamente funzionante circoli senza le istruzioni d'uso dirette ad evitare danni) e i difetti da rischio di sviluppo (sconosciuti al momento dell'immissione sul mercato ed dunque imprevedibili ed emergenti esclusivamente in un secondo momento).

Tali distinzioni non sono però utili ad accertare il nesso di causalità fondamentale per individuare una responsabilità penale del soggetto. Il difetto potrebbe, infatti, dipendere da fasi produttive composte da una miriade di soggetti (che possono provenire da differenti entità collettive collaboranti tra loro) che devono affrontare un'indefinita quantità di scenari. [645] Risulterà quindi complesso individuare il soggetto specifico che ha omesso di valutare una data situazione che poi si sia verificata in concreto. Nel caso in cui si dovesse superare tale difficoltà di individuazione, si potrebbe liberamente applicare tale disciplina.[646] Altrimenti si rischierebbe di ledere il principio di personalità della responsabilità penale.

Nel diritto penale, dunque, non è possibile riconoscere la responsabilità da prodotto in capo al singolo soggetto senza rischiare di rientrare in una responsabilità oggettiva, però è possibile addossare un margine di responsabilità al produttore del veicolo autonomo qualora si dovesse provare che il danno è direttamente dovuto al difetto di prodotto cagionato da una violazione degli *standard* di diligenza.

Ancora più complesso il caso in cui non sia possibile individuare con certezza l'errore da attribuire alla fase di programmazione. In tale situazione non potrà quindi essere addebitata una responsabilità per danno da prodotto. Bisogna quindi sottolineare l'autonomia di azione di tali *fully self-driving cars* che le porta ad essere non più semplici oggetti dell'azione ma anche (in parte) soggetti della stessa dato che, inoltre, non si basano esclusivamente su algoritmi programmati in modo fisso dall'essere umano, ma seguono meccanismi di *machine learning* al fine di auto-apprendere dall'esperienza ed aumentare il proprio bagaglio di conoscenze degli scenari della strada. Dunque, il veicolo può modificare i suoi comportamenti in modo imprevedibile per rispondere agli *input* che gli provengono dall'esterno rivelandosi come una *black box* in cui non è possibile comprendere a pieno le ragioni di scelta. Quindi risulta ancora più complesso risalire allo specifico errore di valutazione commesso dal singolo programmatore a causa di un'assenza di una possibile prevedibilità *ex ante* del comportamento del sistema.

---

[645] G. LATTANZI, P. SEVERINO, A. GULLO, op.cit., Volume I, Parte Seconda "La disciplina della responsabilità da reato delle persone giuridiche", M. PELISSERO, E. SCAROINA, V. NAPOLETANI, Capitolo 1 "Principi generali", Paragrafo 7 (pag. 154 – 161) "Il principio di autonomia della responsabilità dell'ente".

[646] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 5 (pag. 338 – 341) "L'addebito colposo del sinistro stradale nelle auto completamente autonome, fra danno da prodotto, imprevedibilità tecnologica e "vuoto" di responsabilità".

Si genera, dunque, in questi casi, un vuoto di tutela dato che il sinistro sembrerebbe direttamente riconducibile al veicolo autonomo: restando necessariamente al di fuori di una disciplina penale.

Arriviamo quindi ad un'alternativa: autorizzare l'evoluzione tecnologica fino alla completa automazione dei veicoli che comporterebbe vantaggi in tema di sicurezza stradale ma che lascerebbe scoperte da tutela alcune situazioni concrete, oppure, proibire tale nuova tecnologia limitando la globale fruizione dei benefici ad essa connessi. La scelta di politica legislativa che verrà fatta in futuro dipenderà principalmente dai timori della società all'affidare la propria vita ad un sistema autonomo di intelligenza artificiale.[647]

## ***2. L'intelligenza artificiale e la responsabilità dell'operatore sanitario***

Come accennato nel primo capitolo, lo sviluppo dell'intelligenza artificiale consente una notevole evoluzione nel campo delle tecniche diagnostiche e di intervento in campo medico e chirurgico.

Ad oggi è possibile affidarsi a sistemi intelligenti per il monitoraggio del paziente, per sostenerlo nella somministrazione dei farmaci e per il riconoscimento di patologie attraverso immagini radiologiche. Inoltre, fortemente all'avanguardia risulta essere la possibilità di effettuare interventi chirurgici a distanza tramite appositi strumenti tecnici che permettono, attraverso l'impiego di braccia robotiche, di agire sul paziente in sicurezza. Inoltre, grazie all'analisi dei *big data* che viene rapidamente svolta dagli algoritmi intelligenti, è possibile realizzare uno *screening* sulla popolazione in modo da comprendere quali siano le principali cause di morte nella società così da ridurre l'incremento della mortalità:[648] in tal modo, la ricerca medica potrà studiare la popolazione e innovarsi rendendo rapida la messa in commercio di nuovi farmaci[649] e potendo prevenire precocemente possibili patologie agendo di conseguenza sul singolo paziente.[650]

Di conseguenza, tramite tali sistemi di monitoraggio e prevenzione, sarà possibile generare un grande quantitativo di dati sulla salute del paziente che potranno essere scambiati a distanza e inseriti in una banca dati a cui potrà accedere lo *staff* clinico di riferimento in modo protetto (e nel rispetto della *privacy* dell'interessato) al fine di giungere ad una diagnosi precoce di possibili malattie. In tal modo, tramite dispositivi non invasivi e di facile utilizzo da parte del paziente, si potrà quindi tenere costantemente sotto controllo la situazione clinica del soggetto individuando immediatamente parametri di rischio, così da prevenire efficacemente alcune malattie, ridurre la degenza *post* operatoria o diagnostica in ospedale, avere maggiore sicurezza nelle diagnosi e nelle decisioni mediche ed infine, ridurre i costi di cura del paziente[651]

---

[647] A. CAPPELLINI, "*Profili penalistici delle self-driving cars*" op. cit. Paragrafo 5 (pag. 338 – 341) "L'addebito colposo del sinistro stradale nelle auto completamente autonome, fra danno da prodotto, imprevedibilità tecnologica e "vuoto" di responsabilità".
[648] A. LONGO E G. SCORZA, op. cit., Capitolo 6 (pag. 183 - 198) "L'AI in sanità".
[649] E. SANTORO, "*Sanità, ecco le tre innovazioni che la cambieranno quest'anno*", in *Agendadigitale.eu*, 20 febbraio 2019: https://www.agendadigitale.eu/sanita/sanita-ecco-le-tre-innovazioni-che-la-cambieranno-questanno/
[650] A. LONGO E G. SCORZA, op. cit., Capitolo 6 (pag. 183 - 198) "L'AI in sanità".
[651] *Idem.*

Questo fenomeno viene definito come "telemedicina". Essa permette, quindi, di realizzare una rete ospedaliera per monitorare i pazienti favorendo la prevenzione e permettendo ai soggetti di rimanere autosufficienti, autogestendosi dalle proprie abitazioni in sicurezza tramite dispositivi medici portatili (per il monitoraggio della glicemia, dell'ossigeno del sangue, della frequenza cardiaca) interconnessi tra loro capaci di comunicare al medico possibili anomalie e di adattare il loro comportamento in base alle informazioni ricevute, garantendo la somministrazione e il dosaggio adatto di farmaci.[652]

È, quindi, possibile affermare che i sistemi di intelligenza artificiale, se usati correttamente, aiutano gli esseri umani a leggere meglio la realtà in cui vivono. [653] Grazie al meccanismo di *machine learning,* che lavora su grandi quantità di dati dei pazienti, è possibile riuscire ad individuare i primi segnali di patologie invisibili al medico umano.[654] Verrebbe così garantita una medicina predittiva e preventiva prima della comparsa dei sintomi di malattie croniche grazie all'accesso istantaneo al *set* di dati del paziente.[655] Questo perché i sistemi di *machine learning* imparano dall'esperienza e se addestrati efficacemente sulla base di dati sanitari hanno l'abilità di superare le capacità di un medico persona fisica. Infatti, il programma di intelligenza artificiale è progettato per imitare il ragionamento clinico svolto da un medico attraverso il trattamento automatico del linguaggio naturale (che permette alla macchina di apprendere dall'analisi del discorso spontaneo): in questo modo, il sistema intelligente impara ad associare i sintomi alle patologie mediche più o meno rare tramite l'utilizzo di modelli statistici e probabilistici fortemente complessi che permettono di valutare i possibili fattori di rischio del paziente.[656] Quindi, l'utilizzo di *big data* e dell'intelligenza artificiale permette una medicina basata su ciò che risulta evidente (non più al singolo medico) ma a seguito di un'analisi automatizzata dei dati raccolti che porta a galla situazioni altrimenti non evidenti.

L'efficienza di tale meccanismo diagnostico è legata all'effettiva realizzazione di un fascicolo sanitario elettronico che consente l'aggiornamento continuo ed il monitoraggio a distanza dei parametri clinici del paziente.[657] Tale fascicolo risulta essere una delle maggiori fonti di *big data* sanitari utili per raggiungere l'apice della medicina personalizzata. Esso permette la realizzazione di un *database* di documenti analizzabili da strumenti di intelligenza artificiale.[658] Gli obiettivi di tale fascicolo sono agevolare l'assistenza del paziente, facilitare l'integrazione delle diverse competenze professionali, realizzare una base consistente di informazioni migliorando così i servizi di prevenzione, diagnosi e cura.[659]

---

[652] D. MARINO, A. MICELI, D. NACCARI CARLIZZI E G. QUATTRONE, "*Telemedicina, cos'è e come farla in Italia: tecnologie e finalità, un modello possibile*", in *Agendadigitale.eu*, 8 aprile 2020: https://www.agendadigitale.eu/sanita/telemedicina-come-farla-in-italia-le-tecnologie-le-finalita-un-modello-possibile/

[653] E. SANTORO, "*Sanità, ecco le tre innovazioni che la cambieranno quest'anno*", op. cit.

[654] G. F. ITALIANO, "*Intelligenza artificiale, che errore lasciarla agli informatici"*, in *Agendadigitale.eu*, 11 giugno 2019: https://www.agendadigitale.eu/cultura-digitale/intelligenza-artificiale-che-errore-lasciarla-agli-informatici/

[655] D. MARINO, A. MICELI, D. NACCARI CARLIZZI E G. QUATTRONE, op cit.

[656] E. SANTORO, "*L'algoritmo è un buon pediatra: così migliorano le diagnosi dell'AI*", in *Agendadigitale.eu*, 5 marzo 2019: https://www.agendadigitale.eu/cultura-digitale/lalgoritmo-e-un-buon-pediatra-cosi-migliorano-le-diagnosi-dellai/

[657] A. LONGO E G. SCORZA, op. cit., Capitolo 6 (pag. 183 - 198) "L'AI in sanità".

[658] SERGIO PILLON, "*Sanità digitale in Italia, le tendenze e le priorità del 2019*", in *Agendadigitale.eu*, 30 gennaio 2017: https://www.agendadigitale.eu/sanita/sanita-digitale-in-italia-le-tendenze-e-le-priorita-del-2019/

[659] A. F. PATTARO, "*Fascicolo Sanitario Elettronico, cos'è, a che serve e come attivarlo*" in *Agendadigitale.eu*, 16 settembre 2021: https://www.agendadigitale.eu/sanita/fascicolo-sanitario-elettronico-cose-e-a-che-punto-e-la-guida/

Tale sistema diagnostico – assistenziale sarà utile per garantire un'efficiente telemedicina (fondamentale per i soggetti più fragili). [660]

L'ultima frontiera in materia di telemedicina e di fascicolo sanitario è rappresentata dalla creazione di veri e propri gemelli digitali del paziente. Il gemello digitale è la raffigurazione elettronica di ogni informazione medica rilevante, costantemente aggiornata del paziente. Come tale rappresenta il più alto grado di medicina personalizzata, con tutte le conseguenze positive che ne derivano: possibilità di effettuare *test* clinici sul paziente digitale (senza ricorrere dunque alla sperimentazione animale) di produrre e somministrare farmaci personalizzati per uno specifico paziente, di analizzare le cellule digitali dell'individuo per poter prevenire future malattie del soggetto, e finanche, di realizzare una vera e propria clinica virtuale rappresentativa dell'intera popolazione. Le difficoltà sono molte dato che le variabili di incidenza sulla salute dei soggetti sono potenzialmente infinite. Ma, nonostante tali criticità, i modelli digitali degli organi umani permettono numerosi vantaggi in tema di diagnosi e terapie, facilitando l'operato del medico anche in campo di ricerca. Inoltre, le informazioni del gemello digitale sarebbero sempre a disposizione del medico curante del paziente che, potrà sempre monitorare le sue condizioni di salute. [661]

Inoltre, sotto un altro profilo, l'utilizzo dell'intelligenza artificiale in campo medico consente altresì di utilizzare operatori virtuali per l'assistenza del paziente. A tal fine sono stati elaborati una serie di strumenti: le *chatbot* come *Chat Yourself* nata con lo scopo di aiutare i pazienti con l'*Alzheimer* o l'assistente virtuale *Filos* che riesce a rispondere alle domande più comuni dei pazienti oncologici. Inoltre, molto innovativa è anche *Babylon App* che, tramite un sistema di intelligenza artificiale, permette una consulenza medica basata sulla conoscenza dell'anamnesi personale, dei sintomi e delle linee guida mediche. Essa risulta talmente tanto efficace che il *National Healt Service* inglese l'ha scelta come strumento di *triage* su un gruppo di cittadini a nord di Londra.

Tra gli ultimi approdi della tecnologia medica è possibile utilizzare le c.d. pillole intelligenti, dotate di sensori particolari in collegamento *Bluetooth* con un'*app* sullo *smartphone,* volte al monitoraggio del paziente e alla somministrazione dei famaci. Esse raccolgono dati dallo stomaco del soggetto per un mese rilasciando gradualmente il farmaco. In questo modo sarà più rapida ed efficace la rilevazione di patologie o di infezioni nel paziente.[662] Inoltre, in ambito psichiatrico sussistono pillole (*Abilify My Cite*) che includono un sensore in grado di segnalare l'effettiva somministrazione del farmaco da parte di soggetti con disturbi mentali.[663]

Ciò premesso, il ricorso alla medicina intelligente, per quanto all'avanguardia da un punto di vista scientifico, pone una serie di problematiche dal punto di vista etico e giuridico, potendo influire negativamente sul perno del rapporto medico-paziente: la fiducia. Come sappiamo, infatti, i sistemi di *machine learning*

---

[660] D. MARINO, A. MICELI, D. NACCARI CARLIZZI E G. QUATTRONE, op. cit.
[661] G. PORRO, "*Il gemello digitale del corpo umano: la nuova frontiera della medicina personalizzata*" in *Agendadigitale.eu*, 15 gennaio 2022: https://www.agendadigitale.eu/sanita/il-gemello-digitale-del-corpo-umano-la-nuova-frontiera-della-medicina-personalizzata/
[662] E. SANTORO, "*Sanità, ecco le tre innovazioni che la cambieranno quest'anno*", op.cit.
[663] E. SANTORO, "*Tecnologie in Sanità, ecco tre prodotti innovativi di provata efficacia*", in *Agendadigitale.eu*, 5 febbraio 2018: https://www.agendadigitale.eu/sanita/tecnologie-sanita-tre-prodotti-innovativi-provata-efficacia/

vengono paragonati ad una *black box*.[664] Proprio per questo, gran parte degli studiosi sostengono la necessità di far rimanere tali sistemi come un mero sostegno al medico, senza mai sostituirlo del tutto: sia per ragioni medico-sanitarie che per ragioni di responsabilità in caso di errore. È fondamentale che la decisione finale sia dell'essere umano e che sia il medico in persona a validare la decisione a cui è giunto l'algoritmo prima di iniziare quello specifico trattamento sanitario sul paziente.[665]

Tali problematiche sono dovute sia all'oscurità (intrinseca) dell'algoritmo, sia alla carenza di un'adeguata formazione del personale medico-sanitario relativa all'uso corretto dei sistemi intelligenti. È necessario, quindi, una formazione specifica sui macchinari dotati di intelligenza artificiale.[666] Per raggiungere tale obiettivo è fondamentale aumentare gli investimenti.

I futuri medici dovranno saper raccogliere i dati (tutelandone la riservatezza) attraverso *database* e algoritmi intelligenti, inoltre, dovranno saper utilizzare apparecchiature di monitoraggio da remoto oltre, ovviamente, ai sistemi di intelligenza artificiale specifici per il loro campo di conoscenza.[667] In più, per riuscire ad avere una sanità digitale effettiva con professionisti formati sul tema, è necessaria una strategia globale in ambito clinico – sanitario al fine di raggiungere la piena digitalizzazione dei processi e delle cartelle cliniche e l'effettiva possibilità di garantire un'assistenza sanitaria a domicilio.[668]

Sussistono però dei rischi connessi all'utilizzo dell'intelligenza artificiale in ambito sanitario.

In *primis* è possibile osservare il problema dei *bias* algoritmici[669] che sono dovuti principalmente a dati immessi in modo errato o contenenti discriminazioni fin dall'origine: infatti, i dati vengono raccolti dall'ambiente esterno che di per sé è influenzato da pregiudizi che, essendo spesso radicati in un determinato contesto sociale, sono complessi da eliminare. La macchina, dunque, riflette le discriminazioni umane: è importante agire sulla società per evitare che queste si ripercuotano sui sistemi intelligenti. A maggior ragione, il problema si amplia qualora esse vengano immesse in un sistema sanitario intelligente perché comporterebbe discriminazioni sociali tra soggetti che andranno ad impattare sulla loro salute violando l'articolo 3 e 32 della Costituzione.[670]

---

[664] G. F. ITALIANO, op. cit.
[665] E. SANTORO, "*L'algoritmo è un buon pediatra: così migliorano le diagnosi dell'AI*", op. cit.
[666] G. F. ITALIANO, op. cit.
[667] N. MARINO, "*Intelligenza Artificiale e medicina, serve fare formazione di medici e pazienti: ecco perché*", in *Agendadigitale.eu*, 25 marzo 2019: https://www.agendadigitale.eu/sanita/intelligenza-artificiale-e-medicina-serve-fare-formazione-di-medici-e-pazienti-ecco-perche/
[668] M. LEONI, M. PAPARELLA E S. SOLVI, "*Sanità 4.0, una strategia integrata e globale per la trasformazione digitale*" in *Agendadigitale.eu*, 13 febbraio 2019: https://www.agendadigitale.eu/sanita/sanita-4-0-una-strategia-integrata-e-globale-per-la-trasformazione-digitale/
[669] Z. OBERMEYER, B. POWERS, C. VOGELI E S. MULLAINATHAN, "*Dissecting racial bias in an algorithm used to manage the health of populations*", in "Science", 25 ottobre 2019, Vol. 366, Issue 6464 (pag.447 – 453).
[670] L. MISCHITELLI, "*I pregiudizi dell'intelligenza artificiale in Sanità: perché si creano e come prevenirli*", in *Agendadigitale.eu*, 23 aprile 2021: https://www.agendadigitale.eu/sanita/i-pregiudizi-dellintelligenza-artificiale-in-sanita-perche-si-creano-e-come-prevenirli/

Uno studio pubblicato nel 2019 su *Science Magazine,* ad esempio, ha dimostrato che il sistema *Optum* della *United Health Group* (uno dei più utilizzati in America per individuare i pazienti con esigenze sanitarie complesse) discrimina i pazienti di colore.[671]

Ciò che è importante è, dunque, garantire la diversità dei dati e la loro alta qualità, al fine di sviluppare un sistema intelligente che sia realmente equo e rappresentativo della società. Il che significa anche permettere una circolazione dei dati sanitari tra i vari sistemi per consentire un corretto addestramento dell'algoritmo (nel pieno rispetto della *privacy* dei soggetti); in aggiunta, per assicurare la qualità dei dati è necessario che gli stessi siano accompagnati da un'adeguata documentazione sulla provenienza, sullo scopo, sui limiti e sui potenziali errori. Tutto ciò connesso ad un monitoraggio continuo.[672]

Come sappiamo (e come vedremo nel corso di questo paragrafo), in base alle norme sulla responsabilità medica, se l'operatore sanitario dovesse peccare di negligenza, sarà lui stesso a rispondere dei danni arrecati al paziente. Ma in tal caso potrebbe essere aggiunta anche una responsabilità in capo al programmatore qualora lo stesso non abbia realizzato un *set* di dati abbastanza rappresentativi per addestrare il sistema intelligente utilizzato a sostegno del medico. Solo in tal modo si potrà spronare verso la realizzazione di sistemi sanitari basati sull'intelligenza artificiale che siano più inclusivi e meno discriminatori verso le minoranze.[673]

Per questo è necessario realizzare una modalità di valutazione dei sistemi di intelligenza artificiale al fine di eliminare i rischi di discriminazioni (tramite la disciplina ancora in fase di sviluppo in Europa grazie alla normativa citata nel primo capitolo): così muovendosi, sarà possibile incoraggiare una cultura dell'inclusività attraverso una collaborazione multidisciplinare tra esperti clinici ed esperti di intelligenza artificiale.[674]

Tra le criticità connesse all'utilizzo di sistemi di medicina intelligente, si rinviene altresì l'utilizzo e la gestione dei *big data* sanitari nel rispetto della *privacy* dei pazienti.[675]

Vi sono, poi, le problematiche di *cybersecurity* sanitaria, visto il rilevante rischio che le informazioni sensibili dei singoli siano soggette ad aggressione da parte di *hacker*. Infatti, la *cybersecurity* delle strutture e dei sistemi sanitari è fortemente connessa alla tutela dei pazienti dato che un attacco *cyber*-terroristico, indirizzato verso un sistema sanitario intelligente, potrebbe rischiare di bloccare un intero ospedale mettendo in pericolo numerose vite. Inoltre, potrebbe causare la fruizione di terapie errate e la somministrazione di

---

[671] Z. OBERMEYER, B. POWERS, C. VOGELI E S. MULLAINATHAN, op cit.
Questo perché, il *risk score*, utilizzato per prestare assistenza a pazienti in condizioni di bisogno, dava priorità a individui bianchi, in quanto utilizzava il costo delle cure per valutare la gravità della malattia. Infatti, dato che i soggetti di colore (in genere) ricevono un'assistenza più economica, secondo l'algoritmo, necessitano di meno cure. In tal modo, il sistema va a sovrastimare i costi sanitari per i pazienti di colore; riducendo, di conseguenza, l'accesso alle cure per essi. Dunque, andando a cambiare l'algoritmo, sparirebbe la discriminazione in tale ambito. (F. PASQUALE, "*La cura degli umani*", LUISS University Press, 3 giugno 2021: https://luissuniversitypress.it/nuove-leggi-della-robotica-frank-pasquale-estratto/).
[672] L. MISCHITELLI, "*I pregiudizi dell'intelligenza artificiale in Sanità: perché si creano e come prevenirli*", op. cit.
[673] F. PASQUALE, op. cit.
[674] L. MISCHITELLI, "*I pregiudizi dell'intelligenza artificiale in Sanità: perché si creano e come prevenirli*", op. cit.
[675] Un'interessante inchiesta del "*Wall Street Journal*" del 2019 ha fatto emergere che *Google* in accordo con *Ascension* (*network* americano di istituti sanitari) ha raccolto dati (come i risultati di laboratorio, le diagnosi mediche, i ricoveri e i dati personali) di 50 milioni di pazienti nell'ambito del "*Project Nightingale*" con lo scopo di modernizzare i sistemi di sanità digitale, senza che i medici e i pazienti ne fossero a conoscenza.

farmaci non corrispondenti alle esigenze dei pazienti. Oltre alla tutela della riservatezza dei dati sensibili degli assistiti. [676]

Inoltre, sussiste un notevole rischio di deresponsabilizzazione dell'essere umano dovuto alla traslazione del peso della scelta medica al sistema intelligente, con evidenti ripercussioni sulla concezione dell'autonomia di giudizio e di responsabilità per errore medico.[677]

L'abilità e la sensibilità umana non possono essere sostituite da una macchina. In ultima analisi, il dibattito sul ricorso alla medicina intelligente è ancora aperto: se, da un lato, è indubbio che l'apporto umano è insostituibile, dall'altro non può sottacersi che il supporto delle macchine consente una serie di attività ontologicamente precluse all'essere umano (ad esempio l'analisi di un numero elevato di dati in pochissimo tempo) di fondamentale utilità ai fini della prevenzione delle malattie e della cura del paziente. [678]

D'altro canto, sotto il profilo giuridico, l'utilizzo delle macchine in campo medico, impone un adeguamento del vigente sistema di norme che regola la responsabilità del sanitario (di cui si tratterà nel paragrafo 2.3).

### *2.1 I cyborg e il potenziamento umano*

Nel mondo reale esistono i *cyborg*, esseri fantascientifici al limite tra umani e macchine dotati di forza e intelletto superiori? Sorvolando scenari in stile "*Terminator*",[679] ad oggi esiste una branca della scienza denominata "biorobotica" che consiste nell'inserire nel corpo umano supporti informatici e robotici a fini terapeutici e di potenziamento fisico generando un'ibridazione uomo-macchina grazie alla connessione con il sistema nervoso che permette al paziente di comandare le protesi tramite impulsi celebrali.

Grazie a tale tecnologia è possibile realizzare sistemi di stimolazione celebrale volti al controllo del tremore per soggetti affetti dal *Parkinson,* ripristinare le funzioni motorie perdute e potenziare apparati fisici e cognitivi perfettamente funzionanti grazie a *software* e *hardware* incorporati nel paziente e controllati tramite l'attività celebrale.

Il potenziamento umano ha quindi effetti sia sul piano intellettivo che fisico. Nel primo caso si rientra in un potenziamento dei processi mentali di organizzazione ed elaborazione delle informazioni che incide sulla comprensione e sulla memorizzazione del mondo esterno superando le abilità di un "normale" cervello

[676] SERGIO PILLON, "*Il rischio clinico nella medicina digitale, che cos'è e perché è importante*", in *Agendadigitale.eu*, 11 aprile 2018: https://www.agendadigitale.eu/sanita/il-rischio-clinico-nella-medicina-digitale-che-cose-e-perche-e-importante/
[677] A. LONGO E G. SCORZA, op. cit., Capitolo 6 (pag. 183 - 198) "L'AI in sanità".
[678] F. PASQUALE, op. cit.
[679] J. CAMERON, "*Terminator*", 1984 .

biologico.[680] Esso viene solitamente svolto attraverso l'utilizzo di farmaci che amplificano le potenzialità del cervello, ma nulla impedisce di realizzare gli stessi effetti (se non migliori) grazie a sistemi robotici di stimolazione celebrale. Nel secondo caso si potenziano le prestazioni fisiche ed estetiche del soggetto sano incidendo sulla funzionalità degli arti e del corpo umano grazie alla realizzazione di protesi ed esoscheletri robotici (anche per migliorare le condizioni di vita di persone affette da disabilità o anziane). Il potenziamento fisico può riguardare, inoltre, organi interni come impianti cocleari (un orecchio artificiale utile per ripristinare l'udito), impianti retinali (realizzati al fine di ripristinare la vista). Inoltre, le interfacce cervello-macchina possono essere utilizzate per comunicare all'esterno da parte di soggetti affetti dalla *locked-in syndrome* ed anche per controllare sedie a rotelle e protesi.

In questo modo, grazie a sensori e impianti robotici immessi nell'organismo umano si avrà la possibilità di utilizzare arti elettronici con funzionalità pari o superiori ad arti naturali. Tali sistemi, dotati di sensori per percepire l'ambiente circostante, saranno connessi alla rete al fine di raccogliere ed elaborare rapidamente tutte le informazioni acquisite.

Con la rapidità dell'evoluzione tecnologica in atto ci stiamo spingendo verso il transumanesimo in cui umani interamente biologici si dovranno confrontare con ibridi uomo-macchina. La strada è ancora lunga se pensiamo al fatto che ad oggi l'intelligenza artificiale imita le capacità cognitive umane senza riuscire a sostituirle; però, vedendo le protesi guidate "con il pensiero" gli approdi di tali tecnologie potranno portare a scenari straordinari.

Com'è possibile immaginare, si pongono, in tale ambito, numerose perplessità etico giuridiche: il regime normativo da applicare al soggetto ibrido, i limiti applicabili alla trasformazione del corpo per evitare di superare il concetto di fisicità umana, il significato di salute perfetta, i rischi di discriminazione sociale e le responsabilità in caso di lesioni.

È dunque fondamentale, analizzare preliminarmente la distinzione tra modificazioni terapeutiche o di potenziamento del corpo umano.

Nella pratica, la distinzione tra terapia curativa (ossia il recupero delle funzionalità perdute) e di potenziamento (ossia l'accrescimento delle funzionalità esistenti) non è così limpida come può sembrare. Infatti, numerosi trattamenti medici non sono volti esclusivamente a curare il paziente ma anche a prevenire la malattia (medicina preventiva), ad alleviare le sofferenze (cure palliative), a modificare deformazioni estetiche (chirurgia plastica) e ad impedire gravidanze (dispositivi contraccettivi).

Sicuramente sono considerate legittime le alterazioni del corpo in caso di trattamenti terapeutici: infatti, la Direttiva 93/42/CEE del 14 giugno 1993 considera come dispositivo medico "*qualunque strumento, apparecchio, impianto, software, sostanza o altro prodotto […] destinato ad essere impiegato sull'uomo a fini di diagnosi, prevenzione, controllo, terapia o attenuazione di una malattia; di diagnosi, controllo, terapia,*

[680] 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 1 (pag. 1 – 7) "Biorobotica, interfacce cervello-macchina e potenziamento umano: filosofia precauzionale e euristica di avversione al rischio".

*attenuazione o compensazione di una ferita o di un handicap […]*".[681] Quindi, quando la trasformazione sia connessa all'obiettivo della tutela della salute non si pongono problematiche in relazione al considerare il corpo così potenziato come rientrante sotto lo stesso regime cui è assoggettato il corpo interamente biologico. Ciò perché lo strumento robotico inserito nel corpo ha una ragione strumentale rispetto alla tutela della vita della persona.

Mentre sorgono dei dubbi nel caso di utilizzo per il potenziamento di soggetti sani.[682] Potremmo infatti definire il potenziamento come ogni intervento (non finalizzato a curare una malattia) realizzato allo scopo di migliorare la condizione umano oltre quanto necessario per tutelare una buona salute grazie all'utilizzo di strumenti biorobotici e di farmaci che amplificano il funzionamento fisico e psichico migliorando le capacità umane. [683] Di conseguenza, i dubbi sono connessi alla limitazione degli atti dispositivi del proprio corpo e alla tutela dell'integrità fisica (art. 5 c.c.) nel caso di alterazioni non connesse alla tutela della salute del paziente.

Il diritto alla salute, infatti, implica un'astensione da parte dei consociati a porre in essere comportamenti che possano cagionare ad altri malattie, menomazioni o infermità. Tale diritto è rimesso all'autodeterminazione dell'individuo titolare, a due condizioni: che gli atti dispositivi del proprio corpo non siano contrari alla legge, all'ordine pubblico, e al buon costume; e che non cagionino una diminuzione permanente dell'integrità fisica del soggetto. Infatti, in base all'articolo 5 del Codice Civile, gli unici trapianti ammessi a titolo gratuito tra vivi riguardano il rene, parti di fegato, il pancreas, il polmone e l'intestino; interventi di modificazione dei caratteri sessuali, sterilizzazione volontaria e vendita di parti staccate dal corpo (come i capelli).[684]

Quindi, in base a tali principi, la distinzione si basa sull'esistenza o meno della preesistente patologia connessa alla realizzazione dell'intervento biorobotico.

Nonostante tali tentativi di classificazione, effettuare una distinzione netta tra trattamento curativo e di potenziamento è in realtà fuorviante dato che il secondo comprende il primo; dunque, tale separazione ha natura meramente teorica, in quanto, nella pratica, i due aspetti si intersecano (come nel citato caso delle vaccinazioni).[685] Ciò è dovuto innanzitutto alla complessa definizione del concetto di salute che va a ricomprendere sia stati fisici che psichici (risultando fortemente soggettivo) e inoltre al fatto che ogni terapia comporta in sé un potenziamento dato che l'obiettivo della medicina è migliorare la vita degli esseri umani.[686]

Quindi, nonostante le tecniche di puro potenziamento non curativo sembrerebbero porsi al di fuori della tutela costituzionale del diritto alla salute inteso come diritto fondamentale degli individui, in realtà, in virtù di un'interpretazione estensiva dell'articolo 32 della Costituzione in combinato disposto con l'articolo 3 della

[681] Decreto lgs. 24 febbraio 1997, n. 46 emendato col D.lgs. 25 gennaio 2010, n. 37 – Recepimento Direttiva 2007/47/CE, "*Attuazione della Direttiva 93/42/CEE del 14 giugno 1993 concernente i dispositivi medici*".
[682] E. PALMERINI, op. cit., Paragrafo 7 "Cyborg e human enhancement tra vecchio...".
[683] O. ERONIA, "*Potenziamento umano e diritto penale il "caso" dell'enhancement cognitivo*", Riv. it. dir. e proc. pen., fasc. 3, 2012, pag. 975, Paragrafo 2 "Che cos'è il potenziamento umano?".
[684] A. TORRENTE, P. SCHLESINGER, op. cit., Paragrafo 63 (pag. 132 – 137) "Diritto alla salute".
[685] O. ERONIA, op. cit., Paragrafo 2 "Che cos'è il potenziamento umano?".
[686] 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".

Carta, tutti i trattamenti medici curativi o migliorativi della vita di un paziente con patologie (o senza) sono costituzionalmente tutelati. Così, anche nel caso di terapie sperimentali, in assenza di alternative terapeutiche e con la necessaria presenza del consenso informato del paziente.[687] Infatti, in virtù del principio di autodeterminazione dei soggetti (articolo 2, 13 e 32 della Costituzione, articolo 8 della CEDU), ad oggi, tramite il consenso, alla luce del recente orientamento della Corte Costituzionale[688], il paziente può liberamente esprimere la propria volontà di trattamento medico (con il limite di pratiche illecite).

Il consenso ad un trattamento, sperimentale, di routine, curativo o di potenziamento basato su strumenti dotati di intelligenza artificiale, può sembrare ad oggi scontato e banale. Non è così: si parla di consenso, per la prima volta a livello internazionale nel codice di Norimberga il quale afferma che la persona non può essere mai oggetto di sperimentazione e di ricerca senza il suo consenso. L'essere umano non può essere considerato come uno strumento in mano alla scienza.

Tra le fonti di diritto internazionale deve essere citata inoltre, la Convenzione per la tutela dei diritti dell'uomo e della dignità dell'essere umano rispetto alle applicazioni della biologia (altrimenti nota come Convenzione di Oviedo) del 1997,[689] non ancora sottoscritta dall'Italia ma continuamente citata dalla giurisprudenza, riguarda i diritti etici fondamentali dell'uomo e mira a tutelare i diritti delle persone in campo biomedico.

Tutt'oggi vi è una grande fiducia nei confronti della scienza, infatti, l'articolo 9 Costituzione afferma che la Repubblica ha l'obbligo di promuovere lo sviluppo della cultura e della scienza. L'obbligo della nostra società è di porre strumenti e istituti tali per non impedire lo sviluppo della scienza ma, allo stesso tempo, tutelare le persone.

Le principali problematiche bioetiche in relazione all'utilizzo di sistemi di intelligenza artificiale al fine di potenziare il corpo umano consistono, inoltre, nell'incapacità di prevedere (con certezza scientifica) gli effetti dannosi per la salute umana nel lungo termine. La questione si amplifica in ragione dell'incidenza di tali tecnologie sulle capacità psico-fisiche dei soggetti.

Sul punto, taluni ritengono che un approccio troppo cauto rischi di risolversi in uno sterile appiattimento ed in un ingiustificato limite al progresso medico-scientifico. Altri, invece, basandosi sulla carenza di prove scientifiche in ordine alla dannosità del potenziamento biorobotico, considerano il ricorso a tale tecnologia come un dovere morale: l'utilizzo di tali dispositivi non deve essere, dunque, vietato a priori dall'ordinamento, ma può esserne (e anzi deve esserne) valutata l'applicazione caso per caso.[690]

Bisogna però porsi preliminarmente una questione: cosa significa malattia dal punto di vista giuridico-penalistico?

---

[687] *Idem.*

[688] Sentenza della Corte Costituzionale n. 242 del 25 settembre 2019.

[689] "*Convenzione per la protezione dei Diritti dell'Uomo e della dignità dell'essere umano nei confronti dell'applicazioni della biologia e della medicina: Convenzione sui Diritti dell'Uomo e la biomedicina*", Consiglio d'Europa, Oviedo, 4 aprile 1997.

[690] 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".

Il codice Zanardelli fece proprio il concetto di malattia intesa come qualsiasi alterazione anatomica o funzionale dell'organismo. Negli anni Novanta, con la Sentenza Franciolini, si superò tale concezione affermando che "*il concetto clinico di malattia richiede il requisito essenziale di una riduzione apprezzabile di funzionalità, a cui può anche non corrispondere una lesione anatomica e quello di un fatto morboso in evoluzione a breve o lunga distanza, verso un esito che potrà essere la guarigione perfetta, l'adattamento a nuove condizioni di vita oppure la morte. Ne deriva che non costituiscono malattia, e quindi non possono integrare il reato di lesioni personali, le alterazioni anatomiche, a cui non si accompagni una riduzione apprezzabile della funzionalità*."[691] Successivamente, la Corte di Cassazione nella Sentenza Pagnani, ha affermato che: "*la nozione di malattia giuridicamente rilevante non comprende tutte le alterazioni di natura anatomica, che possono in realtà anche mancare, bensì solo quelle alterazioni da cui deriva una limitazione funzionale o un significativo processo patologico ovvero una compromissione delle funzioni dell'organismo, anche non definitiva ma comunque significativa*".[692] Nello stesso anno, la questione della definizione del concetto giuridico di malattia viene risolta dalle Sezioni Unite della Corte di Cassazione, secondo cui, per malattia deve intendersi un processo patologico evolutivo necessariamente accompagnato da una più o meno rilevante compromissione dell'assetto funzionale dell'organismo, con esclusione delle mere alterazioni anatomiche non incidenti sul profilo funzionale della persona.[693] Il tutto, ovviamente, modulato in base al concetto di salute inteso sia come integrità psico-fisica del soggetto che come autodeterminazione dell'individuo e della sua personalità.[694] Come sappiamo, il diritto alla salute rientra nel catalogo dei diritti fondamentali come tutela dell'integrità psico-fisica della persona in base ad azioni lesive di terzi, e come diritto a ricevere trattamenti sanitari positivi garantiti dallo stato. [695]

In ambito giuridico, si pone inoltre, il dilemma della potenziale disuguaglianza tra soggetti potenziati e non, foriero di, come nel film "*Gattaca* – La porta dell'universo"[696], insormontabili distinzioni di classe, tra individui superiori e persone "normali", che comporterebbero gravi conseguenze sociali e di emarginazione.

Cosa significa essere umani? Come accennato anche nel secondo capitolo, tale questione è forse tra le più dibattute tra i filosofi del diritto. Non è possibile definire un limite netto secondo cui un soggetto potenziato non sia più considerabile come essere umano. Infatti, una persona dotata di protesi biorobotiche non può non essere considerata un essere umano e lo stesso vale per un soggetto che riesce a guidare una sedia a rotelle tramite gli impulsi del proprio cervello.

Per giungere ad una regolamentazione legislativa del potenziamento umano bisogna adottare approcci basati su un'analisi del caso concreto. [697]

---

[691] Sentenza della Cassazione penale, Sez. IV del 14 novembre 1996, n. 10643.
[692] Sentenza della Cassazione penale, Sez. IV del 19 marzo 2008, n. 17505.
[693] Sentenza SS. UU. del 18 dicembre 2008, n. 2437.
[694] O. ERONIA, op. cit., Paragrafo 3.2 "Segue. ... e in ambito giuridico. Uno sguardo particolare alle applicazioni giurisprudenziali".
[695] O. ERONIA, op. cit., Paragrafo 3.3 "Dal diritto alla salute ...".
[696] A. NICCOL, "*Gattaca – La porta dell'universo*", 1997.
[697] O. ERONIA, op. cit., Paragrafo 7 "Quali i modelli per una possibile regolamentazione?".

Non possono, inoltre, sottacersi i rischi potenziali per la sicurezza della persona generati da possibili errori ingegneristici e di programmazione nella realizzazione e addestramento.[698] Inoltre sussistono forti rischi di attacchi alla *cybersecurity* di tali sistemi, essendo gli stessi connessi alla rete. Aggressioni di questo tipo riuscirebbero a generare grossi danni alla salute dei soggetti e alla loro riservatezza.

Le regole relative alla responsabilità penale da prodotto difettoso potrebbero essere applicate solo in caso di malfunzionamento del dispositivo o nel caso in cui non siano stati predisposti appositi dispositivi di sicurezza informatica volti ad impedire aggressioni virtuali tutelando il soggetto possessore. Si deve valutare dunque l'idoneità dei sistemi di *cybersecurity*, necessari per evitare intromissioni esterne, immessi nel prodotto robotico di potenziamento umano. [699]

In tale scenario si inseriscono problematiche in tema di responsabilità penali per malfunzionamenti dei sistemi di trattamento potenziante che possono comportare danni al soggetto stesso o a terzi. Chi risponde in caso di commissione di reati tramite l'utilizzo ad esempio dell'arto bionico? Non sempre è possibile ricostruire precisamente le motivazioni della realizzazione di tale azione (se dovuta ad un malfunzionamento del sistema o se connessa ad una volontà cosciente del soggetto potenziato). [700]

La soluzione al problema potrebbe essere mutuata dal regime di responsabilità applicato agli illeciti commessi dai veicoli autonomi: se il danno arrecato a sé stessi o ad altri deriva dal malfunzionamento del prodotto, la responsabilità può essere ascritta al programmatore o al produttore della macchina. Tuttavia, a differenza delle *self-driving cars*, in tal caso si pone altresì il problema, di cui si dirà al paragrafo 2.3, della responsabilità del medico che non abbia agito correttamente durante l'installazione del sistema intelligente nel corpo del paziente.[701]

### *2.2 I robot* chirurghi

Giunti a questo punto della trattazione, è necessario premettere una rapida classificazione dei *robot* utilizzabili in ambito sanitario al fine di analizzare le loro implicazioni in ambito penale.

Essi possono essere considerati come modulari, di servizio, sociali, mobili e autonomi.

I primi consistono nel potenziare altri sistemi ed essere utilizzati per eseguire diverse funzioni; i secondi si occupano di gestire le attività logistiche di *routine* evitando un aggravio di lavoro agli operatori sanitari; i terzi hanno l'abilità di interagire con altri esseri umani; i quarti permettono di ridurre le interazioni tra il medico e il paziente (fondamentale in questo periodo di pandemia) grazie alla capacità di muoversi secondo binari

[698] E. PALMERINI, op. cit., Paragrafo 6 "*I problemi della sicurezza e della responsabilità nel mercato emergente della robotica*".
[699] E. PALMERINI, op. cit., Paragrafo 7 "Cyborg e human enhancement tra vecchio...".
[700] 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".
[701] E. PALMERINI, op. cit., Paragrafo 6 "*I problemi della sicurezza e della responsabilità nel mercato emergente della robotica*".

stabiliti; ed infine, i *robot* autonomi possiedono sistemi di rilevamento che grazie alla tecnologia *LiDAR* consentono di mappare gli ambienti e muoversi al loro interno.

Oltre a tali categorie, è possibile distinguere anche i *robot* ospedalieri, chirurgici, i *nanorobot*, gli esoscheletri e i *robot* per l'assistenza.

I primi hanno la capacità di distribuire farmaci, campioni di laboratorio e altri materiali all'interno delle strutture sanitarie. I secondi, possedendo braccia meccaniche collegate ad attrezzatura chirurgica, vengono controllati dal medico al fine di realizzare operazioni e interventi con maggiore precisione. [702] In Italia sono oltre cento i *robot* chirurghi. Vengono utilizzati soprattutto per interventi alla prostata, ma è possibile impiegarli in una vastissima gamma di operazioni. [703] I *nanorobot*, hanno l'abilità di muoversi all'interno del corpo umano somministrando farmaci e eseguendo operazioni di microchirurgia. Gli esoscheletri, invece, rilevano segnali elettrici del corpo del paziente in modo tale da permettere il movimento del corpo. Infine, i *robot* per l'assistenza forniscono supporto ai pazienti presso le proprie abitazioni; essi sono fondamentali per il decorso *post* operatorio, per la riabilitazione e per l'assistenza agli anziani e disabili.[704]

Nonostante la rapida evoluzione scientifica, la robotizzazione della sanità è fortemente limitata a causa dei costi notevolmente elevati (una singola apparecchiatura può avere un valore variabile dai due ai tre milioni di euro più le cospicue spese di manutenzione annuali). [705]

Al fine di comprendere meglio il mondo dei *robot* chirurghi e prima di proseguire con la trattaione relativa alla responsabilità medica, giova produrre un semplice esempio. Si pensi al trauma che il singolo vive quando entra in una sala operatoria: sdraiato su un lettino asettico, attorniato dai volti di infermieri e medici coperti dalle mascherine che lo osservano avvolti da un alone di luce bianca al *neon.* Come ci si sentirebbe se, prima di addormentarsi per l'effetto dell'anestesia, non si vedesse nessun volto, nessuna persona, ma solo delle braccia robotiche che si muovono collaborando con un chirurgo che si trovi dall'altra parte del globo?

Ad oggi, gestire operazioni a lunga distanza non è ancora pienamente possibile a causa del ritardo del segnale di arrivo del comando impartito alla macchina, ma grazie all'evoluzione di sistemi basati sull'intelligenza artificiale che renderanno i *robot* chirurghi sempre più autonomi, questo sarà possibile. Per ora si può essere operati da un medico che non si trovi nella sala operatoria ma che realizzi l'intervento grazie al sostegno di un *robot* chirurgo che agirà sul corpo del paziente guidato dall'essere umano:[706] la telechirurgia è la prossima frontiera. [707]

---

[702] G. MAGLIO, "*Robotica in sanità, come vengono impiegati nell'assistenza agli anziani: vantaggi e limiti*", in *Agendadigitale.eu*, 23 aprile 2021: https://www.agendadigitale.eu/sanita/robot/

[703] M. MORUZZI, "*Robot sanitari alla sfida autonomia: la svolta «quinta dimensione»*", in *Agendadigitale.eu*, 21 ottobre 2020: https://www.agendadigitale.eu/sanita/robot-sanitari-alla-sfida-autonomia-la-svolta-quinta-dimensione/

[704] G. MAGLIO, op. cit.

[705] M. MORUZZI, op. cit.

[706] L. MISCHITELLI, "*Così i robot aiutano i chirurghi a operare meglio: progressi e prospettive della tecnologia*", in *Agendadigitale.eu*, 4 giugno 2021: https://www.agendadigitale.eu/sanita/cosi-i-robot-aiutano-i-chirurghi-a-operare-meglio-progressi-e-prospettive-della-tecnologia/

[707] F. CALIMERI E A. MARZULLO, "*Telechirurgia, le operazioni "a distanza" saranno la norma? Ostacoli e prospettive*", in *Agendadigitale.eu*, 1 settembre 2021: https://www.agendadigitale.eu/sanita/telechirurgia-le-operazioni-a-distanza-saranno-la-norma-ostacoli-e-prospettive/

I macchinari intelligenti, impiegati in ambito medico (soprattutto in campo ortopedico), hanno abilità superiori all'essere umano in quanto a precisione. Già fino a poco tempo fa era possibile assistere a interventi realizzati da un chirurgo che pilotava a distanza, tramite *computer*, il *robot* che agiva sul paziente.[708] Non si tratta di *robot* autonomi, è il chirurgo a dirigere i movimenti tramite una console posta in sala operatoria. Ma l'obiettivo è aumentare la distanza tra il *robot* e il chirurgo permettendo, ad esempio di realizzare operazioni urgenti a distanza su soldati in guerra (riducendo le difficoltà e i costi dovuti allo spostamento di persone).

L'aumento di tale distanza risulta però complesso a causa dei possibili cali di connessione che renderebbero impossibile il trasporto di grandi quantità di dati in modo rapido (immagini endoscopiche e dati del *robot*). [709]

Ad oggi, è però possibile osservare *robot* chirurghi che agiscono autonomamente senza la necessità di un'*equipe* umana che li segua grazie agli esperimenti realizzati dai ricercatori dalla *University of California – Berkeley* sul sistema chirurgico "*da Vinci*" costituito da due braccia meccaniche. Tramite la tecnica della *computer vision*, i *robot* riescono ad eseguire le operazioni chirurgiche in autonomia grazie alle reti neurali: il sistema impara analizzando grandi quantità di dati raccolti dalle telecamere degli stessi *robot* chirurghi che registrano video di ogni operazione. In questo modo, seguendo i movimenti realizzati dalla macchina sotto il controllo dell'essere umano, il *robot* autonomo impara come agire senza l'intervento del chirurgo.

L'autonomia del *robot,* in sala operatoria, si misura sulla sua capacità di agire in base all'obiettivo prefissato dal programmatore a cui si aggiungono gli *input* ambientali dati dal caso concreto e dalle indicazioni del medico.

Tale tipologia di *robot* chirurghi non è ancora stata testata su un essere vivente essendo i dispositivi ancora in fase di addestramento ma la rapidità di evoluzione impone di soffermarsi fin da subito sui risvolti etico-giuridici connessi al fenomeno. Il *robot* riesce infatti a superare la destrezza e la velocità umana. In questo modo, però, si rischierebbe di disumanizzare le operazioni chirurgiche eliminando del tutto l'elemento umano. L'obiettivo in realtà è un altro: sostenere i medici e aumentare il tasso di successo delle operazioni. [710] La robotizzazione della sanità e delle operazioni chirurgiche non ha lo scopo di sostituire i medici ma solo di supportarli: è infatti il sanitario ad avere il controllo e a decidere le azioni da compiere.

Per ora, i *robot* agiscono sempre sotto lo stretto controllo dell'essere umano e non in modo pienamente autonomo. Ed è, quindi, già possibile assistere alla presenza di *robot* all'interno delle sale operatorie, in sede di diagnosi, fisioterapia e assistenza al paziente. Con l'evoluzione dell'intelligenza artificiale potremmo giungere alla realizzazione di *robot* sempre più autonomi rispetto alle indicazioni del chirurgo (che arriverebbe ad avere un ruolo di mero controllore della correttezza della procedura, per intervenire solo in caso di emergenza). [711]

---

[708] L. MISCHITELLI, "*Così i robot aiutano i chirurghi a operare meglio: progressi e prospettive della tecnologia*", op. cit.
[709] F. CALIMERI E A. MARZULLO, op. cit.
[710] L. MISCHITELLI, "*Così i robot aiutano i chirurghi a operare meglio: progressi e prospettive della tecnologia*", op. cit.
[711] M. MORUZZI, op. cit.

Le problematiche giuridiche sono inoltre simili a quelle affermate precedentemente in tema di auto a guida autonoma: l'intelligenza artificiale non riesce a gestire l'imprevisto come l'essere umano.[712]

Infatti, come per le auto a guida autonoma, anche per i *robot* chirurghi è stata realizzata una scala di tre livelli di autonomia. Al primo livello, i sistemi seguono i comandi imposti dal medico, al secondo livello i *robot* sono in grado di realizzare una procedura chirurgica sotto la supervisione umana e al terzo, invece, riescono a svolgerla senza alcun controllo umano.

È dunque importante comprendere come debbano essere interpretati gli *standard* medici di diligenza oggetto delle linee guida e delle buone pratiche clinico-assistenziali posti alla base della colpa medica (*ex* art. 590 *sexies* c.p.).

Infatti, per quanto riguarda i sistemi sottoposti al controllo umano, la responsabilità è da attribuire al medico che dirige il *robot*. Se invece, il danno arrecato dovesse dipendere dal difetto del macchinario, si dovrà valutare la disciplina relativa alla responsabilità da prodotto difettoso (come per i veicoli autonomi).[713]

### *2.3 La responsabilità medica in relazione all'utilizzo dei sistemi intelligenti: profili penalistici*

Come abbiamo analizzato fino ad ora, i sistemi di intelligenza artificiale applicati all'ambito medico possono avere risvolti sorprendenti ma al contempo, in caso di errori è importante comprendere i diversi piani di imputazione della responsabilità del sanitario che li ha utilizzati.

I *robot* chirurghi possono commettere errori che, anche se infinitesimali, possono determinare la vita o la morte del paziente. Anche i dispositivi medici intelligenti impiegati nelle diagnosi possono essere soggetti ad errori che possono avere conseguenze disastrose per la salute del paziente. Infatti, è sufficiente cambiare pochi *pixel* di un'immagine clinica per ingannare il sistema facendogli credere che sussista una malattia in realtà assente. Lo stesso vale per piccole variazioni sulla descrizione delle condizioni del paziente che potrebbero portare a diagnosi completamente opposte. I motivi che possono spingere taluno ad ingannare un sistema intelligente sono dei più disparati: dalla frode assicurativa alla commissione del delitto perfetto. [714]

Avendo più volte escluso una responsabilità in capo al sistema stesso, è necessario individuare un soggetto persona fisica cui ascrivere la responsabilità. Infatti, i sistemi intelligenti impiegati in ambito sanitario, verranno considerati come strumenti nelle mani del medico.

---

[712] L. MISCHITELLI, "*Così i robot aiutano i chirurghi a operare meglio: progressi e prospettive della tecnologia*", op. cit.
[713] C. PIERGALLINI, op. cit., Paragrafo 3 "Danno da prodotto e intelligenza artificiale sotto il controllo umano".
[714] A. LONGO E G. SCORZA, op. cit., Paragrafo 3 (pag. 88 - 147) "Vantaggi e rischi non economici dell'AI - I rischi etici, sociali e politici dell'AI - Algoritmi che discriminano".

Quindi, nel momento in cui un sistema intelligente consigli ad un operatore sanitario di seguire un determinato trattamento per il paziente, il soggetto avrà la possibilità di scegliere se affidarcisi ciecamente o se discostarsi da tali indicazioni facendo riferimento alle sue esperienze o a quelle di altri esperti umani.

Nel primo caso, se la macchina ha commesso un errore, questo può dipendere da un errato inserimento dei dati da parte dell'operatore (nel qual caso potrà rinvenirsi una responsabilità del sanitario) oppure, se il medico ha operato correttamente, l'errore può dipendere da un difetto del *software* (in tal caso risponderà il produttore ai sensi della direttiva 85/374/CEE[715]).

Problemi più complessi sussistono nel momento in cui il sistema proponga una terapia dannosa senza che sussista alcun difetto nel *software*. Tali situazioni dipendono dal fatto che i *robot* sono dotati del meccanismo di *machine learning* consistente in un apprendimento basato sull'esperienza. In tal caso il medico che ha seguito le indicazioni errate sarà considerato responsabile se il danno è frutto di una sua imperizia nel controllo e nella valutazione delle indicazioni fornitegli basandosi sul criterio dell'*"homo eiusdem condicionis"*.

Nel secondo caso, se invece il medico decide di discostarsi dall'operato del sistema intelligente, dovrà giustificare le sue ragioni. Se in tal modo dovesse generare un danno al paziente, sarà il medico stesso ad essere considerato responsabile delle lesioni arrecate. Ciò rende complesso per il medico discostarsi dalle indicazioni fornite dal *robot*, con conseguente applicazione implicita del principio "*in dubio, pro machina*": il rischio, chiaramente, è quello di un eccessivo appiattimento dell'operato del medico. In sostanza, il totale affidamento del medico al sapere delle macchine alimenta la c.d. medicina difensiva, che, tra le altre cose, si pone in contrasto con un principio cardine in tema di assistenza medico-sanitaria: "*primum non nocere*".[716]

Viceversa, è importante che il medico sappia gestire il rischio clinico cercando, al contempo, di migliorare le proprie prestazioni grazie ad un costante aggiornamento e ad una cautela e professionalità nell'erogazione delle cure, al fine di garantire la sicurezza dei pazienti.[717]

A seguito di un trattamento medico arbitrario, il sanitario potrà rispondere per lesioni o per omicidio preterintenzionale (se il paziente decede).

Sussistono diverse tesi connesse alla responsabilità del medico. Secondo un primo orientamento la posizione del sanitario, essendo finalizzata alla tutela della salute del paziente, sarebbe esente da responsabilità (a prescindere dall'esito del trattamento). Il secondo orientamento, invece, pone una distinzione tra esito fausto o infausto: l'esito fausto (o positivo) consiste in un miglioramento della salute del paziente e dunque, non sussistendo il requisito della malattia (necessario per configurare il delitto di lesioni) non può essere addossata la responsabilità al medico (anche se il paziente non aveva richiesto quello specifico trattamento). Per esito infausto (o negativo) si fa riferimento ad un peggioramento della situazione di salute del paziente a seguito

[715] 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.
[716] 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*»".
[717] SERGIO PILLON, "*Il rischio clinico nella medicina digitale, che cos'è e perché è importante*", op. cit.

della realizzazione dell'azione medica: in tal caso sussiste la malattia o la morte, e di conseguenza si considererà responsabile il medico se ha posto in essere tali azioni con dolo (volontariamente) o con colpa (se l'operato del sanitario ha violato le regole del corretto operare). Una terza tesi va a considerare come esito infausto qualunque tipologia di azione realizzata sul corpo del paziente in assenza del suo consenso e ciò comporterebbe sempre una responsabilità in capo al medico. L'ultima tesi, anche detta rigorista, considera ogni trattamento medico senza il consenso del paziente come delitto di lesioni perché c'è sempre oggettivamente un'alterazione dell'organismo del paziente.

Nel 2008 le Sezioni Unite[718] della Cassazione accolgono la tesi che differenzia tra esito fausto (assenza lesioni personali perché assenza di malattia) e esito infausto che comporta una responsabilità del medico in caso di colpa (per non rispetto delle *leges artis* e del consenso) o per dolo (se manca un interesse curativo o se il sanitario agisce consapevolmente in assenza del consenso).

Analizzando rapidamente l'evoluzione storica della disciplina della responsabilità medica, inizialmente, si graduava la colpa considerando responsabile il medico solo per colpa grave o dolo, successivamente il grado della colpa è stato eliminato da tale contesto in quanto rilevante solo in caso di commisurazione della pena in concreto e dunque la disciplina è diventata più rigida e meno garantista per il medico. Per queste ragioni, in ambito sanitario si è assistito ad un drastico aumento del fenomeno della medicina difensiva che consiste nella somministrazione di analisi diagnostiche ridondanti e superflue, onerose e stressanti per il paziente al solo scopo di evitare di poter ricadere in forme di responsabilità colposa del medico e non con l'obiettivo di assicurare la salute del paziente.[719]

Per tali ragioni, con il decreto-legge Balduzzi[720], convertito in legge nel novembre 2012, si tentò di individuare dei criteri di limitazione della responsabilità.

Secondo la citata legge, il medico che si attiene alle linee guida e alle buone pratiche clinico-assistenziali sarà responsabile solo per colpa grave, mai per colpa lieve. Le riflessioni che si sollevarono in relazione alla legge riguardavano due aspetti. In primis, l'indeterminatezza del concetto di colpa lieve, in relazione al quale sono state poste questioni di illegittimità costituzionale per denunciare l'irragionevolezza di escluderla solamente in ambito medico. In secondo luogo, l'incertezza dell'affidabilità delle linee guida e delle buone pratiche clinico-assistenziali. Infatti, le stesse non sono norme precise, ma raccomandazioni da seguire in alcune situazioni fornite da società scientifiche e dalle associazioni tecnico-scientifiche delle professioni sanitarie.[721] Le definizioni, che vedremo poco più avanti, sottolineano come sia le linee guida che le buone pratiche clinico-assistenziali siano sicuramente fondamentali perché permettono di innalzare gli *standard* di

[718] Sentenza SS. UU. penali del 18 dicembre 2008, n. 2437.
[719] C. CUPELLI, "*I profili penali della legge Gelli – Bianco (art. 590 sexies c.p.)*", estratto dal Volume a cura di G. IUDICA "*La tutela della persona nella nuova responsabilità sanitaria – Responsabilità civile e previdenza*", Giuffrè Francis Lefebvre editore, Milano 2019; Paragrafo 1 (pag. 237 – 239) "Introduzione".
[720] Decreto-legge 13 settembre 2012, n. 158 (c.d. Decreto Balduzzi), convertito in legge 8 novembre 2012, n. 189.
[721] MINISTERO DELLA SALUTE - Governo clinico e sicurezza delle cure, "*Sistema nazionale Linee Guida – SNLG*", 21 gennaio 2022: https://www.salute.gov.it/portale/sicurezzaCure/dettaglioContenutiSicurezzaCure.jsp?lingua=italiano&id=4835&area=qualita&menu=lineeguida

perizia esigibile e di uniformare i criteri di valutazione della colpa medica al fine di limitare così l'uso della c.d. medicina difensiva; sono necessari, però, criteri selettivi che ne assicurino l'adeguatezza.[722]

Quindi, a seguito di tali problematiche, nel 2017, la legge c.d. Gelli – Bianco[723] ha introdotto nel codice penale l'articolo 590 *sexies* c.p. al fine di dare certezza all'esonero di responsabilità del medico. Con esso, viene disciplinata la responsabilità colposa per morte o lesioni personali in ambito sanitario tramite l'imposizione di una sanzione penale nel caso in cui si commettano i fatti di cui agli articoli 589 e 590 c.p. nell'esercizio di una professione sanitaria. La norma deroga rispetto ai citati articoli di omicidio e lesioni nell'ipotesi in cui l'evento si verifichi a causa di imperizia. In tal caso, la punibilità è esclusa qualora siano state rispettate le raccomandazioni previste dalle linee guida come definite e pubblicate ai sensi di legge o dalle buone pratiche clinico-assistenziali; sempre che tali raccomandazioni siano adeguate al caso concreto.

Al primo comma, l'articolo 590 *sexies* c.p. rinvia alle pene previste per i reati di lesioni e di omicidio colposo nel caso in cui tali fattispecie siano poste in essere in ambito sanitario. Nel secondo comma, però, esclude la punibilità per imperizia se vengono rispettate le linee guida approvate ai sensi di legge e le buone pratiche clinico-assistenziali che siano adeguate al caso concreto. [724]

La legge Gelli – Bianco fa riferimento al valore e all'autorità del sapere scientifico codificato nelle linee guida (approvate dall'Istituto Superiore di Sanità) e nelle buone pratiche clinico-assistenziali (poste in secondo piano e solo in assenza delle prime). Esse consistono in un insieme variegato di *leges artis* tradotte in regole di diligenza professionale.

Una delle problematiche principali, rilevate in dottrina, in relazione a questa riforma è l'assenza di chiarezza nelle definizioni. [725]

Tentando di dare una definizione di linea guida valida a livello internazionale, può dirsi che essa sia una raccomandazione di esperti basata su esperienze condivise in letteratura medico-scientifica in relazione a comportamenti clinici con lo scopo di assistere i medici nelle pratiche assistenziali da fornire ai pazienti. Di conseguenza, tali raccomandazioni non possono essere considerate di carattere vincolante e consistono in *leges artis* (contenenti regole di prudenza, perizia e diligenza professionale) a cui si deve attenere un medico competente in base alla situazione clinica concreta.[726]

È fortemente dibattuto il reale significato di buone pratiche clinico-assistenziali. [727]

Infatti, neppure a livello internazionale c'è una definizione condivisa di buone pratiche clinico-assistenziali. Per tale ragione, la Corte Suprema di Cassazione è intervenuta più volte in proposito cercando

[722] C. CUPELLI, op. cit., Paragrafo 1 (pag. 237 – 239) "Introduzione".
[723] Legge 8 marzo 2017 n.24 recante "*Disposizioni in materia di sicurezza delle cure e della persona assistita, nonché in materia di responsabilità professionale degli esercenti le professioni sanitarie*".
[724] C. CUPELLI, op. cit., Paragrafo 2 (pag. 239) "I risvolti penalistici della legge Gelli – Bianco: un primo inquadramento".
[725] F. CEMBRANI, "*Irresponsabilità penale del medico e qualità metodologica del sapere scientifico codificato medical and methodological quality of the scientific code*", Rivista Italiana di Medicina Legale (e del Diritto in campo sanitario), fasc.2, 1 aprile 2019, pag. 645; Paragrafo 1 "Introduzione".
[726] F. CEMBRANI, op. cit., Paragrafo 3 "La codificazione del sapere scientifico".
[727] F. CEMBRANI, op, cit., Paragrafo 1 "Introduzione".

di interpretare in modo corretto il termine. [728] Nel 2013 ha affermato che si tratta di protocolli rigidi e definiti di comportamento clinico la cui violazione ricade sempre in un comportamento definibile imprudente o negligente.[729] Nel 2018 ha chiarito come esse vengano realizzate grazie ad un processo di elaborazione concettuale per offrire indicazioni ai medici su come scegliere il trattamento terapeutico più appropriato al caso concreto. [730]

È inoltre fondamentale valutare la qualità del sapere scientifico codificato in base alle modalità di selezione riferite agli *standard* internazionali e ad alcuni criteri principali di analisi di tali raccomandazioni: gli obiettivi e gli ambiti applicativi, il coinvolgimento dei portatori di interesse (gli *stakeholders*), il rigore metodologico, la chiarezza espositiva, l'applicabilità in concreto ed infine, l'indipendenza editoriale di chi le fornisce (valutando gli eventuali conflitti di interessi). [731]

Altro problema è dovuto all'assenza di criteri rigidi e selettivi che possono generare distorsioni connesse ai contrasti tra le scuole scientifiche che, a loro volta, possono portare alla realizzazione di linee guida disomogenee che, invece di aiutare il medico nell'esercizio delle pratiche sanitarie, finiscono per provocare una maggiore incertezza del suo operato.

L'idea di inserire le linee guida e le buone pratiche clinico-assistenziali, all'interno della disposizione che regola la responsabilità colposa in ambito sanitario, è legata al rispetto di vincoli penalistici connessi alla possibilità di essere posti a conoscenza (prima di tenere la condotta) della liceità o meno di quel determinato comportamento garantendo, al contempo, determinatezza alla fattispecie colposa in esame.

I dubbi posti in passato in relazione alla gradazione della colpa sembrerebbero sparire dato che non si nota più un'espressa distinzione tra colpa lieve e grave ma il riferimento alle linee guida e buone pratiche clinico-assistenziali ne fa sorgere di nuovi in ordine al considerare l'elemento soggettivo come colpa generica o come colpa specifica, dato che, da una parte punisce (sempre) l'imprudenza e negligenza mentre dall'altra punisce l'imperizia solo con riferimento al mancato rispetto di tale sapere scientifico codificato.[732]

Paradossalmente è infatti più semplice comprendere quando tale deroga non si applica: in caso di assenza di linee guida e di buone pratiche clinico-assistenziali o nel caso in cui non siano adeguate al caso concreto; nel caso di comportamenti negligenti o imprudenti tenuti dal medico; e nel caso in cui non vengano rispettate tali linee guide e buone pratiche.

Risulta difficile rispondere all'interrogativo: quando si applica?

Nel tentativo di fornire una soluzione al quesito, la giurisprudenza ha espresso diversi orientamenti, due estremi e uno intermedio.

---

[728] F. CEMBRANI, op. cit., Paragrafo 3 "La codificazione del sapere scientifico".
[729] Sentenza della Cassazione Penale, n. 16237 del 2013.
[730] Sentenza della Cassazione Penale, n. 47748 del 2018.
[731] F. CEMBRANI, op. cit., Paragrafo 4 "Sapere scientifico codificato e colpa penale".
[732] C. CUPELLI, op. cit., Paragrafo 3 (pag. 240 – 242) "Il ruolo delle «nuove» linee guida certificate".

Con la Sentenza Tarabori,[733] nel 2017, i giudici di legittimità interpretano l'articolo 590 *sexies* c.p., introdotto dalla legge Gelli-Bianco, come non applicabile: in ambiti non governati da linee guida; ovvero in situazioni in cui, nel caso concreto, le linee guida o le buone pratiche clinico-assistenziali non risultino adeguate alla cura del paziente; infine, in caso di errore nell'esecuzione materiale dell'atto chirurgico pur correttamente impostato secondo le raccomandazioni ufficiali.

Dunque, l'ambito applicativo della norma viene limitato ai casi in cui si verifichino eventi che sono stati disciplinati in modo concretamente appropriato da dette linee guida accreditate.

La Corte fa poi riferimento al momento in cui il medico sceglie le linee guida ma poi non le applica correttamente in fase esecutiva. A causa dei dubbi dovuti all'imperizia in caso di scelta adeguata del sapere scientifico, la pronuncia sottolinea il rilievo, anche in ambito penale, dell'articolo 2236 c.c. (situazioni tecnico scientifiche nuove o situazioni di urgenza) in quanto espressione di un principio di razionalità secondo cui la rilevanza della colpa grave e può essere applicata in ambito penale per valutare l'addebito di imperizia nel caso di effettive difficoltà nel caso concreto. Secondo tale articolo, bisogna valutare attentamente il caso concreto, le difficoltà avute dal medico, l'atipicità della situazione e le motivazioni del medico in modo da valutare la colpa su due piani: soggettivo (la perizia del medico nel caso concreto) e oggettivo (le circostanze particolari in cui il medico ha agito). In tal modo si va a tutelare la classe medica e si evita la medicina difensiva. [734]

Nella Sentenza Cavazza,[735] invece, i giudici di legittimità esprimono il principio secondo cui la riforma, volta a escludere la responsabilità in caso di imperizia lieve, tutelerebbe maggiormente il diritto alla salute ponendo un limite al fenomeno della c.d. medicina difensiva. Secondo tale pronuncia, con la novella si punisce solo l'imperizia nella fase esecutiva che sussiste successivamente ad una corretta scelta delle linee guida.

La sentenza si discosta dalla Tarabori nel considerare il tema della gravità della colpa ritendendo applicabile la causa di non punibilità anche nel caso di colpa grave in quanto nella norma non si fa riferimento al grado della colpa.

Viene così introdotta una causa di esclusione della punibilità per la sola imperizia nel caso in cui il sanitario rispetti le *leges artis* adeguate al caso concreto. L'unica ipotesi di rilevanza penale dell'imperizia sanitaria riguarderà il caso di applicazione di linee guida inadeguate al caso concreto mentre il medico rimarrà esente da responsabilità anche qualora sia incorso in un'imperita applicazione di adeguate linee guida scelte correttamente.

I giudici della Suprema Corte affermano infine il principio secondo cui l'articolo 590 *sexies* c.p. prevede una causa di non punibilità solo nel caso di imperizia e a prescindere dal grado della colpa.

---

[733] Sentenza della Cassazione Penale, Sez. IV, n. 28187 del 20 aprile – 7 giugno 2017.

[734] C. CUPELLI, op. cit., Paragrafo 4 (pag. 242 – 244) "Una fattispecie ad hoc per la responsabilità penale dell'esercente la professione sanitaria: il nuovo articolo 590 sexies c.p. e il contrasto sul perimetro applicativo".

[735] Sentenza della Cassazione Penale, Sez. IV, n. 50078 del 19 ottobre – 31 ottobre 2017.

Dunque, la Corte estende il regime di favore a tutte le ipotesi di imperizia nella fase esecutiva di linee guida adeguate e pertinenti al caso concreto.[736]

A risolvere il contrasto interpretativo sono intervenute le Sezioni Unite penali della Corte di Cassazione con la Sentenza n. 8770 del 2018[737] la quale ha stabilito che l'esercente la professione sanitaria risponde, a titolo di colpa, per morte o lesioni personali derivanti dall'esercizio di attività medico-chirurgica: a) se l'evento si è verificato per colpa (anche "lieve") da negligenza o imprudenza; b) se l'evento si è verificato per colpa (anche "lieve") da imperizia quando il caso concreto non è regolato dalle raccomandazioni delle linee-guida o dalle buone pratiche clinico-assistenziali; c) se l'evento si è verificato per colpa (anche "lieve") da imperizia nella individuazione e nella scelta di linee-guida o di buone pratiche clinico-assistenziali non adeguate alla specificità del caso concreto; d) se l'evento si è verificato per colpa "grave" da imperizia nell'esecuzione di raccomandazioni di linee-guida o buone pratiche clinico-assistenziali adeguate, tenendo conto del grado di rischio da gestire e delle speciali difficoltà dell'atto medico. [738]

L'applicazione di tali linee guida deve sempre essere connessa alle specificità del caso concreto al fine di poter garantire un margine di discrezionalità al medico che potrà discostarsi da tali raccomandazioni in base alle specifiche esigenze cliniche del singolo paziente. Il problema è che sussiste un grande margine di discrezionalità del giudice in relazione all'accertamento di tale discostamento ma il giudizio di adeguatezza va valutato in prospettiva *ex ante* in base alle circostanze (linee guida, buone pratiche clinico-assistenziali e fattori del caso concreto) conosciute o conoscibili dal medico nell'atto di presa in carico del paziente.[739]

La legge Gelli-Bianco è stata fortemente criticata anche perché per il *reo* rimane più favorevole il decreto-legge Balduzzi. Infatti, con riferimento ai fatti commessi prima della legge, in caso di rispetto delle *leges artis* ma con negligenza e imprudenza lieve, la legge Balduzzi è più favorevole perché esclude la colpa lieve e nel caso in cui si rispettino le *leges artis* ma con imperizia, se sussiste un errore nel momento selettivo, è più favorevole la Balduzzi, se invece sussiste un errore con imperizia lieve in fase esecutiva, il soggetto sarà esente da punibilità sia per la legge Balduzzi che per la legge Gelli-Bianco.

La decisione delle Sezioni Unite ha sollevato rilevanti perplessità dato che sembra tornare a fare riferimento alla gradazione della colpa (anche se non espressi nella norma), un elemento nuovo in *malam partem* con effetti limitativi della punibilità.[740]

All'esito della sintetica analisi compiuta, è quindi possibile affermare che i sistemi di intelligenza artificiale applicati in ambito sanitario possono avere sicuri effetti sulla responsabilità del medico. Ad oggi, come affermato precedentemente, i *robot* chirurghi, gli algoritmi di sostegno diagnostico e di monitoraggio

[736] C. CUPELLI, op. cit., Paragrafo 4 (pag. 242 – 244) "Una fattispecie ad hoc per la responsabilità penale dell'esercente la professione sanitaria: il nuovo articolo 590 sexies c.p. e il contrasto sul perimetro applicativo".
[737] Sentenza della Cassazione Penale, Sezioni Unite, n. 8770 del 22 febbraio 2018: https://www.biodiritto.org/ocmultibinary/download/3259/31842/8/ba374071a4619dfb28edf5d0587fb316/file/cass-pen-sez-un-2018-8770.pdf
[738] F. CEMBRANI, op. cit., Paragrafo 4 "Sapere scientifico codificato e colpa penale".
[739] C. CUPELLI, op. cit., Paragrafo 3 (pag. 240 – 242) "Il ruolo delle «nuove» linee guida certificate".
[740] C. CUPELLI, op. cit., Paragrafo 5 (pag. 244 – 253) "La soluzione delle Sezioni Unite e qualche perplessità residua".

non sono interamente autonomi e sottostanno ad un costante controllo dell'operatore, di conseguenza, bisognerà valutare caso per caso la perizia del medico e il suo rispetto delle linee guida e delle buone pratiche clinico-assistenziali in relazione all'utilizzo dell'intelligenza artificiale.

Infatti, nel giugno 2021 è stato realizzato dall'Organizzazione Mondiale della Sanità un *report* (*Ethics and Governance of Artificial Intelligence for Health*)[741], suddiviso in nove sezioni, contenente indicazioni e linee guida in merito all'applicazione e uso dell'intelligenza artificiale in campo medico. All'interno del documento viene fornita una definizione, di intelligenza artificiale e di *big data* sanitari, utile per individuare le principali applicazioni di tale tecnologia in medicina (come la ricerca, la gestione dei sistemi sanitari e il monitoraggio della salute pubblica). Vengono poi trattate le leggi e i principi (tutela dei diritti fondamentali, protezione dei dati personali, norme sull'utilizzo dei dati sanitari), anche di stampo etico, da dover rispettare nell'utilizzo della tecnologia innovativa. L'OMS individua inoltre sei principi etici chiave da osservare nell'utilizzo dell'intelligenza artificiale in ambito sanitario: la protezione dell'autonomia degli esseri umani; la promozione del benessere delle persone, della loro sicurezza e del pubblico interesse; la predisposizione di adeguate garanzie in materia di trasparenza, spiegabilità e comprensibilità dei sistemi di intelligenza artificiale; la responsabilità e l'affidabilità delle tecnologie impiegate; l'attuazione di criteri di inclusione ed equità nello sviluppo e uso dell'intelligenza artificiale; ed infine, la realizzazione di sistemi intelligenti sostenibili ed efficaci in relazione alle esigenze dei soggetti.

Vengono poi analizzati i vari rischi (come i *bias*, la *privacy*, la responsabilità, l'impatto sul mondo del lavoro, sul mercato e sul cambiamento climatico) dovuti all'introduzione dei sistemi intelligenti in ambito medico che possono essere limitati grazie alla realizzazione di strumenti giuridici.

Infatti, le ultime tre sezioni analizzano come i diversi *stakeholders* debbano programmare in modo etico i sistemi intelligenti in ambito sanitario tramite il coinvolgimento del pubblico e la dimostrazione della loro affidabilità, predisponendo appositi modelli di monitoraggio e di valutazione; di chi debba essere considerato responsabile, in base alle attuali normative, in caso di uso di tali tecnologie in ambito sanitario; di quali, infine, siano gli elementi necessari per realizzare una *governance* in relazione all'uso dell'intelligenza artificiale in questo contesto (ossia riferita ai dati, alla gestione dei costi e benefici, al settore privato e pubblico).

Infine, il *report* fornisce linee guida pratiche che devono essere seguite da parte di programmatori, ministeri della salute e operatori sanitari.[742]

Soffermandoci per un attimo sul tema della responsabilità per l'uso dell'intelligenza artificiale nell'assistenza clinica. L'OMS, nel rispondere alle domande sulla responsabilità del medico che si sia affidato al suggerimento proposto dal dispositivo intelligente, successivamente tradottosi in un errore, sottolinea le problematiche relative sia all'eccessiva limitazione sia all'eccessiva liberalizzazione dell'utilizzo dei sistemi

[741] WHO, WORLD HEALTH ORGANIZATION, "*Ethics and Governance of Artificial Intelligence for Health: WHO Guidance*", Ginevra, 2021: https://www.who.int/publications/i/item/9789240029200

[742] M. FASAN, "*OMS – Ethics ad Governance of Artificial Intelligence for Health: WHO Guidance*", in BioDiritto rivista online – Università di Trento, Facoltà di Giurisprudenza, del 28 giugno 2021: https://www.biodiritto.org/Biolaw-pedia/Docs/OMS-Ethics-ad-Governance-of-Artificial-Intelligence-for-Health-WHO-Guidance

intelligenti. Infatti, qualora si dovesse penalizzare il medico a causa di un errore dovuto all'essersi affidato al dispositivo intelligente, si limiterebbe fortemente l'evoluzione scientifica e tecnologica in abito sanitario che frenerebbe lo sviluppo di tali tecnologie in campo medico. Al contempo, qualora si giustificasse sempre il medico che pone in essere un'azione lesiva sul paziente, per il solo fatto di essersi affidato all'intelligenza artificiale, si causerebbe un'eccessiva automatizzazione delle scelte mediche.

Inoltre, nel medesimo *report*, l'OMS sottolinea l'importanza di ampliare la disciplina relativa alla responsabilità da prodotto anche in relazione ai sistemi di intelligenza artificiale in modo tale da far rispondere anche il produttore qualora il danno subito dal paziente dipenda da difetti del *software*. Infatti, riconoscendo tale responsabilità si garantirebbe l'adozione da parte degli sviluppatori di tutte le misure necessarie per ridurre la possibilità di errore. Allo stesso modo, il documento evidenzia come tale responsabilità debba essere adeguata e aggiornata in relazione all'apprendimento automatico dei sistemi intelligenti (non sempre prevedibile dal programmatore).

L'OMS dà inoltre la possibilità di interrompere il nesso di causalità tra le azioni del programmatore e la lesione del paziente nel caso di intervento del medico, considerato come "*learned intermediary*", qualora il programmatore/produttore abbia fornito adeguate informazioni sui rischi connessi all'utilizzo del dispositivo medico intelligente.

L'OMS, in quanto organizzazione internazionale non può quindi imporre una responsabilità del sanitario uguale per tutte le Nazioni e dunque invita i legislatori dei singoli Stati ad introdurre sistemi di responsabilità che siano idonei a disciplinare anche tale fenomeno. Il *report* elenca infatti diverse raccomandazioni tra cui: la necessità che le agenzie internazionali garantiscano un costante aggiornamento delle loro linee guida in relazione all'intelligenza artificiale.[743]

Nelle raccomandazioni specifiche per i singoli operatori sanitari, il *report* citato suddivide l'utilizzo dell'intelligenza artificiale in tre aree tematiche: l'adeguatezza della tecnologia di intelligenza artificiale, in cui si pone al centro la prevenzione del rischio, la trasparenza, la tutela della *privacy* e l'eliminazione dei *bias* per ridurre gli errori e le discriminazioni in fase di utilizzo; l'adeguatezza del contesto in cui la stessa verrà impiegata, grazie ad una valutazione del rapporto rischi-benefici; e infine, l'effettivo utilizzo del sistema da parte dell'operatore sanitario, garantendo centralità al giudizio umano mai sostituito dai suggerimenti forniti dal dispositivo intelligente.[744]

Sarebbe dunque opportuno realizzare delle linee guida accreditate specifiche per l'intelligenza artificiale per far si che il medico possa affidarsi ad esse potendosi così muovere all'interno di binari stabiliti e certi per non rischiare di ricadere in responsabilità dovendosi affidare a linee guida e buone pratiche clinico-assistenziali frammentarie e non aggiornate. Solo tramite una corretta formazione del sanitario e attraverso

[743] WHO, WORLD HEALTH ORGANIZATION, "*Ethics and Governance of Artificial Intelligence for Health: WHO Guidance*", op. cit., Paragrafo 8 (pag. 76 – 80) "Liability regimes for artificial intelligence for health".
[744] WHO, WORLD HEALTH ORGANIZATION, "*Ethics and Governance of Artificial Intelligence for Health: WHO Guidance*", op. cit., Paragrafo A3 (pag. 146 – 148) "Considerations for health-care institutions and provider".

realizzazione di linee guida al passo con i tempi, in aggiunta al rispetto, da parte dei programmatori di tali sistemi, dei principi guida forniti dall'Organizzazione Mondiale della Sanità, si potrà prevenire il rischio che l'utilizzo di sistemi di intelligenza artificiale in campo medico possa provocare morte o lesioni.

Allo stato attuale sembra non esserci spazio per un'esenzione della responsabilità dell'operatore sanitario dovuta ad errore del sistema intelligente, in quanto non è possibile imputarla al *robot* essendo effettivamente un mero strumento nelle mani del medico; salvo, naturalmente, ipotesi di responsabilità da prodotto difettoso nel caso in cui i danni siano dovuti ad un difetto del *software*.

## ***3. Cybercrime e i reati informatici "in senso ampio"***

Le scelte di politica legislativa legate all'incriminazione di comportamenti connessi all'uso di *software* pericolosi provengono principalmente da fonti sovrannazionali o europee che vengono trasfuse nei codici penali nazionali. Per questo nel diritto penale informatico troviamo una forte vicinanza tra le varie legislazioni statali. Non sempre gli organismi sovrannazionali impongono l'incriminazione penale, anzi, spesso lasciano facoltà ai singoli Stati sulla scelta di quale tipologia di sanzione usare (penale, civile o amministrativa), dunque, di frequente, lo stesso comportamento viene sanzionato in modo differente nei diversi ordinamenti.

Gli strumenti utilizzati dal Consiglio d'Europa per incentivare gli Stati a incriminare le condotte inerenti ai programmi informatici sono: la Convenzione europea sulla tutela dei servizi ad accesso condizionato e dei servizi di accesso condizionato, firmata a Strasburgo il 24 gennaio 2001[745] (punisce tutte le condotte poste in essere a fini commerciali riguardanti la fabbricazione, produzione, importazione, distribuzione, vendita, noleggio, possesso, installazione, manutenzione, sostituzione di dispositivi illeciti) e la *Cybercrime Convention* firmata a Budapest il 23 novembre 2001[746] (punisce chi, al fine di commettere un reato informatico vende, distribuisce, mette a disposizione o possiede codici di accesso, *password o software* principalmente concepiti o destinati alla commissione di reati informatici contro la riservatezza, la disponibilità e l'integrità di dati o di sistemi informatici). Nei dispositivi illeciti rientrano anche i programmi informatici concepiti o adattati al fine di rendere possibile l'accesso a sistemi informatici senza l'autorizzazione del fornitore di servizi. [747]

In ambito dell'Unione europea, la Direttiva 2009/24/CE sulla tutela giuridica dei programmi per elaboratore obbligava gli Stati a punire le condotte aventi ad oggetto qualsiasi mezzo usato al fine di facilitare

---

[745] Council of Europe, "*European Convention on the Legal Protection of Services based on, or consisting of, Conditional Access*", Strasbourg, 24.1.2001: https://rm.coe.int/1680080623

[746] Council of Europe, "*Convention on Cybercrime*", Budapest, 23.11.2001: https://rm.coe.int/1680081561

[747] R. WENIN E G. FORNASARI, "*Diritto penale e modernità - Le nuove sfide fra terrorismo, sviluppo tecnologico e garanzie fondamentali*", Atti del convegno Trento, 2 e 3 ottobre 2015; I. SALVADORI, IV sessione, (pag. 309 – 439), 2017 Articolo "*Il diritto penale dei software «a duplice uso»*"; Paragrafo 5 (pag 382 - 396) "Gli obblighi di incriminazione dei software «a duplice uso» nelle fonti sovranazionali*"*.

la rimozione non autorizzata o l'elusione di dispositivi tecnici applicati a protezione di un programma informatico. La Direttiva 2014/62/UE[748] relativa al rafforzamento della tutela per mezzo di sanzioni penali e altre sanzioni contro la falsificazione di monete in relazione all'introduzione dell'euro, prevede l'obbligo di punire il fatto di produrre fraudolentemente, ricevere, ottenere o possedere strumenti o programmi informatici che sono per loro natura atti a falsificare o alterare monete. La Decisione quadro 2001/413/GAI[749], relativa alla lotta contro le frodi e le falsificazioni di mezzi di pagamento diversi dai contanti, impone sanzioni per chi produce fraudolentemente, riceve, ottiene, vende, cede ad altri programmi per il *computer* appositamente allestiti per la contraffazione o falsificazione di strumenti di pagamento, ovvero programmi per il *computer* il cui scopo sia la commissione di una frode informatica. La Direttiva 2013/40/UE[750], relativa agli attacchi contro i sistemi di informazione, punisce l'intenzionale e illegittima fabbricazione, vendita, approvvigionamento per uso, importazione, distribuzione, messa a disposizione di un programma per *computer* destinato a commettere un illecito contro la riservatezza informatica, l'integrità o la disponibilità di dati o di sistemi informatici.

I programmi informatici possono essere divisi in tre gruppi: quelli principalmente concepiti o adattati per la commissione di un reato; quelli oggetto di una promozione, pubblicità o commercializzazione con la finalità di commettere un reato; quelli il cui scopo consiste nella commissione di un reato.[751]

Dunque, per mezzo dell'incriminazione dei comportamenti antecedenti aventi ad oggetto un *software* pericoloso, il legislatore anticipa la punibilità rispetto alla consumazione di un altro più grave delitto sanzionando condotte che si collocano in una fase antecedente al secondo reato. Ciò al fine di permettere alle autorità di svolgere indagini in un momento anteriore alla commissione dell'illecito più grave e al fine di creare una barriera protettiva in modo da ostacolare la commissione di fatti che minacciano un interesse superiore.

Negli ultimi anni, l'automazione dei *cybercrime* ha permesso lo sviluppo di programmi informatici che eseguono in automatico diverse condotte criminose senza interazione con l'essere umano (tranne per quanto riguarda l'installazione del *software*). Con l'avvento dell'intelligenza artificiale si spera che in un prossimo futuro cambi quindi la politica criminale, fino ad ora incentrata sui reati informatici commessi esclusivamente dall'essere umano.[752]

---

[748] Direttiva 2014/62/UE del Parlamento Europeo e del Consiglio, sulla "*protezione mediante il diritto penale dell'euro e di altre monete contro la falsificazione e che sostituisce la decisione quadro 2000/383/GAI del Consiglio*", del 15 maggio 2014: https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32014L0062&from=EN

[749] Decisione Quadro del Consiglio (2001/413/GAI) "*relativa alla lotta contro le frodi e le falsificazioni di mezzi di pagamento diversi dai contanti*" del 28 maggio 2001: https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32001F0413&from=IT

[750] Direttiva 2013/40/UE del Parlamento Europeo e del Consiglio, "*relativa agli attacchi contro i sistemi di informazione e che sostituisce la decisione quadro 2005/222/GAI del Consiglio*", del 12 agosto 2013: https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32013L0040&from=SK

[751] R. WENIN E G. FORNASARI, "*Diritto penale e modernità - Le nuove sfide fra terrorismo, sviluppo tecnologico e garanzie fondamentali*", Atti del convegno Trento, 2 e 3 ottobre 2015; I. SALVADORI, IV sessione, (pag. 309 – 439), 2017 Articolo "*Il diritto penale dei software «a duplice uso»*"; Paragrafo 5 (pag 382 - 396) "Gli obblighi di incriminazione dei software «a duplice uso» nelle fonti sovranazionali*"*.

[752] R. WENIN E G. FORNASARI, op. cit. I. SALVADORI, IV sessione, (pag. 309 – 439), 2017 Articolo "*Il diritto penale dei software "a duplice uso"*"; Paragrafo 9 (pag. 423 - 443) "Sui presupposti per una legittima incriminazione dei software «a duplice uso»".

Ad oggi, non esiste una normativa penale *ad hoc* ed idonea a punire i reati commessi dall'intelligenza artificiale. Si ha esclusivamente la possibilità di estendere la disciplina già esistente in merito ad altre fattispecie penali, anche a tali innovative situazioni criminali.

L'intelligenza artificiale come già affermato, si comporta come *dual-use technology*, potendo essere utilizzata sia a scopi leciti che illeciti (potendosi sostituire in tutto o in parte all'essere umano nelle attività illecite) anche in rete. [753]

Tale tecnologia, grazie all'elevato quantitativo di dati che costantemente immettiamo in rete o su dispositivi connessi ad essa, rischia di essere utilizzata al fine di aggredire molti beni giuridici meritevoli di tutela penale come il patrimonio (truffe mediante *e-mail*, *scam, phishing*, *pharming*, *smishing*), l'onore (diffamazione su *blog* o *social network*), la proprietà intellettuale, l'autodeterminazione in ambito sessuale (*child-grooming, revenge porn* e la diffusione *on-line* di immagini pedopornografiche, la realizzazione di video pornografici tramite l'utilizzo della tecnologia *deep-fake*), la riservatezza (*data-espionage*) e l'integrità dei sistemi informatici (accesso abusivo a sistemi informatici, *hacking* o *cracking,* impedire il corretto funzionamento di sistemi di istituzioni pubbliche tramite *DoS* o *DDoS attacks*). Tali illeciti vengono realizzati grazie all'utilizzo di *malware* (*software* malevoli che si insidiano illegalmente nei sistemi informatici) per giungere alla commissione anche di attacchi *cyber*-terroristici, di *cyber-war* e di *cyber espionage*. [754]

L'intelligenza artificiale può infatti essere utilizzata per realizzare gli attacchi di *spear* o *spray phishing* dato che gli algoritmi intelligenti hanno l'abilità di analizzare le abitudini degli utenti *on-line* potendo dunque inviare sofisticate *e-mail* fraudolente aumentando le possibilità di indurre in errore i destinatari riuscendo così ad ottenere i loro dati sensibili per commettere ulteriori illeciti. Il *phishing* consiste infatti nell'utilizzo di *e-mail* e siti *web* fraudolenti per raccogliere informazioni personali, finanziarie e sensibili delle vittime. Le *e-mail* sembrano provenire da organizzazioni legittime chiedendo al destinatario di verificare le informazioni sull'*account* inducendolo in inganno a fornire informazioni sensibili utilizzate e rivendute per commettere reati connessi al furto di identità. A tali *e-mail* o siti *web* possono essere incorporati collegamenti che portano al *download* di *malware* (*trojan* o *key logger*) infettando così il sistema della vittima per carpirgli il maggior numero di informazioni. La differenza tra lo *spear* e lo *spray phishing* è incentrata sul numero dei destinatari e sulla qualità stessa del messaggio inviato. Il primo consiste in un attacco rivolto ad un destinatario specifico individuato *ex ante* e di conseguenza l'*e-mail* risulterà fortemente personalizzata e basata su meccanismi psicologici calibrati sulla specifica vittima. Mentre il secondo consiste nell'invio dell'*e-mail* a più soggetti differenti e non predeterminati e dunque il contenuto del messaggio risulterà standardizzato.

I *software* basati sull'intelligenza artificiale possono essere inoltre programmati per accedere abusivamente a sistemi informatici, per realizzare i (*distributed) denial of service attacks* (*DoS* o *DDoS*) e per

[753] I. SALVADORI, op. cit., Paragrafo 2, "Applicazioni dell'AI e possibile impatto sul diritto penale".
[754] R. WENIN E G. FORNASARI, op. cit. I. SALVADORI, IV sessione, (pag. 309 – 439), 2017 Articolo "*Il diritto penale dei software "a duplice uso"*"; Paragrafo 1 (pag. 369 - 376) "Introduzione".

diffondere *malware*. [755] Si tratta di un nuovo tipo di attività criminale in base alla quale viene sfruttato il modo in cui i *computer* collegati in rete comunicano al fine di sopraffare la rete stessa e negare il servizio. Per raggiungere tale obiettivo, il *DoS* tempesta di richieste il sito *web* per far si che esso non sia più in grado di far fronte al numero di interazioni che sta ricevendo da parte degli utenti. Questo genera notevoli problematiche soprattutto per quanto riguarda le infrastrutture pubbliche. Le modalità per realizzarlo sono differenti: disabilitare i *router* della rete o i punti di accesso *wireless*, effettuare il *mail-bombing* in cui si utilizza un *software* specializzato per l'invio di grandi volumi di *e-mail* ad un unico indirizzo per bloccare il funzionamento del *software* e per replicare un *virus* che sovrasti la rete. Per realizzare un *DDoS*, invece, il soggetto agente va ad arruolare (infettandoli) altri *computer* (c.d. *zombie*) per attaccare la rete di destinazione.

Ovviamente, tali operazioni possono essere realizzate manualmente dagli agenti oppure automatizzate grazie all'impego di *software* che utilizzano l'intelligenza artificiale per agire senza il necessario *input* dell'essere umano.

Incriminando le condotte (come procurarsi o detenere un *malware*) prodromiche e preparatorie per la commissione di ulteriori reati più gravi, si tutelando i beni giuridici finali indicati precedentemente. Si rientra dunque in un diritto penale preventivo in cui si va ad impedire la commissione di futuri reati.

Per evitare che la normativa diventi obsoleta in poco tempo, data la rapidità dell'evoluzione informatica e tecnologica, nella *Cybercrime Convention* è stato utilizzato un linguaggio di ampio respiro, utile per facilitare l'ingresso dell'intelligenza artificiale sotto l'alveo della disciplina penale dei *cybercrime*, ma che comporta però rischi di indeterminatezza a causa dell'elasticità dei concetti; ad esempio, è complesso comprendere quali siano i *software* malevoli o meno a causa della loro caratteristica di *dual-use* (lecito e illecito).[756]

I *software* a duplice uso possono essere suddivisi in due categorie: multifunzionali e multiuso. I primi hanno l'abilità di svolgere più funzioni di cui almeno una è illegittima. Per incriminarli bisogna valutare se la loro funzione illecita prevalga sulla lecita. I secondi, invece, contengono una funzionalità dannosa intrinseca che può essere impiegata in ambito illecito ma anche lecito (come i *trojan horse*). [757]

I *software* possono inoltre essere concepiti (per volontà del creatore) o adattati (per volontà dell'utilizzatore) al fine di realizzare un illecito.

Dunque, in base all'articolo 6 della *Cybercrime Convention* chiunque detenga, utilizzi, produca o distribuisca tali tipologie di *software* con l'intenzionalità di commettere l'illecito sarà soggetto a sanzione penale. Questa formulazione esclude dall'incriminazione i *software* concepiti o adattati in buona fede ma di per sé idonei a commettere l'illecito.

Inoltre, al medesimo articolo si fa riferimento ai *software* "principalmente" concepiti o adattati per commettere illeciti. Quindi anche nel caso di *software* multifunzionali, questi devono essere stati concepiti in

[755] I. SALVADORI, op. cit., Paragrafo 2, "Applicazioni dell'AI e possibile impatto sul diritto penale".
[756] R. WENIN E G. FORNASARI, op. cit. I. SALVADORI, IV sessione, (pag. 309 – 439), 2017 Articolo "*Il diritto penale dei software "a duplice uso"*"; Paragrafo 1 (pag. 369 - 376) "Introduzione".
[757] R. WENIN E G. FORNASARI, op. cit. I. SALVADORI, IV sessione, (pag. 309 – 439), 2017 Articolo "*Il diritto penale dei software "a duplice uso"*"; Paragrafo 4 (pag. 380 - 382) "Nozione e caratteristiche dei software "a duplice uso".

primo luogo per la commissione di reati. Tale formulazione è criticabile per quanto concerne l'eccessiva limitazione dell'oggetto materiale dato che un *software* può essere concepito per diversi scopi, ma per essere sottoposti alla disciplina penale devono essere stati concepiti principalmente con tale obiettivo. Ed in tal modo viene esclusa la rilevanza di comportamenti comunque pericolosi.[758]

Nel codice penale vengono previste alcune fattispecie di reati in base ai *software* a duplice uso.

All'articolo 615 *quater* c.p. si punisce la detenzione e la diffusione (ricomprendenti gli atti di procurare, riprodurre, diffondere, comunicare e/o consegnare, fornire indicazioni o istruzioni idonee) abusiva di codici di accesso a sistemi informatici o telematici al fine di procurare a sé o ad altri un ingiusto profitto, ovvero di arrecare ad altri un danno.

Tale articolo, con lo scopo di tutelare la riservatezza delle modalità di sicurezza per l'accesso ai sistemi informatici o telematici, viene inserito tra i delitti contro l'inviolabilità del domicilio sanzionando condotte preparatorie per l'accesso abusivo al sistema informatico punito dall'articolo 615 *ter* c.p.. Le due norme appena citate, secondo un recente orientamento della Corte di Cassazione non possono concorrere tra di loro in quanto sono poste a tutela del medesimo bene giuridico (l'inviolabilità del privato domicilio) che l'articolo 615 *quater* c.p. protegge in misura meno ampia rispetto all'articolo 615 *ter* c.p. (che protegge il domicilio informatico inteso come spazio ideale di esclusiva pertinenza della persona fisica o giuridica). Inoltre, come affermato poco fa, l'articolo 615 *quater* c.p., punendo condotte prodromiche alla realizzazione del delitto di accesso abusivo al sistema informatico *ex* articolo 615 *ter* c.p., ne costituisce un antefatto necessario. Si passa quindi da condotte meno invasive a condotte più invasive che necessariamente presuppongono le prime. Quindi l'articolo 615 *quater* c.p. costituisce l'antecedente non punibile dell'articolo 615 *ter* c.p. e rimane in esso assorbito.[759]

Le due condotte sanzionate dall'articolo 615 *quater* c.p. riguardano l'acquisizione del codice di accesso per il sistema informatico protetto da misure di sicurezza e la messa a disposizione (di tale codice) a terzi.

Solitamente per realizzarle vengono impiegati *software* specializzati con l'abilità di captare tali informazioni: gli algoritmi intelligenti si prestano molto bene per tali scopi, grazie alla loro capacità di calcolare le possibili *password* di accesso, di analizzare i dati dei sistemi selezionando le informazioni utili e la loro abilità di superare le barriere poste a sicurezza del sistema informatico riuscendo a scovare le lacune nei modelli di *cyber*-sicurezza aggirando l'ostacolo e immettendosi abusivamente nel sistema.

Con l'articolo 615 *quinquies* c.p., modificato a seguito della legge di ratifica ed esecuzione della *Cybercrime Convention* (legge n. 48/2008),[760] si punisce invece il delitto di diffusione di apparecchiature, dispositivi o programmi informatici diretti a danneggiare o interrompere (anche temporaneamente) un sistema

[758] R. WENIN E G. FORNASARI, op. cit. I. SALVADORI, IV sessione, (pag. 309 – 439), 2017 Articolo "*Il diritto penale dei software «a duplice uso»*"; Paragrafo 5 (pag. 382 - 396) "Gli obblighi di incriminazione dei software "a duplice uso" nelle fonti sovranazionali".
[759] Sentenza della Cassazione Penale, Sez. II, n.21987/2019 del 20 maggio 2019.
[760] Legge 18 marzo 2008, n. 48, "*Ratifica ed esecuzione della Convenzione del Consiglio d'Europa sulla criminalità informatica, fatta a Budapest il 23 novembre 2001, e norme di adeguamento dell'ordinamento interno*", pubblicata nella *Gazzetta Ufficiale* n. 80 del 4 aprile 2008 - Supplemento ordinario n. 79: https://www.parlamento.it/parlam/leggi/080481.htm

informatico. Le condotte tipiche (ossia l'acquisizione, la produzione, la riproduzione, l'importazione, la diffusione, la comunicazione, la consegna o la messa a disposizione) devono avere ad oggetto apparecchiature, dispositivi o programmi informatici.

Sia l'articolo 615 *quater* c.p. che per l'articolo 615 *quinques* c.p., consistono in reati comuni di pericolo, quindi, per integrarli non è richiesta la realizzazione del danno in quanto risulta sufficiente la diffusione delle chiavi di accesso o la realizzazione di un sistema idoneo ad arrecare un rischio di danneggiamento. Come elemento soggettivo, è necessario il dolo specifico consistente nella volontà e consapevolezza di danneggiare illecitamente un sistema informatico o telematico. Con tali norme, il legislatore ha l'obiettivo di tutelare la riservatezza informatica e la fruizione del sistema informatico da parte del gestore. La *ratio* è quindi l'estensione del concetto di luoghi di dimora (come proiezione spaziale della persona) andando a ricomprendere anche la sicurezza dei propri sistemi informatici. Si tratta di reati che vanno a punire il pericolo di commissione della condotta di cui all'articolo 615 *ter* c.p.

Tale articolo punisce, infatti, l'accesso abusivo (ossia un'intromissione estranea alla volontà del gestore o possessore) ad un sistema informatico o telematico protetto da misure di sicurezza. Anch'esso tutela la riservatezza informatica sanzionando due condotte: l'intromissione non autorizzata e il mantenimento di tale accesso al sistema informatico contro la volontà di chi ha in diritto di escluderlo.

Con l'art. 617 *quinquies* c.p. si punisce (fuori dai casi consentiti dalla legge) chiunque installi apparecchiature (compresi i programmi informatici che permettono al *reo* di accedere da remoto e intercettare le comunicazioni in entrata o in uscita) atte a intercettare, impedire, o interrompere comunicazioni relative ad un sistema informatico. Di conseguenza, la liceità o meno di tali condotte dipende da specifiche norme di legge che le permettono.[761] La *ratio* di tale disposizione del codice penale è la tutela della riservatezza delle comunicazioni telematiche. Si tratta di un reato che va a punire i fatti prodromici alla commissione del reato di cui all'articolo 617 *quater* c.p. ossia l'intercettazione, l'impedimento, l'interruzione o la rilevazione illecita (con un connotato di fraudolenza) di comunicazioni informatiche o telematiche.

Nell'incriminare tali tipologie di azioni, il legislatore fa ricorso alla struttura dei delitti di ostacolo (punendo condotte basate sul mero possesso di un programma a duplice uso) e dei reati preparatori (punendo condotte utili a facilitare la realizzazione di reati da parte del *reo* o di terzi) generando quindi una categoria di reati c.d. di prevenzione. [762]

Nel nostro sistema giuridico, come accennato precedentemente, non sussistono norme che trattino i sistemi di intelligenza artificiale come mezzi di commissione del reato. Anche se, ad esempio nel reato di *stalking* (art. 612 *bis* c.p.) e nel reato impropriamente definito come *revenge porn* (art. 612 *ter* c.p.) si fa

[761] R. WENIN E G. FORNASARI, op. cit. I. SALVADORI, IV sessione, (pag. 309 – 439), 2017 Articolo "*Il diritto penale dei software "a duplice uso"*"; Paragrafo 7 (pag. 404 - 414) "L'incriminazione dei software "a duplice uso" nel diritto penale italiano".
[762] R. WENIN E G. FORNASARI, op. cit. I. SALVADORI, IV sessione, (pag. 309 – 439), 2017 Articolo "*Il diritto penale dei software "a duplice uso"*"; Paragrafo 9 (pag. 423 - 443) "Sui presupposti per una legittima incriminazione dei software «a duplice uso»".

espresso riferimento all'utilizzo di strumenti informatici o telematici nella commissione del reato per andare ad applicare la pena più grave.

Infatti uno *stalker* potrebbe utilizzare una *chatbot* intelligente allo scopo di generare nei confronti della vittima un perdurante stato di ansia o di paura o un fondato timore per la sua incolumità. Ed inoltre, grazie all'intelligenza artificiale risulta estremamente facile realizzare contenuti offensivi o illeciti manipolando immagini, sostituendo volti o voci in *file* video realizzando contenuti *fake* di natura pornografica o pedopornografica. [763]

Di conseguenza è interessante analizzare l'incidenza dell'intelligenza artificiale in tali tipologie di reati informatici "in senso ampio".

### *3.1 Le chatbot e i reati di odio e discriminazione*

Grazie all'esperimento di Turing (*the imitation game*) è stato dimostrato come l'intelligenza artificiale, a seguito della notevole evoluzione tecnologica avvenuta negli ultimi anni, ha la grande abilità di riuscire a simulare il linguaggio umano rendendo quasi impossibile la distinzione tra un interlocutore persona fisica o artificiale. Abbiamo già trattato, nel primo capitolo, delle varie sedute "psichiatriche" avvenute tra la prima *chatbot* "*Eliza*" (denominata "*The Doctor*") realizzata da Joseph Weizenbaum nel 1964 e "*Parry*", *chatbot* ideata dallo psichiatra Kenneth Colby nel 1972 che imitava un soggetto affetto da malattie mentali.

Le capacità di tali sistemi intelligenti vengono oggigiorno sfruttate da gran parte della collettività come supporto quotidiano (ad esempio *Alexa*, *Siri* e *Cortana*) e dalle società per garantire una facile interlocuzione con il pubblico e per pubblicizzare i loro prodotti (le *chatbot* del supporto clienti e le *chatbot* presenti sui *social network*).[764]

Nonostante i notevoli effetti positivi dovuti all'utilizzo di apparecchiature intelligenti, i rischi rimangono dietro l'angolo. Un esempio recentissimo che dimostra la pericolosità degli assistenti vocali è dato da *Alexa* che ha proposto una rischiosissima *challenge*, chiamata "*penny challenge*", in voga sul *social network* denominato *Tik Tok* a una bambina di dieci anni spingendola a inserire parte degli spinotti di un caricabatterie nella presa di corrente, lasciandone per metà fuori, al fine di poter appoggiare una moneta tra i due poli esposti. Il *software* ha imparato dagli utenti di *Tik Tok* e ha deciso di riproporlo alla bambina (rientrante nel *target* della *challenge*). A seguito dell'incidente, *Amazon* ha aggiornato il sistema e ha assicurato che *Alexa* non proporrà più sfide pericolose.[765]

[763] I. SALVADORI, op. cit., Paragrafo 7, "L'agente artificiale come mezzo di esecuzione di un reato".
[764] S. QUINTARELLI, op. cit., Paragrafo 1.3 (pag. 21 - 24) "Chatbot. Siri, Alexa, Cortana e i loro cugini".
[765] ANSA Redazione, "*Assistenti poco smart, Alexa propone sfida rischiosa a bimba*", Milano, 30/12/2021.

Il *machine learning*, applicato in contesti sociali reali, rischia di generare notevoli problematiche: impara dalla società automatizzando i suoi comportamenti (sia positivi che negativi).

Ancora, si può ricordare di "*Tay* (*Thinking About You*)", una *chatbot* realizzata da *Microsoft* per sperimentare e condurre ricerche sulla comprensione conversazionale, resa accessibile via *Twitter* il 23 marzo 2016. Il sistema era stato progettato al fine di simulare le conversazioni di un'adolescente americana nelle interazioni con gli altri utenti della piattaforma. Come tutti i meccanismi di *machine learning*, anche *Tay* era programmata per imparare dagli stimoli esterni: in tal caso, i messaggi ricevuti dagli utenti di *Twitter*.[766] Il problema fu che *Tay* fu inondato di messaggi inneggianti l'odio raziale, il nazismo e l'utilizzo di droghe pesanti, di conseguenza, la *chatbot* imparò tale linguaggio e iniziò a *tweettare* nello stesso modo generando nell'arco di sedici ore dall'immissione *on-line* oltre 96000 *tweet* osceni rispondendo alle provocazioni di altri utenti:[767] *Microsoft* fu costretta a sospendere il servizio.

Lo stupore dei programmatori e della società produttrice fu connesso al fatto che *Tay*, in laboratorio, funzionava perfettamente, ma, una volta immesso nella rete, gli *input* esterni, hanno portato la *chatbot* a commettere reati di odio e di discriminazione.[768]

Com'è possibile osservare, le *chatbot* contribuiscono anch'esse alla proliferazione dell'*hate speech* e delle *fake news* sui *social network* imparando dagli utenti della rete in modo autonomo grazie al *machine learning*. Gli *input* appresi dall'esterno, modificano il comportamento dell'intelligenza artificiale che si adatta e si evolve in modo oscuro anche agli stessi programmatori. Inoltre, l'elevato numero di operatori che agiscono per la realizzazione del sistema intelligente rende, come negli altri casi citati precedentemente, difficile l'individuazione dei responsabili[769] (nel caso di *Tay* risulterebbero coinvolti gli sviluppatori, i *designer*, l'azienda e coloro che interagivano con la *chatbot*).[770]

Le *chatbot* sono l'emblema dell'applicazione del *natural language processing*, ossia la capacità delle macchine di elaborare le parole realizzando frasi e discorsi simili ad un essere umano. L'intelligenza artificiale immagazzina tutte le conversazioni che avvengono sui *social network*, gli articoli pubblicati sui giornali *on-line* e sui *blog*: il *web* è la sua fonte primaria di conoscenza ed evoluzione, in questo modo impara a capire le nostre modalità di comunicazione per riuscire a formulare frasi e discorsi di senso compiuto. Di conseguenza, se il linguaggio usato dalla società *on-line* è razzista, misogino, negazionista e nazista, lo sarà anche il linguaggio dell'intelligenza artificiale.[771]

La manifestazione odiosa del pensiero sui *social network* comporta la nascita di un dibattito in relazione alla possibilità di utilizzare il diritto penale come limitazione della liberà di espressione quando la stessa sfoci

---

[766] S. QUINTARELLI, op. cit., Paragrafo 1.3 (pag. 21 - 24) "Chatbot. Siri, Alexa, Cortana e i loro cugini".
[767] E. HUNT, "*Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter*", The Guardian, 24/03/2016: https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter
[768] S. QUINTARELLI, op. cit., Paragrafo 1.3 (pag. 21 - 24) "Chatbot. Siri, Alexa, Cortana e i loro cugini".
[769] I. SALVADORI, op. cit., Paragrafo 2, "Applicazioni dell'AI e possibile impatto sul diritto penale".
[770] 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*".
[771] E. CAPONE, "*Le intelligenze artificiali fra razzismo e questione etica*", 06/04/2021, IT Italian.Tech, Parte del gruppo GEDI e la Repubblica: https://www.italian.tech/2021/04/06/news/le-intelligenze-artificiali-fra-razzismo-e-questione-etica-299491573/

in condotte discriminatorie e di incitamento all'odio. L'intensificarsi di questo fenomeno è dovuto alla crisi economica che stiamo vivendo che ha portato ad un lento depauperamento del bagaglio culturale degli utenti *on-line* e una disinformazione verso le Costituzioni democratiche nate a seguito delle guerre mondiali. Ed inoltre, l'aumento della diffusione di tali messaggi discriminatori è dovuta all'evoluzione tecnologica in campo di comunicazione che permette oggi, anche grazie alle *chatbot* e all'intelligenza artificiale, di rendere "virali" tali contenuti. La tecnologia moderna ha il vantaggio di poter diffondere (quasi) ovunque il sapere e le informazioni, ma allo stesso tempo rende estremamene semplice l'accesso a contenuti discriminatori e alle *fake news* che spingono gli utenti più deboli a far loro il pensiero odioso e a diffonderlo ulteriormente. Inoltre, se tale contenuto viene a contatto con una *chatbot* intelligente essa imparerà come un bambino, a replicare il messaggio, automatizzando la diffusione odiosa del pensiero con effetti fortemente impattati sulla popolazione globale: i problemi sono dunque la scala e la rapidità di auto-alimentazione del pensiero.

Tramite il *web* è quindi possibile commettere numerosi illeciti connessi alla manifestazione del pensiero come ad esempio l'apologia con finalità di terrorismo, commessa dai c.d. *foreign fighters* nell'ambito del *cyber*-terrorismo, punita dall'articolo 414 comma 3 c.p. che sanziona l'apologia di reato con l'aggravante speciale ad effetto speciale in caso di utilizzo del mezzo informatico. L'idea di tali criminali è quella di innescare processi di radicalizzazione e di addestramento per finalità terroristiche sul *web* tramite la messa in circolazione di idee radicali, di video violenti che incitano all'odio. Questo scopo può essere raggiunto ed enfatizzato attraverso l'utilizzo di *chatbot* che agiscono senza la necessità di un essere umano connesso.[772]

L'aggravante della commissione tramite l'utilizzo di strumenti informatici del reato di apologia o di istigazione a commettere reati, in realtà, non la troviamo solo in relazione ai reati di terrorismo o crimini contro l'umanità. Essa, è infatti riscontrabile (ad effetto comune) al medesimo articolo 414 c.p. per l'istigazione o l'apologia (considerabile come un'istigazione indiretta) a commettere qualsiasi delitto o contravvenzione e nell'articolo 302 c.p. nel caso di commissione dei delitti non colposi previsti ai Capi I e II del Titolo I del Libro II del codice penale. Ed è evidente che l'intelligenza artificiale rientri a pieno nel concetto di mezzo informatico.

Si tratta di norme poste a tutela dell'ordine pubblico, essendo condotte che generano allarme sociale e rischi per la collettività e dato l'elevato pericolo di tali azioni (commesse pubblicamente con coscienza e volontà), questi reati sono posti in deroga all'articolo 115 c.p. secondo cui non è punibile l'istigazione se non è seguita dalla commissione del reato. Secondo la giurisprudenza maggioritaria ci troviamo difronte ad un reato di pericolo concreto di cui bisogna valutare l'idoneità della condotta a provocare reati. Sono reati comuni che possono essere commessi da chiunque.

[772] P.M. SABELLA, *"La manifestazione odiosa del pensiero in Internet. Responsabilità individuali e dell'Internet Service Provider*", Lezione del 23 marzo 2020, Luiss Guido Carli, Dipartimento di Giurisprudenza, Cattedra di Diritto penale 2, Titolari di insegnamento – Prof.ri E. GALLUCCI E M.N. MASULLO, Paragrafo 1 (pag. 1 – 2) "La manifestazione odiosa del pensiero in Internet. Il ruolo del diritto penale fra tutela dell'individuo e libertà di manifestazione del pensiero ai tempi dei social network. Una premessa".

Il concetto di "odio" è complesso da inserire all'interno di una norma a causa del suo connotato etico-emozionale e rischia di risultare indeterminato nel diritto penale.

L'*hate speech*, nonostante non possieda una descrizione universalmente condivisa, può essere definito come quel discorso finalizzato a promuovere l'odio e il disprezzo nei confronti di individui o gruppi a causa della loro connotazione razziale, etnica, religiosa, culturale o di genere.[773] Sono dunque espressioni pubbliche discriminatorie e razziste verso soggetti o gruppi specifici con l'effetto di alimentare o consolidare pregiudizi e ostilità al fine di emarginare il "diverso" fino a giungere ad una sua totale disumanizzazione.[774] L'azione può avere ad oggetto una o più persone con una motivazione basata sul pregiudizio dovuto ad una discriminazione della vittima da parte del *reo*. [775] L'azione del diritto penale volta a tutelare l'uguaglianza tra gli individui deve essere bilanciata con il diritto fondamentale alla libertà di espressione (che però non può mai spingersi fino alla lesione dei diritti fondamentali altrui).[776]

Per trasformare un reato comune in un reato d'odio, il *reo* deve selezionare la vittima per ragioni discriminatorie (anche e principalmente razziali) verso il gruppo a cui appartiene la persona offesa.[777] Il concetto di razzismo può essere ripreso dalla Sentenza 196583/1993 della Corte di Cassazione in cui si afferma che tale nozione "*indica le dottrine che postulano quale presupposto del divenire storico l'esistenza di razze superiori ed inferiori, le prime destinate al comando, le seconde alla sottomissione"*.[778] La Corte EDU invece, riprendendo l'articolo 14 della CEDU, descrive il comportamento discriminatorio come il trattare in modo differente, salvo giustificazione oggettiva e ragionevole, persone che si trovano in situazioni comparabili: si tratta di una definizione di ampio respiro che permette di essere applicata in notevoli situazioni.[779]

I reati di odio sono stati introdotti dalla legge n. 654 del 13 ottobre 1975[780] in relazione alla criminalizzazione dei reati di diffusione di idee razziste, di incitamento alla discriminazione, alla violenza razzista e di associazione finalizzata ad incitare all'odio o alla discriminazione. Tale legge fu successivamente modificata dalla legge n. 205 del 25 giugno 1993 (c.d. legge Mancino)[781] introducendo una circostanza

---

[773] G. PINO, "*Discorso razzista e libertà di manifestazione del pensiero*", *Pol. dir.*, 11, 2008 (pag. 287 ss.).
[774] V. NARDI, "*I discorsi d'odio nell'era digitale: quale ruolo per l'internet service provider?*" (pag. 268 – 288) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018. In Diritto penale Contemporaneo, Rivista trimestrale 2/2019; Paragrafo 1 (pag. 269 – 272) "Premessa: fenomenologia dei discorsi d'odio 2.0".
[775] P.M. SABELLA, op. cit., Paragrafo 2 (pag. 2 – 4) "Il discorso e il delitto d'odio nel diritto internazionale, sovranazionale e nella giurisprudenza. Una definizione".
[776] V. NARDI, op. cit., Paragrafo 1 (pag. 269 – 272) "Premessa: fenomenologia dei discorsi d'odio 2.0".
[777] P.M. SABELLA, op. cit., Paragrafo 2 (pag. 2 – 4) "Il discorso e il delitto d'odio nel diritto internazionale, sovranazionale e nella giurisprudenza. Una definizione".
[778] Sentenza della Cassazione Penale, Sez. I, n. 196583 del 30 settembre 1993.
[779] P.M. SABELLA, op. cit., Paragrafo 2 (pag. 2 – 4) "Il discorso e il delitto d'odio nel diritto internazionale, sovranazionale e nella giurisprudenza. Una definizione".
[780] Legge 13 ottobre 1975, n. 654, "*Ratifica ed esecuzione della convenzione internazionale sull'eliminazione di tutte le forme di discriminazione razziale, aperta alla firma a New York il 7 marzo 1966*": https://www.gazzettaufficiale.it/eli/id/1975/12/23/075U0654/sg
[781] Legge 25 giugno 1993, n. 205 "*Conversione in legge, con modificazioni, del decreto-legge 26 aprile 1993, n. 122, recante misure urgenti in materia di discriminazione razziale, etnica e religiosa*", entrata in vigore il 27 giugno 1993: https://www.gazzettaufficiale.it/eli/id/1993/06/26/093G0275/sg

aggravante legata alla finalità di discriminazione e di odio. Una successiva modifica ci fu con la legge n. 85 del 24 febbraio del 2006[782] che ha sostituito i termini "diffondere" e "incitare" con "propagandare" e "istigare".

L'articolo 604 *bis* c.p. inserito successivamente a tali riforme, nel 2018, va a punire la propaganda e l'istigazione a delinquere per motivi di discriminazione raziale, etnica e religiosa. L'articolo è collocato all'interno del Titolo XII del Libro II del codice penale, al Capo III aggiungendo la nuova Sezione I *bis* "Dei delitti contro l'eguaglianza".

Si tratta di reati comuni (che possono essere commessi da chiunque).

L'articolo contiene in sé otto fattispecie: la propaganda di idee fondate sulla superiorità o sull'odio raziale o etnico; l'istigazione a commettere atti di discriminazione per motivi razziali, etnici, nazionali o religiosi; la commissione di atti di discriminazione per motivi razziali, etnici, nazionali o religiosi; l'istigazione a commettere violenza per motivi razziali, etnici, nazionali o religiosi; la commissione di atti di violenza o di provocazione alla violenza per motivi razziali, etnici, nazionali o religiosi; la creazione di organizzazioni, associazioni, movimenti o gruppi il cui scopo è l'incitamento alla discriminazione o alla violenza per i motivi sopra citati; la partecipazione a suddette organizzazioni; ed infine, alla propaganda o istigazione e incitamento fondati sulla negazione o apologia della *Shoah*, dei crimini di genocidio, dei crimini contro l'umanità e dei crimini di guerra. [783]

Per propaganda si intende un'azione il cui risultato è volto ad influire sul comportamento altrui tramite una diffusione idonea a raccogliere consensi in base all'idea divulgata.[784] E l'istigazione necessita di un'attività diretta a convincere terzi a tenere quella determinata condotta rafforzandone il proposito criminoso o facendo insorgere un proposito prima inesistente. Si tratta di reati plurioffensivi che vanno a colpire l'ordine pubblico e la dignità umana.[785] Essi sono considerati come reati di pura condotta e di pericolo astratto in quanto non interessa che l'azione abbia prodotto effetti nell'immediatezza del fatto per far si che il reato si consideri perfezionato.[786] Per le condotte sancite al comma 1 lett. a) dell'articolo 604 *bis* c.p. è sufficiente la sussistenza del dolo generico di realizzarle con coscienza e volontà, mentre per le fattispecie espresse al comma 1 lett. b), del medesimo articolo, è necessario il dolo specifico. [787]

La terminata digressione teorica è utile per inserire in un contesto normativo le problematiche connesse alla realizzazione di tali attività illecite attraverso l'utilizzo di tecnologie informatiche e a maggior ragione l'influenza dell'intelligenza artificiale su tali reati.

Navigando nella rete, è facilissimo imbattersi in commenti (soprattutto sui *social network*) che incitano l'odio e la discriminazione. La rapidità della connessione ha accentuato un fenomeno non di per sé legato allo sviluppo tecnologico. La velocità istantanea di diffusione dei messaggi che possono giungere immediatamente

---

[782] Legge 24 febbraio 2006, n. 85, "*Modifiche al codice penale in materia di reati di opinione*", pubblicata nella *Gazzetta Ufficiale* n. 60 del 13 marzo 2006: https://web.camera.it/parlam/leggi/060851.htm

[783] P.M. SABELLA, op. cit., Paragrafo 2.1 (pag. 5 – 6) "E nel diritto interno. L'art. 604-bis c.p.".

[784] Sentenza della Corte Costituzionale n. 87 del 22 giugno 1966.

[785] Sentenza della Cassazione Penale, Sez. III, n. 36906 del 14 settembre 2015.

[786] Sentenza della Cassazione Penale, Sez. III, n. 37581 del 7 maggio 2008.

[787] Sentenze della Cassazione Penale, Sez. III, n. 36906/2015 e 37581/2008.

ai destinatari avendo un impatto notevole su un numero quasi infinito di utenti, la difficoltà nell'eliminare definitivamente quel messaggio grazie alla ricondivisione da parte di altri utenti, la possibilità di raggiungere quasi l'interezza del globo connesso e la difficoltà per gli Stati di cooperare tra loro in un ambito transnazionale, si presta ad essere un terreno fertile per tali tipologie di illeciti;[788] inoltre, con l'avvento dell'intelligenza artificiale, la ricondivisione può essere addirittura automatizzata e di conseguenza avere effetti ancora più impattanti sulla società.

A differenza delle due situazioni di rilevanza penale, in cui può essere coinvolta l'intelligenza artificiale, citate nei precedenti paragrafi (veicoli autonomi e *robot* chirurghi) in cui la responsabilità finiva per ricadere principalmente su due categorie di figure (il creatore/produttore e l'utilizzatore), nel caso dei reati d'odio commessi all'interno del *web*, la figura dell'*Internet Service Provider* (prestatore di servizi o gestore di piattaforme *on-line*) assume un'ulteriore complessa rilevanza.

Infatti, già prima della forte diffusione dell'intelligenza artificiale nella quotidianità, tale figura riscontrava in ambito penale un difficile inquadramento in relazione alla sua responsabilità per i reati di odio commessi dagli utenti.

L'*Internet Service Provider* può essere chiamato a rispondere a titolo di concorso con il singolo autore del *post* odioso avendo consentito (o comunque non impedito) la condivisione del messaggio? In proposito si riscontrano in dottrina tre diverse posizioni.

La prima pone l'*Internet Service Provider* sullo stesso piano degli altri utenti, considerandolo privo di un dovere-potere di controllo sulle condotte altrui. Di conseguenza, la sua responsabilità penale, è limitata alle ipotesi di concorso commissivo doloso nella condotta altrui. [789] La condotta commissiva dovrà consistere nell'immissione o diffusione dolosa di dati in quanto non può considerarsi sufficiente la mera fornitura degli accessi alla rete.[790]

La seconda, invece, impone all'*Internet Service Provider* di svolgere attività di controllo e censura *ex ante* sul materiale caricato dagli utenti. In tal modo, si tutelano maggiormente i soggetti terzi configurando una responsabilità per reato omissivo improprio *ex* art. 40 cpv. c.p. per non aver impedito l'evento,[791] oppure in combinato disposto con l'articolo 110 c.p. a titolo di concorso omissivo nel reato realizzato dall'utente. In

[788] V. NARDI, op. cit., Paragrafo 1 (pag. 269 – 272) "Premessa: fenomenologia dei discorsi d'odio 2.0".
[789] V. NARDI, op. cit., Paragrafo 3 (pag. 274 – 278) "I paradigmi di responsabilizzazione dell'*Internet Service Provider* nel formante legislativo, dottrinale e giurisprudenziale".
[790] B. PANATTONI, "*Gli effetti dell'automazione sui modelli di responsabilità: il caso delle piattaforme online*" (pag. 33 – 58) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018. In Diritto penale Contemporaneo, Rivista trimestrale 2/2019; Paragrafo 3 (pag. 39 – 41) "La responsabilità penale degli ISP in caso di caricamento e/o diffusione di contenuti illeciti in rete".
[791] V. NARDI, op. cit., Paragrafo 3 (pag. 274 – 278) "I paradigmi di responsabilizzazione dell'*Internet Service Provider* nel formante legislativo, dottrinale e giurisprudenziale".

tal caso si farà riferimento ad un mancato controllo e censura preventiva del contenuto illecito: una responsabilità *ex ante* operante prima che i dati vengano resi disponibili in rete. [792]

Una posizione intermedia è infine posta dal terzo modello di responsabilità secondo cui l'*Internet Service Provider* ha l'obbligo di svolgere un controllo *ex post* su quanto condiviso dagli utenti. Esso ha quindi due doveri: l'obbligo di denuncia degli illeciti (di cui viene a conoscenza) in base al dovere di collaborazione con le autorità e l'obbligo di eliminare il materiale illecito. In tale ultima ipotesi, a differenza della seconda in cui si pone come controllore, l'*Internet Service Provider* ha un ruolo di tutore secondo un modello di responsabilità da reato omissivo proprio.[793] In tal caso si farà riferimento ad una mancata rimozione del contenuto illecito: una responsabilità *ex post* connessa al perdurare della disponibilità del contenuto illecito.[794]

In realtà, ad oggi, non esiste alcuna norma che impone all'*Internet Service Provider* di agire per evitare o per rimuovere un contenuto illecito, di conseguenza, non è possibile configurare una responsabilità omissiva (che sia essa propria o impropria). Infatti l'articolo 17 del D.lgs. n.70 del 2003[795] esclude l'esistenza di un obbligo generale di sorveglianza, gravante sull'*Internet Service Providers*, sui contenuti caricati dagli utenti.[796] Anche se la giurisprudenza[797], in virtù dell'obbligo di rimozione del materiale illecito (*ex* art. 16 del D.lgs. n. 70 del 2003), tende a riconoscere la possibilità di configurare in capo all'*Internet Service Provider* una forma di responsabilità a titolo di concorso omissivo con il *reo* che ha diffuso il messaggio di odio nel caso in cui l'*Internet Service Provider* fosse venuto a conoscenza del *post* senza adottare condotte necessarie per interrompere la diffusione del *hate speech*. [798]

Il problema di configurare una responsabilità penale *ex post* in capo all'*Internet Service Provider* per omessa rimozione di un contenuto illecito concerne la valutazione della sussistenza o meno della posizione di garanzia in capo a tali soggetti. Basandosi sul modello *notice and takedown,* che impone un dovere di rimozione immediata dei contenuti illeciti connesso ad un dovere di informazione dell'Autorità competente a seguito della presa coscienza di un contenuto illecito sulla piattaforma, tale sistema risulta applicabile, secondo il D.lgs. n. 70 del 2003, esclusivamente nell'ambito del diritto amministrativo e non permette di configurare penalmente una posizione di garanzia in capo all'*Internet Service Provider*.[799] Di conseguenza, in virtù della posizione di signoria detenuta dallo stesso sulla sua piattaforma e della rilevanza sociale di tali soggetti, dovrebbe essere ipotizzata la realizzazione di una normativa penale idonea ad imporre tali obblighi di controllo

[792] B. PANATTONI, "*Gli effetti dell'automazione sui modelli di responsabilità: il caso delle piattaforme online*", op. cit., Paragrafo 3 (pag. 39 – 41) "La responsabilità penale degli ISP in caso di caricamento e/o diffusione di contenuti illeciti in rete".
[793] V. NARDI, op. cit., Paragrafo 3 (pag. 274 – 278) "I paradigmi di responsabilizzazione dell'*Internet Service Provider* nel formante legislativo, dottrinale e giurisprudenziale".
[794] B. PANATTONI, "*Gli effetti dell'automazione sui modelli di responsabilità: il caso delle piattaforme online*", op. cit., Paragrafo 3 (pag. 39 – 41) "La responsabilità penale degli ISP in caso di caricamento e/o diffusione di contenuti illeciti in rete".
[795] Decreto Legislativo 9 aprile 2003, n. 70, "Attuazione della direttiva 2000/31/CE relativa a taluni aspetti giuridici dei servizi della società dell'informazione nel mercato interno, con particolare riferimento al commercio elettronico", GU n.87 del 14 aprile 2003: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2003;70
[796] P.M. SABELLA, op. cit., Paragrafo 3 (pag. 6 – 9) "Quale responsabilità per gli Internet Service Providers ed *hate speech*".
[797] Sentenza della Cassazione Penale, Sez. V, n. 54946, del 27 dicembre 2016.
[798] V. NARDI, op. cit., Paragrafo 3.1 (pag. 278 – 280) "(segue) Gli obblighi di rimozione successivi alla commissione del reato: quale modello sanzionatorio per l'ISP?".
[799] P.M. SABELLA, op. cit., Paragrafo 3 (pag. 6 – 9) "Quale responsabilità per gli Internet Service Providers ed *hate speech*".

e rimozione al fine di tutelare gli utenti del servizio (impossibilitati a proteggere autonomamente il loro bene giuridico leso) da azioni illecite *on-line*. Inoltre, l'attribuzione della posizione di garanzia può trovare riscontro negli obblighi normativi (*extra* penali) sanciti dall'art. 16 comma 1 D.lgs. n. 70/2003; anche se, in base a tale articolo l'elemento soggettivo deve consistere nell'effettiva conoscenza del contenuto dell'illecito riscontrabile in un dolo diretto di partecipazione. È da evidenziarsi, dunque, l'opportunità di un adeguamento della disciplina[800] che però genera al contempo due perplessità che necessitano di essere superate: il rischio di azioni eccessivamente censuranti da parte dell'*Internet Service Provider* al fine di evitare di incorrere in una sanzione penale, andando però a limitare eccessivamente la liberà di espressione degli utenti della piattaforma, e la rilevanza penale di tale condotta omissiva. In questa seconda problematica, in modo da essere considerata non come un mero *post factum* non punibile, l'azione omissiva, intervenuta successivamente alla condotta dell'utente, deve aver avuto un valore causale alla realizzazione del reato preso in esame. In concreto, infatti, la fase successiva al caricamento di un contenuto odioso *on-line* ha una notevole portata offensiva in conseguenza della diffusione e dalla possibilità di ricondivisione dello stesso da parte di altri utenti generando un effetto moltiplicatore del messaggio e protraendo dunque l'offesa al bene giuridico protetto. Di conseguenza, l'inazione dell'*Internet Service Provider* potrebbe essere considerata non come mero *post factum* ma come azione a contenuto offensivo.[801]

All'interno dei *social network*, l'utente è responsabile delle proprie azioni a prescindere dal numero di condivisioni. La diffamazione a mezzo *social* è quindi paragonata a quella su un giornale in quanto ha la capacità di raggiungere un numero indeterminato di persone, ma l'*Internet Service Provider* non è paragonato al direttore del giornale in tema di responsabilità (*ex* art. 57 c.p.) per le affermazioni degli utenti. Le piattaforme sono considerate come dei contenitori e gli *Internet Service Providers* non possono effettuare un controllo costante sugli utenti (per evitare di ricadere nella censura), di conseguenza, come affermato precedentemente, risponderanno penalmente solo se non rimuovono dolosamente il contenuto (a titolo di condotta attiva). [802]

Risultano invariate, nel caso di realizzazione del reato da parte di un utente persona fisica o di una *chatbot* intelligente, le problematiche relative alla responsabilità dell'*Internet Service Provider* per la diffusione di messaggi d'odio sul *web*, in quanto, la configurabilità della responsabilità concerne una rimozione *ex post* del contenuto illecito. Non interessa quindi che il sistema intelligente abbia appreso da un utente senza essere consapevole dell'illiceità del messaggio trasmesso dato che le possibili responsabilizzazioni dell'*Internet Service Provider* riguardano una rimozione successiva del contenuto.

Non è possibile responsabilizzare la *chatbot* intelligente, che, come sappiamo, impara dalle azioni degli utenti tramite il meccanismo del *machine learning*, è però fondamentale che, come già affermato, il produttore

---

[800] B. PANATTONI, "*Gli effetti dell'automazione sui modelli di responsabilità: il caso delle piattaforme online*", op. cit., Paragrafo 4 (pag. 44 – 47) "La configurabilità di una responsabilità penale *ex post* in capo agli ISP".
[801] B. PANATTONI, "*Gli effetti dell'automazione sui modelli di responsabilità: il caso delle piattaforme online*", op. cit., Paragrafo 4.2 (pag. 49 – 51) "Le problematiche poste dalla previsione di una responsabilità *ex post*".
[802] P.M. SABELLA, op. cit., Paragrafo 8 (pag. 5 – 9) "La diffamazione a mezzo stampa o con qualsiasi altro mezzo di pubblicità – Il ruolo dei social network e gli internet service providers (ISP): editori o meri contenitori?".

ponga delle limitazioni al sistema. Lo scopo di tali vincoli, imposti a priori, è quello di impedire la commissione di illeciti da parte della *chatbot* così da non far ricadere il produttore in una possibile responsabilità in concorso con l'*Internet Service Provider*.

### *3.2 Le fake news e gli algoritmi intelligenti*

Nell'"era della disinformazione" è importante soffermare la nostra attenzione sull'emblematico e tragico genocidio iniziato in Myanmar nel 2016 e alimentato anche tramite il supporto di *account* falsi, generati e controllati dal governo militare del Myanmar su *Facebook* al fine di diffondere *fake news* per provocare una rivolta contro la popolazione musulmana *Rohingya*.

*Facebook* fu quindi chiamata a rispondere e a chiudere numerose pagine e *accounts*[803] apparentemente gestite da personaggi famosi del paese ma in realtà utilizzate dal governo per diffondere notizie false per innescare reazioni di odio sulla minoranza.[804] Il *social network* si è rilevato un ambiente favorevole per la proliferazione di atti contrari ai diritti umani. In Myanmar, *Facebook* risulta essere l'unica fonte di informazione per gran parte della popolazione, di conseguenza l'immissione di notizie false (come lo stupro di una donna buddista da parte di un musulmano) genera effetti esorbitanti che si concretizzano in atti di violenza nei confronti della minoranza.[805]

*Facebook* ha ammesso che avrebbe potuto fare di più e ha deciso di investire parte del suo ricavato per contrastare i fenomeni di violenza sulla sua piattaforma tramite l'utilizzo di sistemi di intelligenza artificiale sia per scovare gli *account* falsi ed eliminarli che per segnalare attivamente *post* che infrangono gli *standard* della *community*.[806] Al contempo, la società di proprietà di Mark Zuckerberg non si è mai assunta la responsabilità per i delitti d'odio commessi dal governo militare del Myanmar. [807]

L'intelligenza artificiale può infatti essere utilizzata come strumento per supportare le indagini individuando *accounts* falsi o messaggi di odio, ma può anche essere utilizzata come strumento di disinformazione di massa tramite algoritmi in grado di automatizzare le *fake news*.

Le *fake news* generate possono poi essere divulgate esponenzialmente tramite le *chatbot* che automatizzano il processo di diffusione. [808]

---

[803] "*Removing Myanmar Military Officials From Facebook*", in Meta Newsroom, 28 Agosto 2018: https://about.fb.com/news/2018/08/removing-myanmar-officials/

[804] P. MOZUR, "*A Genocide Incited on Facebook, With Posts From Myanmar's Military*", in The New York Times 16 Ottobre 2018: https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html

[805] "*Facebook ha ammesso di avere sbagliato in Myanmar*", in Il Post, 7 novembre 2018: https://www.ilpost.it/2018/11/07/facebook-ha-ammesso-di-avere-sbagliato-in-myanmar/

[806] S. SU, "*Update on Myanmar*", in Meta Newsroom, 15 Agosto 2018: https://about.fb.com/news/2018/08/update-on-myanmar/

[807] A. SIMONCINI, op. cit., Paragrafo 3.1 (pag. 69 – 71) "Il soggetto "catturato" dallo strumento".

[808] A. LONGO E G. SCORZA, op. cit., Paragrafo 3 (pag. 88 - 147) "Vantaggi e rischi non economici dell'AI - I rischi etici, sociali e politici dell'AI - Algoritmi che discriminano".

Per *fake news* si intende qualsiasi tipologia di falsa informazione realizzata tramite l'utilizzo del *web* e che di conseguenza avrà un impatto maggiore sulla società. Gli scopi a cui sono proiettate sono di diverso tipo. In primo luogo, come abbiamo visto tramite il caso del genocidio in Myanmar, le *fake news* possono essere utilizzate per generare campagne di odio e di discriminazione verso minoranze o anche verso singoli gruppi o soggetti. Inoltre, possono essere realizzate al fine di indirizzare gli elettori verso un candidato politico a discapito di altri (come avvenne nel 2016 in America). Ed infine, le *fake news* hanno anche la capacità di generare profitti attraverso la realizzazione di titoli accattivanti e "a caccia di *click*" per poter sfruttare i guadagni ricavati dalle campagne pubblicitarie che vanno ad ospitare.

Le notizie false sono sempre esistite anche prima della nascita delle piattaforme *social*, ma la rapidità di connessione e di circolazione dei contenuti ha ampliato il fenomeno. Inoltre, è forse possibile connettere fenomeno della disinformazione *on-line* anche alla riduzione del livello medio di cultura globale, all'analfabetismo culturale e alla crisi economica. Ed è proprio su tali debolezze della popolazione che si insinuano azioni criminali che sfruttano la rapidità del *web* e l'automazione degli algoritmi per diffondere false informazioni.[809]

Il diritto penale si è già da tempo prodigato verso la limitazione del fenomeno della diffusione di notizie false prevedendo norme incriminatici di tali condotte. Innanzitutto può citarsi l'articolo 656 c.p. che punisce, con la forma di reato di pericolo astratto, la pubblicazione o la diffusione di notizie false, esagerate o tendenziose atte a turbare l'ordine pubblico. Il concetto di diffusione è inteso in senso ampio come divulgazione con qualsiasi mezzo ad una pluralità di persone, ricomprendendo dunque anche il fenomeno delle *fake news* attraverso le piattaforme *on-line*.

Ovviamente tali tipologie di norme rischiano di collidere con l'articolo 21 della Costituzione che tutela la liberà di manifestazione del pensiero. Nel Codice Rocco, infatti, l'articolo 656 c.p. era volto alla criminalizzazione di opinioni contrarie al regime fascista, ma oggi, la Corte Costituzionale ha ritenuto legittimo tale articolo grazie ad un'interpretazione costituzionalmente orientata. La legittimità della norma è garantita dalla determinazione del contenuto delle notizie oggetto di reato: false (difformi al vero), esagerate (contenenti verità amplificate) e tendenziose (realtà presentata in modo ingannevole e alterato). Di conseguenza, vengono escluse dall'alveo della criminalizzazione tutte le interpretazioni personali del vero, facendoci rientrare solamente le notizie che trasformano il vero in falso. Inoltre, l'articolo 656 c.p. non punisce le mere condotte di pubblicazione o diffusione di per sé, ma esclusivamente in base alla loro idoneità a turbare l'ordine pubblico.[810] Trattandosi di una contravvenzione è sufficiente che il *reo* si trovi in colpa nella

[809] A. COSTANTINI, "*Istanze di criminalizzazione delle fake news al confine tra la tutela penale della verità e repressione del dissenso*" (pag. 60 – 80) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018. In Diritto penale Contemporaneo, Rivista trimestrale 2/2019; Paragrafo 1 (pag. 61 – 63) "*Fake news*, ossia la falsità delle notizie all'epoca di internet e della post-verità".

[810] A. COSTANTINI, op. cit., Paragrafo 2 (pag. 63 – 67) "Il diritto penale di fronte alla diffusione di notizie false *online*: i reati astrattamente configurabili".

diffusione della notizia imponendo in capo alla collettività un obbligo di controllo della veridicità dell'informazione diffusa.

Sussiste, però, nella norma la clausola di sussidiarietà "*se non costituisce più grave reato*" che permette alle *fake news* di essere punite in modo più grave in base a precise disposizioni di legge: il procurato allarme *ex* art. 658 c.p. (che punisce chiunque, annunziando disastri, infortuni o pericoli inesistenti, suscita allarme presso l'Autorità, o presso enti o persone che esercitano un pubblico servizio), il delitto di disfattismo politico *ex* art. 265 c.p. (che punisce chiunque diffonda o comunichi voci o notizie false, esagerate o tendenziose, che possano destare pubblico allarme o deprimere lo spirito pubblico o altrimenti menomare la resistenza della nazione di fronte al nemico) la cui rilevanza è circoscritta ai tempi di guerra, l'aggiotaggio informativo *ex* art. 501 c.p. (incrimina le condotte di pubblicazione o divulgazione di notizie false, esagerate o tendenziose che siano idonee a cagionare un aumento o una diminuzione del prezzo delle merci, ovvero dei valori annessi nelle liste di borsa o negoziabili nel pubblico mercato), il delitto di diffamazione *ex* art. 595 c.p. (rispetto all'immissione sul *web* di contenuti falsi che offendono la reputazione di specifici soggetti) che risulterebbe in tal caso aggravato dal comma 3 del medesimo articolo (utilizzo di qualunque altro mezzo di pubblicità diverso dalla stampa) e il delitto di truffa *ex* art. 640 c.p. (in base al quale le *fake news* potrebbero rientrare in un possibile artificio o raggiro).[811]

Com'è facile comprendere, non si tratta di un fenomeno tipicamente commesso dai sistemi di intelligenza artificiale, ma gli stessi possono avere effetti sull'enfatizzazione delle condotte criminose aumentando esponenzialmente i danni.

La potenza computazionale connessa alla digitalizzazione dell'informazione permette agli algoritmi di raccogliere informazioni sugli utenti realizzando un gruppo esponenziale di *big data* individuando così le caratteristiche principali di gruppi di persone a cui veicolare messaggi di *fake news* in base alle loro caratteristiche. In tal modo, la *fake news* può essere inviata in modo mirato attraverso operazioni di *profiling* e non diffusa a campione verso la genericità degli utenti delle piattaforme. Così le *fake news* possono diventare virali ed espandersi. [812]

In relazione alla responsabilità dell'*Internet Service Provider* si può rinviare a quanto detto al sotto-paragrafo precedente: sono pur sempre gli utenti ad essere responsabili delle proprie azioni sui *social network* che possono essere considerati come dei contenitori in cui l'intelligenza artificiale può intervenire per automatizzarne la diffusione o per individuare i contenuti illeciti con lo scopo di eliminarli. L'intelligenza artificiale opera come uno strumento nelle mani del *reo*.

---

[811] *Idem*.

[812] B. G. BRANCATI, "*L'impatto dell'Intelligenza Artificiale (AI-Artificial Intelligence) sul ciclo di intelligence e sugli strumenti a disposizione per i pianificatori militari e le forze dell'ordine*", Lavoro di gruppo 2° sessione, a cura di B.C. RAMPONI, A.A. HUSSEIN, M. BELLADONNA, D. MARZINOTTO, T.J. MACKINTOSH, P. DELATO, M. MINENNA, P. COPPOLA, K.A. RAHMANI, H. ZAO, B.G. SCAFURI, E. ZIELLO, S. VALIO, M. DRAGONI, A. ADORNI, Centro Alti studi per la difesa, Istituto alti studi per la difesa, 71° Sessione di studio IASD, 2019-2020; Capitolo 3 "Comparazione tra il ciclo di intelligence attuale e le modifiche apportabili", Paragrafo C (pag. 41 - 50) "Generazione dei Deep Fakes".

*3.3 Il fenomeno del deep-fake e la tutela dell'immagine*

La straordinaria somiglianza tra i sistemi di intelligenza artificiale e le azioni dell'essere umano genera notevoli fraintendimenti che possono essere sfruttati dai criminali a loro comodo. Infatti, è possibile realizzare truffe *on-line* attraverso l'utilizzo di *chatbot* che, oltre alle loro potenzialità lecite, possono instaurare conversazioni realistiche con le vittime facendole cadere in errore. Inoltre, tramite il fenomeno del c.d. *deep-fake*[813] (unione tra le parole "*deep learning*" e "*fake*") è possibile utilizzare l'intelligenza artificiale per diffondere video falsi riguardanti una specifica persona rappresentando un evento nei fatti mai accaduto. Il termine nasce nel 2017 quando un utente del *social network Reddit,* il cui *nickname* era "*deepfakes*", utilizzò tale tecnologia per sostituire i volti di attori di video pornografici con quelli di personaggi famosi.

I *deep-fake* sono immagini o video in cui l'intelligenza artificiale riesce a riprodurre in modo realistico il volto e le movenze di un soggetto in base alla volontà dell'autore. In tal modo si creano falsi difficili da riconoscere. Si riesce infatti a realizzare uno scambio di volto molto realistico e quasi impercettibile grazie ad algoritmi di autoapprendimento che mappano il volto, generando una contraffazione fortemente realistica di materiali visivi preesistenti facendo sembrare vero ciò che non è mai accaduto. [814]

Essi non possono essere considerati di per sé dannosi in quanto permetto di alterare i volti anche per scopi legittimi (come l'anonimizzazione di una pubblica denuncia di violenza e la realizzazione di film, pubblicità, giochi o intrattenimento). Ma tali azioni rischiano di essere dannose e lesive per l'immagine dei soggetti falsamente rappresentati. Ad esempio l'attrice Scarlett Johansson ha denunciato la presenza di *deep-fake* che la ritraggono in video pornografici. Con questo fenomeno è possibile aumentare la diffusione di *fake news*, di *hate speech* e di atti di *revenge porn*. Infatti, tramite la realizzazione di video falsi si ha la possibilità di simulare un discorso politico[815] (come nel caso dell'aprile 2018 in cui fu diffuso un video falso di Obama in cui il regista Jordan Peele costruisce un *deep-fake* in cui mostra come apparenti informazioni rilasciate dall'*ex* Presidente degli USA siano in realtà manipolate dallo stesso autore, al fine di dimostrare la pericolosità di tale fenomeno), [816] di realizzare video pedopornografici con la presenza di un falso minore e di incitare all'odio verso una determinata categoria di soggetti.

---

[813] A. LONGO E G. SCORZA, op. cit., Paragrafo 3 (pag. 88 - 147) "Vantaggi e rischi non economici dell'AI - I rischi etici, sociali e politici dell'AI - Algoritmi che discriminano".

[814] R. FINOCCHI, A. PERRI, P. PEVERINI, "*La prova dell'enunciazione. Fotografia digitale, deepfake e pertinenza documentale denegata o rinegoziata*", in EC E|C Rivista dell'Associazione Italiana di Studi Semiotici, XIV, n. 30, 2020 • Mimesis Edizioni, Milano-Udine; Paragrafo 3 (pag. 140 - 142) "Il *deepfake* come dispositivo. L'enunciazione nell'evoluzione della simulazione visiva".

[815] A. LONGO E G. SCORZA, op. cit., Paragrafo 3 (pag. 88 - 147) "Vantaggi e rischi non economici dell'AI - I rischi etici, sociali e politici dell'AI - Algoritmi che discriminano".

[816] R. FINOCCHI, A. PERRI, P. PEVERINI, "*La prova dell'enunciazione. Fotografia digitale, deepfake e pertinenza documentale denegata o rinegoziata*", in EC E|C Rivista dell'Associazione Italiana di Studi Semiotici, XIV, n. 30, 2020 • Mimesis Edizioni, Milano-Udine; Paragrafo 3 (pag. 140 - 142) "Il *deepfake* come dispositivo. L'enunciazione nell'evoluzione della simulazione visiva".

L'algoritmo posto dietro la produzione di video falsi è denominato *GAN* (*Generative Adversarial Networks*) e apprende grazie al meccanismo della sfida tra due reti neurali in competizione in cui una ha l'obiettivo di costruire immagini false di soggetti e l'altra di imparare a riconoscerli. Quindi, con lo scopo di vincere sull'altra rete, la prima realizzerà *deep-fake* sempre più realistici.

I rischi sono dunque, come per le *fake news*, una maggiore disinformazione della società e una perdita di fiducia da parte della stessa nei confronti delle reali fonti di informazioni. [817]

Il fenomeno del *deep-fake* ha l'abilità di unire in sé capacità algoritmiche e informatiche elevate con tecniche di psicologia che permetto di raggiungere l'obiettivo sperato da parte dell'autore attraverso l'influenza esercitata sul pubblico.

Infatti, l'attacco informatico è costituito da quattro fasi: l'analisi iniziale delle informazioni sulla vittima, in tale fase l'utilizzo dell'intelligenza artificiale è massimo realizzando azioni fortemente intrusive nella vita dei soggetti; la costruzione dell'attacco, allo scopo di non permettere più alla vittima di proteggere e recuperare le informazioni sottratte; la fuga, ossia l'eliminazione delle prove dell'intrusione; ed infine, le contromosse utili a sviare e minimizzare i possibili contro attacchi.

In tale ambito gli scenari aperti dalla realizzazione dei *deep-fakes* attraverso l'impiego dell'intelligenza artificiale possono avere un notevole impatto sociale.

Per realizzarlo è necessario partire da un *file* audio ascrivibile ad un soggetto specifico utilizzando un *software* intelligente; successivamente, attraverso l'analisi di immagini relative a espressioni facciali del soggetto, si realizzerà il corretto ordine di movimenti labiali ed espressivi da inserire nel video in base all'audio. In tal modo, si creerà un *file* video che simula un discorso svolto dall'individuo determinato. Ad oggi, è possibile realizzare un video ad altissima qualità in brevissimo tempo, rendendo inoltre fortemente complesso l'individuazione del falso.[818]

Gli attori del fenomeno si dividono in quattro categorie: appassionati di *deep-fake* che utilizzando l'intelligenza artificiale per realizzare video al fine di rivenderli; attivisti politici che hanno lo scopo di costituire campagne virali di divulgazione di messaggi, manipolando l'opinione pubblica e rischiando di generare disinformazione; criminali come terroristi, truffatori o *hacker* che utilizzano i *deep-fake* per commettere illeciti; e infine attori legittimi (come le società cinematografiche).[819]

Tali azioni minano la fiducia della società verso l'informazione veritiera, mettono a rischio la sicurezza nazionale diffondendo pericolose propagande politiche e interferendo con le campagne elettorali. Inoltre la *cybersecurity* risulta anch'essa minacciata dal fenomeno del *deep-fake*. [820]

---

[817] A. LONGO E G. SCORZA, op. cit., Paragrafo 3 (pag. 88 - 147) "Vantaggi e rischi non economici dell'AI - I rischi etici, sociali e politici dell'AI - Algoritmi che discriminano".
[818] B. G. BRANCATI, op. cit., Capitolo 3 "Comparazione tra il ciclo di intelligence attuale e le modifiche apportabili", Paragrafo C (pag. 41 - 50) "Generazione dei Deep Fakes".
[819] M. WESTERLUND, "*The Emergence of Deepfake Technology: A Review*", Technology Innovation Management Review, November 2019 (Volume 9, Issue 11); Paragrafo 5 (pag. 41 - 42) "Who Produces Deepfakes?".
[820] M. WESTERLUND, op cit., Paragrafo 6 (pag. 42 - 43) "The Possible Threats of Deepfakes".

Al momento il fenomeno del *deep-fake* non è oggetto di regolazione civile o di repressione penale specifica ed occorre valutare un'estensione delle norme vigenti tra cui la diffamazione, l'estorsione, la truffa e lo *stalking*.[821]

Infatti, i contenuti *deep-fake*, per quanto non reali, sono sicuramente in grado di offendere gravemente la reputazione del soggetto tramite il *web* e dunque possono integrare le fattispecie di cui all'articolo 595 c.p.[822] La diffamazione risulta essere una fattispecie connessa sia al fenomeno delle *fake news* che al fenomeno del *deep-fake*. Infatti, il reato tutela l'onore della vittima inteso in senso soggettivo come sentimento della propria dignità morale (la somma dei valori che ogni individuo attribuisce a sé stesso) e oggettivo in termini di riflesso della stima e opinione che gli altri hanno in relazione ad un individuo (la reputazione, ossia il patrimonio morale che deriva dall'altrui considerazione). Si tratta di un bene altamente personale con fondamento nella Costituzione agli articoli 2 e 3. Per onore in termini normativi si fa riferimento al valore interiore della persona che risulta sempre meritevole di protezione penale anche per i soggetti privi di reputazione in quanto connesso alla dignità umana e del singolo. [823]

Per la configurazione del delitto di cui all'articolo 595 c.p., è necessaria la comunicazione a più persone di un'offesa all'altrui reputazione. Si tratta di un reato comune a danno di persone fisiche o giuridiche determinate. La condotta di offesa alla reputazione deve possedere tre requisiti fondamentali: l'assenza dell'offeso (dunque un'impossibilità di difesa da parte della vittima valutando *ex ante* l'idoneità a percepire l'offesa), l'attacco all'altrui reputazione (anche se già compromessa) tramite mezzo orale, scritto, stampa o informatico-telematico, e infine, la comunicazione a più persone (terzi estranei devono percepire e comprendere la manifestazione lesiva). Nel caso di utilizzo di *social network* per diffondere l'offesa, si presume la comunicazione a più persone nel momento in cui il messaggio offensivo viene pubblicato sul sito. È un reato istantaneo che si consuma nel momento e luogo della comunicazione lesiva. [824]

Tale delitto è aggravato, dal comma terzo dell'articolo in questione, dalla diffusione del messaggio attraverso mezzo stampa, qualsiasi altro mezzo di pubblicità o atto pubblico. Interpretando la norma è possibile comprendere come rientri tra tali mezzi anche *Internet* inteso come mezzo di comunicazione di massa. [825]

In tale reato rientra anche il fenomeno del *deep-fake* che, creando video o immagini false, può realizzare contenuti diffamatori e falsi che possono essere fatti circolare rapidamente sui *social network* anche *attraverso social bot* di intelligenza artificiale ledendo l'onore delle vittime.

Inoltre, tali contenuti *deep-fake* possono essere realizzati allo scopo di minacciare ed estorcere (ad esempio denaro in cambio della non pubblicazione) qualcosa alla vittima commettendo quindi il reato di cui

---

[821] M. WESTERLUND, op. cit., Paragrafo 8 (pag. 44 - 46) "Methods to Combat Deepfakes".
[822] N. ORDONSELLI, "*«Porno Deepfake»: profili di diritto penale*", in Cyberlaws, 18 gennaio 2021: https://www.cyberlaws.it/2021/porno-deep-fake-profili-penalistici-reato/
[823] P.M. SABELLA, "*Il delitto di diffamazione. Struttura del fatto tipico nella dimensione offline e online*", Lezione del 18 marzo 2020, Luiss Guido Carli, Dipartimento di Giurisprudenza, Cattedra di Diritto penale 2, Titolari di insegnamento – Prof.ri E. GALLUCCI E M.N. MASULLO, Paragrafo 1 (pag. 1 – 2) "Caratteri generali e bene giuridico tutelato – Il concetto di onore".
[824] P.M. SABELLA, op. cit., Paragrafo 4 (pag. 3 – 4) "La condotta offensiva".
[825] P.M. SABELLA, op. cit., Paragrafo 8 (pag. 5 – 9) "La diffamazione a mezzo stampa o con qualsiasi altro mezzo di pubblicità – Il ruolo dei social network e gli internet service providers (ISP): editori o meri contenitori?".

all'articolo 629 c.p. [826] Infatti, il reato di estorsione consiste in un delitto a cooperazione artificiosa della vittima in cui il soggetto attivo costringe attraverso una condotta violenta o minacciosa il soggetto passivo (che subisce il danno) a commettere un'azione od omissione e può essere inoltre definito come delitto predatorio e parassitario. Tale seconda definizione è connessa al fatto che il *reo* vada ad instaurare un rapporto diretto con la vittima al fine di coartarne la volontà con violenza o minaccia. La modalità della condotta sfocia quindi in un condizionamento psicologico della vittima causato dalla violenza o minaccia (ossia la prospettazione di un male ingiusto in caso di mancata adesione alla richiesta) a persone o cose; l'idoneità della condotta va valutata *ex ante* in concreto. Il reato è definibile, quindi, come delitto plurioffensivo colpendo sia il patrimonio che la libertà di autodeterminazione dell'offeso. Si tratta inoltre di un reato a duplice evento: è necessario che si realizzi sia l'ingiusto profitto per il *reo* che il danno per la vittima.[827]

Per di più, il fenomeno del *deep-fake* potrebbe avere rilevanza penale in tema di truffa (*ex* art. 640 c.p.), delitto che punisce la condotta di chiunque, attraverso artifici o raggiri, inducendo taluno in errore, procuri a sé o ad altri un ingiusto profitto con altrui danno. Si tratta anche qui, come per il reato di estorsione di un delitto comune a duplice evento e di cooperazione artificiosa della vittima. Il *deep-fake* potrebbe facilmente rientrare nel concetto di artificio essendo lo stesso una manipolazione della realtà esterna che simula circostanze inesistenti e dissimulando circostante esistenti. Non rientrerebbe invece nel concetto di raggiri che costituiscono una simulazione della realtà tramite parole o argomentazioni. L'idoneità di tali condotte ad indurre in errore (ossia un vizio della volontà consistente in una falsa rappresentazione della realtà causata dall'illecita azione del *reo*) il soggetto va valutata nel caso concreto. [828]

Infine, risulta rilevante l'ipotesi in cui il *deep-fake* (soprattutto in ambito pornografico) vada ad integrare il reato di atti persecutori *ex* art. 612 *bis* c.p. che prevede l'aggravante dell'utilizzo di strumenti informatici o telematici per generare un fondato timore alla vittima. [829] Lo *stalking* è stato inserito all'interno del nostro codice penale dalla legge n. 38 del 23 aprile 2009[830], allo scopo di punire condotte reiterate di minaccia o molestia che causano tre eventi alternativi: il fondato timore per la propria o altrui incolumità, l'alterazione delle abitudini di vita e la generazione di uno stato di ansia o di paura nella vittima.[831] L'inserimento dell'art. 612 *bis* c.p. e l'individuazione del delitto di atti persecutori come reato autonomo è stato fortemente criticato dalla dottrina che lo considerava come superfluo data la possibilità di far ricadere le previste condotte persecutorie all'interno di altre norme del codice penale. Sul punto si è pronunciata anche la Corte

[826] N. ORDONSELLI, op. cit.
[827] E. GALLUCCI, "*Appunti delle lezioni di Diritto penale parte speciale a.a. 2019 – 2020*", presso l'Università Luiss Guido Carli, Dispensa n. 4 "Estorsione e sequestro di persona a scopo di estorsione".
[828] E. GALLUCCI, op. cit., Dispensa n. 6 "La truffa".
[829] N. ORDONSELLI, op. cit.
[830] Legge 23 aprile 2009, n. 38, "*Conversione in legge, con modificazioni, del decreto-legge 23 febbraio 2009, n. 11, recante misure urgenti in materia di sicurezza pubblica e di contrasto alla violenza sessuale, nonché in tema di atti persecutori*", pubblicata nella *Gazzetta Ufficiale* n. 95 del 24 aprile 2009: https://web.camera.it/parlam/leggi/090381.htm
[831] G. FIANDACA E E. MUSCO, "*Diritto penale – Parte Speciale*" Volume II, Tomo primo "*I delitti contro la persona*", Quarta edizione, Zanichelli Editore, 1 gennaio 2013, Bologna, Capitolo 4 "Delitti contro la libertà personale e la libertà morale", Sezione II "Delitti contro la libertà morale", Paragrafo 6 (pag. 224 – 234) "Stalking".

Costituzionale[832] che, nel 2014, ha ritenuto infondata la questione di costituzionalità sollevata per contrasto con l'articolo 25 della Costituzione, motivando sulla scorta dell'indeterminatezza della nuova fattispecie. La questione era stata sollevata affermando che tale fattispecie peccava di poca chiarezza e precisone, e che le condotte non fossero suscettibili di verificazione empirica in modo univoco a causa del fatto che tutti gli eventi colpissero esclusivamente la sfera psichica della vittima la quale varia a seconda del soggetto e che inoltre, risulta difficile da accertare in concreto. La Corte Costituzionale nel dichiararla infondata afferma che tutti i termini dell'art. 612 *bis* c.p. hanno significato chiaro e preciso in concreto, considerando anche il fatto che gli stessi rientrano nel linguaggio penalistico.

L'articolo 612 *bis* c.p. è dunque un autonomo reato comune di evento (non basta solo la condotta) a forma vincolata e abituale (la condotta deve necessariamente consistere in minacce o molestie reiterate). [833] Ma in relazione al tema trattato in questo paragrafo, è interessante soffermarci sul concetto di mezzi informatici intesi in tale norma allo scopo di applicare l'aggravante. Infatti, secondo la Sentenza della Cassazione del gennaio 2019[834], il reato di *stalking* si configura nel momento in cui la condotta minacciosa del *reo* sia idonea a destabilizzare l'equilibrio psichico della persona offesa (a prescindere da un possibile incontro fisico tra vittima e autore). Il contenuto di tali condotte (solitamente consistenti in messaggi, *e-mail*, chiamate, *post* sui *social*) deve essere determinato da una grave intrusione nella sfera intima della persona e dunque bisogna valutare in concreto l'intensità del tenore persecutorio dei messaggi. Il *deep-fake* rientra perfettamente in tale aspetto essendo lo stesso idoneo, soprattutto nel caso del porno *deep-fake* a generare un grave stato di ansia nella persona offesa trovandosi quindi violata (virtualmente) nella sua intimità.

### *3.4 La pedopornografia e il revenge porn ai tempi del deep-fake*

Le problematiche connesse al fenomeno del *deep-fake* riguardano inoltre la realizzazione di video pornografici o pedopornografici virtuali, considerabili come reati di cui all'articolo 612 *ter* c.p. e soprattutto all'articolo 600 *quater* 1 c.p.. Inoltre, qualora tali materiali dovessero circolare nella rete e nei *social network*, il compito-dovere dell'*Internet Service Provider* sarebbe quello di eliminare il video una volta scoperta la sua illiceità in base alle riflessioni espresse precedentemente.[835]

[832] Sentenza della Corte Costituzionale n.172 del 11 giugno 2014.
[833] G. FIANDACA E E. MUSCO, op. cit., Capitolo 4 "Delitti contro la libertà personale e la libertà morale", Sezione II "Delitti contro la libertà morale", Paragrafo 6 (pag. 224 – 234) "Stalking".
[834] Sentenza della Cassazione Penale, Sez. V, n. 61 del 2 gennaio 2019.
[835] B. PANATTONI, "*Gli effetti dell'automazione sui modelli di responsabilità: il caso delle piattaforme online*", op. cit., Paragrafo 3 (pag. 39 – 41) "La responsabilità penale degli ISP in caso di caricamento e/o diffusione di contenuti illeciti in rete".

Infatti, il *deep-fake*, come già accennato può essere utilizzato per realizzare video pornografici falsi andando a modificare il volto o il fisico dei protagonisti al fine di commettere atti di *revenge porn* incriminato a seguito della riforma data dalla legge n.69/2019[836] (c.d. Codice Rosso).[837]

A pochi anni dalla riforma siamo già difronte ad una nuova modalità di *revenge porn* che permette di distruggere la vita di una persona grazie ad un algoritmo intelligente: una nuova frontiera per le molestie *on-line*. Emblematica risulta essere l'applicazione *DeepNude* lanciata nel 2019 che permette di "spogliare" i corpi delle foto.[838] Tale applicazione è stata scaricata da più di 95000 persone ed il creatore è stato costretto a chiuderla per motivi etici; ma il codice sorgente continuò a circolare. Infatti, rapidamente fu realizzato un *bot* su *Telegram* al fine di permettere agli utenti un accesso facile a video e immagini pornografiche realizzate con il *deep-fake* senza il consenso.[839]

La gravità di tale fenomeno è riscontrabile dai dati ricavabili da due *report* realizzati dalla *Sensity* (società di *cybersecurity* che si occupa di monitorare il fenomeno del *deep-fake*). Nel primo[840], del 2019, si riferisce che il 96% dei video *deep-fake on-line* era costituito da video pornografici; nel secondo[841], del 2020, emerge come nel solo mese di luglio le immagini circolanti su *Telegram* di donne virtualmente denudate (attraverso l'intelligenza artificiale) ammontano ad almeno 104.852. Ad oggi, il numero di tali immagini supera i 680.000.

I dubbi consistono nel comprendere se tali azioni possano essere considerate come reato e rientranti sotto la fattispecie di cui all'articolo 612 *ter* c.p (diffusione illecita di immagini o video sessualmente espliciti).

Le immagini realizzate sono estremamente realistiche e capaci di ledere fortemente la dignità e la liberà di autodeterminazione della vittima; ma non esiste nel nostro ordinamento una norma che tuteli le vittime di diffusione di materiali sessualmente espliciti realizzati attraverso l'intelligenza artificiale.

Nonostante ciò, volgiamo per un attimo la nostra attenzione sull'articolo 612 *ter* c.p. che ha lo scopo di tutelare la riservatezza di contenuti multimediali che potrebbero ledere la dignità della vittima punendo la condotta di diffusione senza consenso di immagini o video sessualmente esplicito realizzati o ottenuti da parte della vittima o di terzi al fine di recare danno al soggetto ivi raffigurato. In tale articolo è prevista infatti l'aggravante di commissione del fatto con strumenti informatici o telematici potendo dunque ricomprendere il fenomeno del *deep-fake* (anche se il testo della norma non fa riferimento a contenuti non reali). I dubbi concernono il rispetto del principio di tassatività della norma penale che risulta difficilmente conciliabile con l'inserimento del fenomeno del *deep-fake* nella fattispecie incriminatrice di cui all'articolo sopra citato.

---

[836] Legge 19 luglio 2019, n. 69, "*Modifiche al codice penale, al codice di procedura penale e altre disposizioni in materia di tutela delle vittime di violenza domestica e di genere*" (GU n.173 del 25-7-2019).

837 M. WESTERLUND, op. cit., Paragrafo 7 (pag. 43 - 44) "Current Examples of Deepfakes".

[838] F.M.R. LIVELLI, "*Deepfake e revenge porn, combatterli con la cultura digitale: ecco come*", in Network Digital 360 – Cybersecurity 360, 8 febbraio 2021: https://www.cybersecurity360.it/nuove-minacce/deepfake-e-revenge-porn-combatterli-con-la-cultura-digitale-ecco-come/

[839] N. ORDONSELLI, op. cit.

[840] H. Ajder, G. Patrini, F. Cavalli e L. Cullen, "*The state of deepfakes landscape, threats, and impact*", Deeptrace, Settembre 2019: http://regmedia.co.uk/2019/10/08/deepfake_report.pdf

[841] H. AJDER, G. PATRINI E F. CAVALLI "*Automating Image Abuse, Deepfakes Bots on Telegram*", Sensity, Ottobre 2020: https://www.medianama.com/wp-content/uploads/Sensity-AutomatingImageAbuse.pdf

Attenendoci al significato letterale del testo, ci troveremmo difronte a una peculiare forma di diffusione non consensuale di materiale sessualmente esplicito appartenente alla vittima: ossia il caso in cui questo estrinsechi un intento vendicativo realizzato da un soggetto legato, da una relazione sentimentale attuale, precedente, futura o immaginata, alla vittima. In realtà, oggi si sente frequentemente parlare di *revenge porn* anche in caso di assenza di tale relazione sentimentale o scopo vendicativo. Infatti tale terminologia viene indistintamente utilizzata anche con riferimento al fenomeno in cui si sabotano dispositivi informatici allo scopo di diffondere il materiale sessualmente esplicito contenuto all'interno o la diffusione a fini estortivi.[842]

È quindi possibile notare come l'articolo 612 *ter* c.p. riesca ad arginare solo una parte del tragico problema della pornografia non consensuale risultando dunque in molti ambiti inadeguato.[843]

Infatti, nel rispetto del tenore letterale della norma e del fondamentale principio di tassatività in ambito penale, non è possibile punire, senza rientrare in un'interpretazione estensiva e analogica in *malam partem*, attraverso l'articolo 612 *ter* c.p. anche le condotte aventi ad oggetto rappresentazioni non reali.

Quindi, sarebbe auspicabile una rimodulazione dell'articolo andando ad includere la condotta di formazione, sottrazione e appropriazione di immagini o video sessualmente espliciti senza dover necessariamente attendere la loro diffusione per intervenire penalmente ed inoltre l'inserimento anche un comma apposito relativo al fenomeno del *deep-fake* (che risulta ugualmente lesivo per la vittima). [844]

Per quanto riguarda invece, il caso in cui le immagini virtuali sessualmente esplicite, realizzate con l'utilizzo di sistemi di intelligenza artificiale, raffigurino soggetti minorenni, i problemi interpretativi si attenuano dato che, ai sensi dell'articolo 600 *quater* 1 c.p. (disciplinante il reato di pedopornografia virtuale) si considerano integrati i reati di cui agli articoli 600 *ter* e 600 *quater* c.p. anche nel caso di immagini virtuali.

La virtualità delle immagini viene definita come la realizzazione delle stesse tramite tecniche di elaborazione grafica non del tutto associate a situazioni reali la cui rappresentazione le faccia apparire come tali.

L'effettiva offensività di tale condotta non risulta immediatamente percepibile: la norma è stata inserita al fine di tutelare maggiormente il minore nel caso di diffusione di sue immagini (anche virtuali) sessualmente esplicite. Non si va a tutelare (a differenza dei due articoli immediatamente precedenti, che implicano la presenza di un minore nella rappresentazione fotografica o cinematografica sessualmente esplicita) la liberà di autodeterminazione e l'integrità psico-fisica del minore, in quanto, in tale reato, non è coinvolto nessun minore in prima persona. Vengono infatti parificate (o quasi, nell'articolo 600 *quater* 1 c.p. la pena è diminuita) le condotte di realizzazione e diffusione di video o immagini sessualmente esplicite con minori reali o virtuali (ossia alterazioni o elaborazioni grafiche realizzate anche attraverso l'intelligenza artificiale).[845]

---

[842] N. AMORE, "*La tutela penale della riservatezza sessuale nella società digitale. Contesto e contenuto del nuovo cybercrime disciplinato dall'art. 612 ter c.p.*", in Legislazione penale, 20 gennaio 2020; Paragrafo 2 (pag. 3 – 7) "Il contesto della riforma: la tutela della libertà e della riservatezza sessuale nella società 4.0": http://www.lalegislazionepenale.eu/wp-content/uploads/2020/01/N.-Amore-Approfondimenti-1.pdf

[843] N. AMORE, op. cit., Paragrafo 5.1 (pag. 30 – 33) "Le criticità dell'art. 612-*ter* c.p.".

[844] N. AMORE, op. cit., Paragrafo 6 (pag. 36 – 38) "Una proposta di riforma dell'art. 612-*ter* c.p.".

[845] N. ORDONSELLI, op. cit.

Nonostante tali legittimi dubbi, la Cassazione, nel 2017[846] ha stabilito che tali alterazioni devono essere considerate come reato in quanto alimentano "*l'attrazione per manifestazioni di sessualità rivolte al coinvolgimento di minori*" mettendo in pericolo la personalità in formazione del minore stesso. Di estrema rilevanza, è in questo campo, il tema della funzione del diritto penale, ossia un diritto tradizionalmente sanzionatorio e di *extrema ratio*. Un'anticipazione della tutela realizzata in tal modo, andando a limitare le condotte che manifestano l'attrazione verso i minori, renderebbe il diritto penale educativo e di conseguenza contrario ai principi generali del diritto penale illuministico. In altri termini, si punirebbe il pedofilo, non per il fatto commesso, ma per la sua tendenza sessuale deviante, pur non concretamente manifestata con atti dannosi.

Ancora, nel 2018[847], la stessa sezione della Cassazione ha affermato che la pedopornografia virtuale realizzata attraverso l'utilizzo di fotomontaggi configuri il reato di cui all'articolo 600 *quater* 1 c.p.

Tale norma sembra dunque applicabile anche nel caso di alla realizzazione di immagini pedopornografiche tramite le tecniche del *deep-fake*. Di conseguenza sarebbe opportuno estendere la tutela anche al reato di cui all'articolo 612 *ter* c.p.

In conclusione, per garantire una maggiore tutela delle vittime e per evitare di ricadere in lesioni al principio di tassatività delle norme penali, è importante che il legislatore adotti una normativa *ad hoc* per disciplinare tale innovativo fenomeno criminale capace di avere effetti devastanti sulle vite delle vittime. [848]

Altre problematiche connesse alla tutela dei minori e al *deep-fake* riguardano la possibilità, da parte di predatori sessuali di utilizzare l'intelligenza artificiale per tramutare le loro fattezze in quelle di minori al fine di rendere più efficaci le loro attività di *child grooming* (punita all'articolo 609 *undicies* c.p.).[849] Tale azione consiste nel processo mediante il quale si spinge un minore a stringere un rapporto di "amicizia" con un molestatore nel tentativo di ottenerne la fiducia al fine di convincerlo ad accettare e sottostare ad un'attività sessuale. Non si tratta di un fenomeno nuovo ma sicuramente l'avvento di *Internet* e delle forme di comunicazione *on-line* hanno fornito agli aggressori opportunità maggiori di contatto con i minori ed inoltre il *deep-fake* favorisce il travisamento dell'adulto come minore.

L'azione di adescamento può essere divisa in cinque fasi: *Friendship Forming Stage*, *Relationship Forming Stage*, *Risk Assesstement Stage*, *Exclusivity Stage* e *Sexual Stage*.

La prima concerne nell'approccio con il minore ed è qui che il *deep-fake* viene realizzato e messo alla prova. L'aggressore, dopo aver attentamente studiato il comportamento, le caratteristiche e i gusti della vittima, realizza un profilo falso appositamente creato (anche con la realizzazione di immagini e video tramite il supporto dell'intelligenza artificiale) al fine di attirare l'attenzione del minore e di instaurare un rapporto di fiducia. La seconda fase riguarda la realizzazione di un legame di fiducia con la vittima cercando di entrare il

[846] Sentenza della Cassazione Penale, Sez. III n. 22265 del 13 gennaio 2017.
[847] Sentenza della Cassazione Penale, Sez. III n. 15757 del 9 aprile 2018.
[848] N. ORDONSELLI, op. cit.
[849] M. WESTERLUND, op. cit., Paragrafo 6 (pag. 42 - 43) "The Possible Threats of Deepfakes".

più possibile in confidenza con lui così da sfruttare le informazioni personali che il minore gli fornirà. La fase del *Risk Assesstement Stage* consiste nel tentativo di spostare la relazione in un contesto reale invece che virtuale cercando di concretizzare i suoi scopi. Nella penultima fase l'adescatore indaga sui desideri sessuali ed intimi della vittima allo scopo di generare una relazione semi-amorosa e di dipendenza. Infine, con la fase del *Sexual Stage*, si passa alla violenza o abuso sessuale vero e proprio, fisico o virtuale (attraverso lo scambio di materiale pedopornografico), ai danni del minore.[850]

L'articolo in questione, inserito a seguito della Convenzione di Lanzarote del 2007[851], punisce, con la forma del reato comune caratterizzato dal dolo specifico, chiunque, allo scopo di commettere i reati connessi alla prostituzione, pornografia e violenza o abuso sessuale a danni di un minore, adesca tale minore. Per adescamento, il codice penale si riferisce a qualsiasi atto volto a carpire la fiducia del soggetto anche attraverso artifici, lusinghe o minacce realizzate anche tramite l'utilizzo di mezzi informatici o telematici di comunicazione.

A differenza di quanto previsto della Convenzione, che all'articolo 23 impone agli Stati di inserire una norma idonea a punire le condotte di adescamento (esclusivamente *on-line* o per mezzo di strumenti telematici) che si sono rilevate idonee ad organizzare un effettivo incontro con il minore, l'articolo 609 *undicies* c.p. anticipa la tutela punendo il *reo* già al compimento della fase di costruzione del rapporto di fiducia con il minore mosso da scopi sessuali inserendo così un reato di pericolo indiretto al fine di meglio tutelare la libertà e l'equilibrato sviluppo psico-sessuale del minorenne. [852]

È indubbio che il *deep-fake* rientri in un artificio che possa spingere il minore a fidarsi del *reo* creando ad esempio un falso profilo sui *social network* che sia il più realistico possibile.

### *3.5 Lo stupro e l'abuso sessuale di "minori robotici"*

Come accennato alla fine del primo capitolo, è stato ipotizzata (e al contempo criticata) la possibilità di considerare i sistemi di intelligenza artificiale come persone offese del reato attraverso il paragone tra il loro livello di intelligenza a quella di un animale dotato di un margine di razionalità.

Le condotte di lesione o maltrattamento di tali sistemi non possono essere paragonati al maltrattamento contro familiari e conviventi (*ex* art. 572 c.p.) o maltrattamento di animali (*ex* art. 544 *ter* c.p.) ma, qualora le

[850] L. DELL'AQUILA, "*Il fenomeno del child grooming: l'adescamento di minori dentro e fuori la rete*", in Cyberlaws, 15 maggio 2019: https://www.cyberlaws.it/en/2019/adescamento-minori-child-grooming/

[851] Convenzione del Consiglio d'Europa sulla "*protezione dei minori dallo sfruttamento e dagli abusi sessuali*", conclusa a Lanzarote il 25 ottobre 2007: https://fedlex.data.admin.ch/filestore/fedlex.data.admin.ch/eli/cc/2014/249/20150826/it/pdf-a/fedlex-data-admin-ch-eli-cc-2014-249-20150826-it-pdf-a.pdf

[852] L. DELL'AQUILA, op. cit.

capacità intellettuali di tali sistemi dovessero aumentare a seguito dell'evoluzione tecnologica, il legislatore potrà trovarsi difronte al dilemma di come istituire una tutela penale per tali sistemi.

Esistono infatti i c.d. *sexbot robot* sessuali dotati di caratteristiche umane che grazie all'intelligenza artificiale riescono ad emulare l'attività sessuale ed affettiva di un essere umano. Il *sexbot* è infatti necessariamente provvisto di tre caratteristiche: una forma umana, la capacità di muoversi e l'essere dotato di intelligenza artificiale. Questo li differenzia dalle altre tipologie di *sextoys* i quali non possiedono tutti e tre i requisiti.

Il primo *sexbot* ad essere stato creato, prodotto dalla società "*True Companion*", si chiama "*Roxxxy-Rocky*": la tecnologia è ancora grezza ma i progressi che si faranno nel prossimo periodo in relazione alla creazione di *robot* umanoidi intelligenti, rende necessario un approfondimento sullo sviluppo di tale utilizzo della tecnologia. [853]

I rischi connessi alla realizzazione dei *sexbots*, concernono il pericolo di creazione di *robot* con sembianze di minori allo scopo di simulare commettere atti sessuali non consensuali. [854]

È dunque possibile effettuare un paragone con la disciplina della pedopornografia virtuale.

Infatti, le problematiche riscontrate in relazione all'articolo 600 *quater* 1 c.p. quanto al rischio di assenza di un'effettiva lesione di un bene giuridico appartenente ad una reale vittima sono riscontrabili anche qui, rischiando di ricadere in una sanzione penale per una mera devianza interna del soggetto. Come nella citata Sentenza del 2017[855] in tema di pedopornografia virtuale in base alla quale si afferma che l'articolo 600 *quater* 1 c.p. mira ad evitare l'aumento dell'attrazione verso manifestazioni di sessualità che coinvolgono i minori, anche in questo ambito, il bene giuridico non sarebbe la libertà sessuale del *robot* ma si verrebbe ad anticipare la tutela della sessualità dei minori sanzionando la diffusione di condotte di pedofilia seppur robotica. Si ripropone, dunque, la già affermata questione relativa funzione del diritto penale eccessivamente paternalistico.

Per far ricadere tali azioni sotto una disciplina penalistica è importante che il legislatore riesca ad individuare un concreto bene giuridico tutelabile che venga messo a rischio da parte del *reo*.

Abbiamo già affermato come i *robot*, non avendo uno *status* di persone elettroniche titolari di diritti e doveri, non possono essere considerati come vittime del reato. Essi possono al massimo rilevare come oggetti materiale su cui si scatena la condotta o il mezzo attraverso cui compiere l'illecito penale.

Se tale tecnologia dovesse prendere piede e si dovesse diffondere notevolmente tra la popolazione, l'utilizzo dei *sexbots* andrà a sopperire a tutte quelle fantasie sessuali (devianti o meno) che l'essere umano non ha il coraggio di realizzare con altri esseri umani (non necessariamente arrivando a trattare di reati a sfondo sessuale). L'oggettivizzazione dell'atto sessuale e la ricerca di un piacere incondizionato sempre a

[853] J. DANAHER, "*Robotic Rape and Robotic Child Sexual Abuse: Should They Be Criminalized?*", in Criminal Law & Philosophy, 2017, p. 71 ss.; Paragrafo 2 "What are we talking about?".
[854] I. SALVADORI, op. cit., Paragrafo 8, "L'agente artificiale come oggetto materiale del reato".
[855] Sentenza della Cassazione Penale, Sez. III n. 22265 del 13 gennaio 2017.

disposizione verrebbero dunque ad enfatizzarsi grazie alla possibilità di programmare i *sexbots* nel modo più appagante per l'utente. Questo porterebbe i soggetti che fantasticano su atti di pedofilia o di violenze sessuali a sentirsi legittimati a commetterli su un *robot*, per definizione non umano, che riproduce fedelmente, nelle sembianze e nel comportamento, una persona. Questo, da una parte, andrebbe a fornire una valvola di sfogo per chi non riesce ad appagare i propri desideri sessuali ma dall'altra, rischierebbe di tradursi in un incentivo a commettere gli stessi atti nei confronti di una vittima umana.[856]

A questo punto si pone la domanda relativa alla possibilità di criminalizzare l'abuso sessuale o lo stupro di un *robot* con le sembianze di un minore. La prima questione riguarda il consenso: il *robot* non avendo una morale, non può prestare o negare il suo consenso ma allo stesso tempo può essere programmato per farlo. Quindi, un soggetto che decida di commettere un atto sessuale con un *robot* programmato per negare il suo consenso, può essere considerato come autore di uno stupro robotico? Lo stesso quesito vale per l'abuso sessuale su un "minore robotico".

Tali azioni, in base ad un'interpretazione paternalista del diritto penale, inteso a considerare illeciti comportamenti senza effetti direttamente dannosi sui soggetti, verrebbero criminalizzate in base al pericolo che da esse potrebbe derivare.[857]

È indubbio che chi commette atti di stupro su *sexbot* con sembianze di minori compia atti dimostrativi di un volere immorale esplicitando un'insensibilità verso il significato sociale di tali atti qualora realizzati con minori persone fisiche. [858]

Le perplessità riguardano, quindi, la possibilità di disciplinare penalmente una condotta moralmente sbagliata che però non ha concreti effetti dannosi verso il prossimo dato che gli atti di stupro o abuso sessuale su *sexbots* con sembianze di minori sono di sicuro moralmente sbagliati ma di fatto innocui verso i beni giuridici tutelati dall'ordinamento.

Tornare ad un diritto penale paternalistico non appare soluzione valida: non possono essere criminalizzati atti o fantasie che non hanno effetti lesivi sul prossimo o sulla società. Diversamente opinando, dovremmo criminalizzare tutti i videogiochi che permettono in un mondo virtuale di commettere illeciti e violenze. Il diritto penale interviene per tutelare beni giuridici che potrebbero essere lesi da determinate condotte. È pur vero che esistono i reati di pericolo che vanno ad anticipare la tutela in caso di condotte connotate da un certo grado di offensività prima ancora che il danno si realizzi. Nonostante tale possibilità, il caso in esame non potrebbe nemmeno definirsi come reato di pericolo perché il bene giuridico non verrebbe messo a rischio né concretamente né direttamente; anzi, considerando tali azioni come reato si finirebbe per punire la personalità deviante del *reo* prima che si realizzi alcuna condotta pericolosa o dannosa.

[856] J. DANAHER, op. cit., Paragrafo 2 "What are we talking about?".
[857] J. DANAHER, op. cit., Paragrafo 4 "The Moralistic Premise".
[858] J. DANAHER, op. cit., Paragrafo 5 "The Wrongness Premise".

Di conseguenza dobbiamo valutare se le condotte di violenza sessuale su *robot* con sembianze di minori possano in concreto risultare come lesive di beni giuridici tutelati. [859]

Per concludere, dunque, bisogna quindi sottolineare come i reati a sfondo sessuale criminalizzati all'interno del nostro codice penale si occupano di tutelare l'integrità psico-fisica, la libertà di autodeterminazione sessuale di adulti o minori persone fisiche.[860]

[859] J. DANAHER, op. cit., Paragrafo 4 "The Moralistic Premise".
860 I. SALVADORI, op. cit., Paragrafo 8, "L'agente artificiale come oggetto materiale del reato".

# ***CONCLUSIONE***

"*Credo che alla fine del secolo l'uso delle parole e l'opinione delle persone di cultura saranno cambiate a tal punto che si potrà parlare di macchine pensanti senza aspettarsi di essere contraddetti*."

Con questa citazione di Alan Turing si conclude l'elaborato.

Aveva ragione. Ad oggi è possibile parlare di macchine pensanti; anche se non ancora nella stessa modalità con cui ci si riferisce al pensiero di un essere umano.

I temi trattati in questa tesi sono molteplici e complessi ma tutto ruota intorno ad una "semplice" domanda: è ancora valido il brocardo *machina delinquere non potest*?

All'esito della trattazione sembra possibile dare una risposta affermativa a tale quesito: i sistemi intelligenti non sono ancora abbastanza autonomi da poter essere considerati come responsabili delle proprie azioni. Questo non può però affermarsi come certezza per il futuro nel caso in cui lo sviluppo tecnologico dovesse giungere fino alla realizzazione di un cervello robotico capace di raggiungere le capacità umane.

Nell'attesa siamo chiamati a valutare le possibili forme di responsabilità in capo a soggetti umani e dunque in base agli attuali canoni normativi in tema di responsabilità penale: la sussistenza di una posizione di garanzia in capo ad un soggetto specifico, un nesso causale tra l'azione della persona fisica e l'evento realizzato dalla macchina e soprattutto, l'esistenza di un elemento soggettivo, almeno al livello della colpa, al fine di evitare di ricadere in forme di responsabilità oggettiva.

Sarà configurabile, ovviamente, una responsabilità di tipo doloso in capo al soggetto utilizzatore o programmatore che utilizzi o programmi il sistema intelligente allo scopo di commettere illeciti (come nel caso del *revenge porn* realizzato attraverso il *deep-fake*).

Più complesso il tema della responsabilità medica o da omicidio e lesioni personali stradali, in quanto il soggetto (medico o guidatore) non ha agito con lo scopo di commettere un reato, ma l'evento lesivo si è realizzato comunque: in tal caso bisognerà quindi valutare, per il medico, il rispetto delle linee guida e delle buone pratiche clinico-assistenziale accreditate ai sensi di legge, e per il guidatore, l'osservanza delle disposizioni dettate in materia di circolazione stradale. Solo nel caso in cui tali norme cautelari non vengano rispettate allora potrà essere ipotizzata una responsabilità colposa di tali soggetti.

Inoltre, in generale, nel caso di responsabilità colposa del programmatore e del produttore, qualora gli stessi abbiano messo in commercio un *robot* intelligente consapevoli dei suoi rischi senza fare nulla per impedirli, si applicheranno, non sempre in modo chiaro e lineare, i criteri mutuati dalla responsabilità da prodotto difettoso. Allo stesso modo, non risulta semplice individuare un unico soggetto direttamente responsabile, in quanto, alla realizzazione di ogni singolo sistema intelligente collaboreranno un numero consistente di programmatori, ingegneri, scienziati, produttori e società.

È evidente l'impossibilità di dare una risposta univoca a tutti gli aspetti rilevanti del tema e come affermato fin dall'inizio di questa trattazione, questo non è lo scopo principale dell'elaborato.

Siamo difronte ad un cambiamento, anche nel mondo del diritto, ed è il momento di affrontarlo così da ridurre il divario tra progresso tecnologico ed evoluzione normativa.

# ***FONTI BIBLIOGRAFICHE, SITOGRAFICHE, NORMATIVE, DI GIURISPRUDENZA E FILMOGRAFICHE***

## ***BIBLIOGRAFIA***


ALESSANDRI A. e SEMINARA S., "*Diritto penale commerciale*", Volume 1 "I principi generali", G. Giappichelli Editore, Torino 2018.

AMIDEI A., "*Intelligenza artificiale e responsabilità da prodotto*", in RUFFOLO U., "*Intelligenza artificiale – Il diritto, i diritti, l'etica*", Collana Tech and Law di Giuffrè Francis Lefebvre Edizione, Milano 2020; Parte II "Diritto civile: responsabilità, contratti, persona e privacy", Sezione I "A.I. e responsabilità" Capitolo 2 (pag. 125 – 153).

AMIGONI F., SCHIAFFONATI V., SOMALVICO M., "*Intelligenza artificiale",* Istituto della Enciclopedia Italiana, Treccani, Enciclopedia della Scienza e della Tecnica, Roma, 2008: http://www.treccani.it/enciclopedia/intelligenza-artificiale_(Enciclopedia-della-Scienza-e-della-Tecnica)/

AMORE N., "*La tutela penale della riservatezza sessuale nella società digitale. Contesto e contenuto del nuovo cybercrime disciplinato dall'art. 612 ter c.p.*", in Legislazione penale, 20 gennaio 2020: http://www.lalegislazionepenale.eu/wp-content/uploads/2020/01/N.-Amore-Approfondimenti-1.pdf

ARCIERI S., "*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

ASIMOV I., "*I, Robot*", Mondadori, 1950.

BALDUCCI P. e MACRILLÒ A., "*Esecuzione penale e ordinamento penitenziario*", Giuffrè Francis Lefebvre S.p.a. Milano – 2020.

BASILE F., "*Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine",* in Diritto Penale e Uomo (DPU) – Criminal law and Human Condition, Milano, 2019: https://archiviodpc.dirittopenaleuomo.org/upload/3089-basile2019.pdf

BIRRITTERI E., "*Big Data Analytics e compliance anticorruzione. Profili problematici delle attuali prassi applicative e scenari futuri*" (pag. 289 – 303) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018, in Diritto penale Contemporaneo, Rivista trimestrale 2/2019.

BONNEFON J.F., RAHWAN I. e SHARIFF A., "*The social dilemma of autonomous vehicles*" in Science, Vol. 352, Issue 6293 (pag. 1573 – 1576), 24 giugno 2016: https://www.science.org/doi/10.1126/science.aaf2654

BORRUSO R., "*Informatica Giuridica*" voce dell'Enciclopedia del Diritto – I Aggiornamento – Giuffrè, Milano, 1994.

BORSARI R., "*Intelligenza Artificiale e responsabilità penale: prime considerazioni"* in rivista quadrimestrale di Media Laws, Rivista di Diritto dei Media, 3/2019 (pag. 262 – 268), Milano: https://www.medialaws.eu/wp-content/uploads/2019/11/borsari.pdf

BRANTINGHAM P. J., "*The Logic of Data Bias and its Impact on Placed-Based Predictive Policing*", in Ohio State Journal of Criminal Law, Vol. 15:473 (pag. 473 – 486), 2018: http://paleo.sscnet.ucla.edu/Brantingham-2018-OSJCL.pdf

BRANTS T., DEAN J., OCH F.J., POPAT A.C., XU P., "*Large Language Models in Machine Translation",* Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, (pag. 858 – 867), Prague, 2007, Association for Computational Linguistics: https://aclanthology.org/D07-1090.pdf

BURCHARD C., "*L'intelligenza artificiale come fine del diritto penale? Sulla trasformazione algoritmica della società*", in Rivista Italiana di Diritto e Procedura Penale – n.4 – 2019 (pag. 1909 – 1942).

BURGIO E. e DE SIMONE L., "*Intelligenza Artificiale e responsabilità civile*", 15 aprile 2021, Medialaws, Law and Policy of the Media in a Comparative Perspective: https://www.medialaws.eu/intelligenza-artificiale-e-responsabilita-civile/

CANZIO G., "*Intelligenza artificiale, algoritmi e giustizia penale*", 8 gennaio 2021, in Sistema Penale: https://www.sistemapenale.it/it/articolo/canzio-intelligenza-artificiale-algoritmi-giustizia-penale

ČAPEK K., "*R.U.R.*", 1920.

CAPPELLINI A., "*Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale",* in Criminalia – Annuario di scienze penalistiche, Edizioni ETS, 2018, disCrimen dal 27/03/2019, Pisa: https://discrimen.it/wp-content/uploads/Cappellini-Machina-delinquere-non-potest.pdf

CAPPELLINI A., "*Profili penalistici delle self-driving cars*" (pag. 325 – 353) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018. In Diritto penale Contemporaneo, Rivista trimestrale 2/2019.

CARROZZA M.C., DI MININ A., MONTEMAGNI G., ODDO C., ORVIETO S., *"AI: profili tecnologici Automazione e Autonomia: dalla definizione alle possibili applicazioni dell'Intelligenza Artificiale"* in Bio Law Journal – Rivista di Bio Diritto, n. 3/2019.

CEMBRANI F., "*Irresponsabilità penale del medico e qualità metodologica del sapere scientifico codificato medical and methodological quality of the scientific code*", in Rivista Italiana di Medicina Legale (e del Diritto in campo sanitario), fasc.2, 1 aprile 2019, pag. 645.

CHAN C. e WALKER-OSBORN C., "*Artificial Intelligence and the Law*", ITNOW, Volume 59, Issue 1 (pag. 36–37), 3 marzo 2017: https://doi.org/10.1093/itnow/bwx017

CIRACÌ F., "*Per una teoria critica del digitale: fake-news e postverità alla luce della logica della verosimiglianza*" in R. FEDRIGA, F. CIRACÌ E C. MARRAS, "*Filosofia digitale, Quaderni di filosofia*", Mimesis Edizioni (Milano – Udine), aprile 2021 Fano (Pu).

CONSULICH F., "*Il nastro di möbius. Intelligenza artificiale e imputazione penale nelle nuove forme di abuso del mercato*", in Banca Borsa Titoli di Credito, fasc.2, 1 aprile 2018, pag. 195.

CONTISSA G., LASAGNI G., SARTOR G., "*Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo"* in Pacini Giuridica, rivista trimestrale di Diritto di Internet: Digital Copyright e Data Protection n. 4/2019 (pag. 619 – 634), Pisa: https://dirittodiinternet.it/wp-content/uploads/2019/12/2_Sartor.pdf

COSTANTINI A., "*Istanze di criminalizzazione delle fake news al confine tra la tutela penale della verità e repressione del dissenso*" (pag. 60 – 80) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018. In Diritto penale Contemporaneo, Rivista trimestrale 2/2019.

COSTANZO A., "*Logica e psicologia nel ragionamento giudiziario logic and psychology in judicial reasoning*" in Cassazione Penale, fasc.6, 1 giugno 2017, pag. 2516B.

CUCCO C., "*La partita del diritto penale nell'epoca dei «drone-crimes»*" (pag. 304 – 324) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018. In Diritto penale Contemporaneo, Rivista trimestrale 2/2019.

D'AVACK L., "*La rivoluzione tecnologica e la nuova era digitale: principi etici*", in RUFFOLO U., "*Intelligenza artificiale – il diritto, i diritti, l'etica*", Collana Tech and Law di Giuffrè Francis Lefebvre Edizione Milano 2020; Parte I "Etica", Capitolo 1 (pag. 3 – 29).

DANAHER J., "*Robotic Rape and Robotic Child Sexual Abuse: Should They Be Criminalized?*", in Criminal Law & Philosophy, 2017, (pag. 71 ss).

DEL GATTO S., "*Potere algoritmico, digital welfare state e garanzie per gli amministrati. I nodi ancora da sciogliere*", in Rivista Italiana di Diritto Pubblico Comunitario, fasc.6, 1 dicembre 2020 (pag. 829).

DELL'AQUILA L., "*Il fenomeno del child grooming: l'adescamento di minori dentro e fuori la rete*", in Cyberlaws, 15 maggio 2019: https://www.cyberlaws.it/en/2019/adescamento-minori-child-grooming/

DI GIOVINE O., "*il judge-bot e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione giurisprudenziale) judge-bot and juridical sequences in criminal matters (artificial intelligence and jurisprudential stabilisation)*", in Cassazione Penale, fasc.3, 1 marzo 2020, pag. 951.

DOLCINI E., GATTA G.L., MARINUCCI G., "*Manuale di diritto penale. Parte generale"*, Ottava edizione, Giuffrè Francis Lefebvre S.p.a. Milano – 2019.

DONATI F., "*Intelligenza artificiale e giustizia*", AIC: Associazione Italiana dei Costituzionalisti, Rivista N°: 1/2020 (pag. 415 – 436), 02 marzo 2020: https://www.rivistaaic.it/images/rivista/pdf/1_2020_Donati.pdf

EASTERBROOK F.H., "*Cyberspace and the Law of the Horse*", in University of Chicago Legal Forum 207 (pag. 207 – 216), 1996: https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=2147&context=journal_articles

ERONIA O., "*Potenziamento umano e diritto penale il "caso" dell'enhancement cognitivo*", in Riv. it. dir. e proc. pen., fasc.3, 2012, pag. 975.

EVARISTI G. M., "*L'uso dello smartphone alla guida nei delitti di omicidio e lesioni colpose stradali*" (pag. 456 – 464) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018. In Diritto penale Contemporaneo, Rivista trimestrale 2/2019.

FASAN M., "*OMS – Ethics ad Governance of Artificial Intelligence for Health: WHO Guidance*", in BioDiritto rivista online – Università di Trento, Facoltà di Giurisprudenza, del 28 giugno 2021: https://www.biodiritto.org/Biolaw-pedia/Docs/OMS-Ethics-ad-Governance-of-Artificial-Intelligence-for-Health-WHO-Guidance

FIANDACA G. e MUSCO E., "*Diritto penale – Parte Speciale*" Volume II, Tomo primo "*I delitti contro la persona*", Quarta edizione, Zanichelli Editore, 1 gennaio 2013, Bologna.

FINOCCHI R., PERRI A., PEVERINI P., "*La prova dell'enunciazione. Fotografia digitale, deepfake e pertinenza documentale denegata o rinegoziata*", in EC E|C Rivista dell'Associazione Italiana di Studi Semiotici, XIV, n. 30 (pag. 135 – 144), 2020 • Mimesis Edizioni, Milano-Udine.

FLORIDI L., "*Philosophy of computing and information. 5 Questions*", in Automatic Press/VIP, p.95, Copenhagen, Danimarca, 2008.

GALGANI B., "*Considerazioni sui "precedenti" dell'imputato e del giudice al cospetto dell'IA nel processo penale"* in Sistema Penale, 4/2020 (pag. 81 – 94):

https://www.sistemapenale.it/pdf_contenuti/1586294115_galgani-2020a-precedenti-intelligenza-artificiale-processo-penale.pdf

GALUPPI G., "*Libero arbitrio, imputabilità, pericolosità sociale e trattamento penitenziario*", in Dir. famiglia, fasc.1, 2007, pag. 328.

GIALUZ M., "*Il diritto alla giurisdizione dell'imputato e della vittima tra spinte europee e carenze dell'ordinamento italiano*", in Rivista Italiana di Diritto e Procedura Penale, fasc.1, 1 marzo 2019, pag. 75.

GLESS S., SILVERMAN E., WEIGEND T., "*If robots cause harm, who is to blame? Self-driving cars and criminal liability*", in 19 New Criminal Law Review, 2016, pag. 412 e ss..

GUERINI T., "*Il formante algoritmico all'alba della Justice digital"*, 29 ottobre 2021, in DisCrimen: https://discrimen.it/wp-content/uploads/Guerini-Il-formante-algoritmico.pdf

GULLO A., LATTANZI G., SEVERINO P., "*Responsabilità da reato degli enti"*, Volume I "Diritto Sostanziale", G. Giappichelli Editore, Torino 2020.

HALLEVY G., "*The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control",* in *Akron Intellectual Property Journal*: Vol. 4: Iss. 2, Article 1 (pag. 171 – 201), 2010: https://ideaexchange.uakron.edu/cgi/viewcontent.cgi?article=1037&context=akronintellectualproperty

HOBBES T., "*Leviathan or The Matter, Forme and Power of a Common Wealth Ecclesiastical and Civil",* 1651

IAGNEMMA C., "*Discrezionalità giudiziaria e legislazione penale. Un rapporto da rivisitare nella teoria del reato e nel sistema sanzionatorio*", Rivista Italiana di Diritto e Procedura Penale, fasc.3, 1 settembre 2019, pag. 1431.

ILLUMINATI G., "*Principio di legalità e processo penale - Principle of legality and criminal process*", Cassazione Penale, fasc.10, 1 ottobre 2020, pag. 3517.

IMBRUGLIA D., "*L'intelligenza artificiale (IA) e le regole. Appunti"* in rivista quadrimestrale di Media Laws, Rivista di Diritto dei Media, 3/2020 (pag. 18 – 31), Milano: https://www.medialaws.eu/wp-content/uploads/2020/12/RDM-3-2020-Imbruglia-18-31.pdf

JHONSSON A.K., MORRIS P.H., MUSCETTOLA N., RAJAN K., SMITH B., "*Planning in Interplanetary Space: Theory and Practice",* in AIPS, 2000, https://www.aaai.org/Papers/AIPS/2000/AIPS00-019.pdf

LAVORGNA A. e SUFFIA G., "*La nuova proposta europea per regolamentare i Sistemi di Intelligenza Artificiale e la sua rilevanza nell'ambito della giustizia penale*", in Diritto Penale Contemporaneo, Rivista trimestrale 2/2021.

LIMITI C., "*Intelligenza Artificiale: implicazioni etiche in materia di privacy e diritto penale*", in Ius in itinere Redazione, pubblicato il 09/02/2021, aggiornato il 09/04/2021: https://www.iusinitinere.it/intelligenza-artificiale-implicazioni-etiche-in-materia-di-privacy-e-diritto-penale-35424

LONGO A. e SCORZA G., "*Intelligenza Artificiale: impatto sulle nostre vite, diritti e libertà",* Mondadori Education Università, Firenze, 2020.

MAGRO M. B., "*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.

MAGRO M.B., "*Decisione umana e decisione robotica un'ipotesi di responsabilità da procreazione robotica*", in La Legislazione penale, Giustizia Penale e nuove tecnologie, 2020, Dipartimento di Giurisprudenza, Università degli Studi di Torino, Torino: http://www.lalegislazionepenale.eu/wp-content/uploads/2020/05/Magro-Giustizia-penale-e-nuove-tecnologie.pdf

MANES V., "*L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*", in RUFFOLO U., "*Intelligenza artificiale – il diritto, i diritti, l'etica*", 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 6 (pag. 547 – 571).

MAUGERI A.M., "*L'uso di algoritmi predittivi per accertare la pericolosità sociale: una sfida tra evidence based practices e tutela dei diritti fondamentali*", in *Archivio penale* 2021, online (pag. 1 – 37), in https://archiviopenale.it/File/DownloadArticolo?codice=79045524-aaff-4497-88fc-be501eff1988&idarticolo=27135

MINSKY M.L., "*Semantic information processing*", Cambridge, 1969.

MONTESQUIEU, "*De l'espirit des lois*", 1748.

MULLAINATHAN S., OBERMEYER Z., POWERS B. e VOGELI C., "*Dissecting racial bias in an algorithm used to manage the health of populations*", in "Science", 25 ottobre 2019, Vol. 366, Issue 6464 (pag.447 – 453): https://www.science.org/doi/10.1126/science.aax2342

NARDI V., "*I discorsi d'odio nell'era digitale: quale ruolo per l'internet service provider?*" (pag. 268 – 288) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018, in Diritto penale Contemporaneo, Rivista trimestrale 2/2019.

ORDONSELLI N., "*«Porno Deepfake»: profili di diritto penale*", in Cyberlaws, 18 gennaio 2021: https://www.cyberlaws.it/2021/porno-deep-fake-profili-penalistici-reato/

ORLANDO S., "*I limiti della tutela penale del trattamento illecito dei dati personali nel mondo digitale*" (pag. 178 – 200) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018, in Diritto penale Contemporaneo, Rivista trimestrale 2/2019.

PAGALLO U., "*Intelligenza Artificiale e diritto. Linee guida per un oculato intervento normativo*", in Sistemi intelligenti (ISSN 1120-9550) Fascicolo 3, dicembre 2017 (pag. 615 – 636), Il Mulino – Rivisteweb, Bologna.

PALMERINI E., "*Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*", in Responsabilità Civile e Previdenza, fasc.6, 2016, pag. 1815B.

PALMISANO M., "*L'abuso di mercato nell'era delle nuove tecnologie. Trading algoritmo e principio di personalità dell'illecito penale*" (pag. 129 - 147) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018, in Diritto penale Contemporaneo, Rivista trimestrale 2/2019.

PANATTONI B., "*Gli effetti dell'automazione sui modelli di responsabilità: il caso delle piattaforme online*" (pag. 33 – 58) in "*Nuove frontiere tecnologiche e sistema penale. Sicurezza informatica, strumenti di repressione e tecniche di prevenzione*", IX Corso di formazione interdottorale di Diritto e Procedura penale "Giuliano Vassalli" per dottorandi e dottori di ricerca. AIDP Gruppo Italiano, Siracusa International Institute for Criminal Justice and Human Rights – Siracusa, 29 novembre – 1 dicembre 2018, in Diritto penale Contemporaneo, Rivista trimestrale 2/2019.

PANATTONI B., "*Intelligenza artificiale: le sfide per il diritto penale nel passaggio dall'automazione tecnologica all'autonomia artificiale*", in Diritto dell'Informazione e dell'Informatica (Il), fasc.2, 1 aprile 2021, pag. 317.

PARODI C. E SELLAROLI V., "*Sistema penale e intelligenza artificiale: molte speranze e qualche equivoco", in* Diritto Penale Contemporaneo, Fascicolo 6/2019 (pag. 47 – 71): https://archiviodpc.dirittopenaleuomo.org/pdf-viewer/?file=%2Fpdf-fascicoli%2FDPC_6_2019.pdf#page=47

PIERGALLINI C., "*Intelligenza artificiale: da 'mezzo' ad 'autore' del reato?"*, in Rivista Italiana di Diritto e Procedura Penale, fasc.4, 1 dicembre 2020, pag. 1745.

PINO G., "*Discorso razzista e libertà di manifestazione del pensiero*", in *Politica del diritto,* 11, 2008 (pag. 287 – 305).

PLATONE, "*Le Leggi*" IV secolo a.C.

PONZANELLI, "*Responsabilità per danno da computer: alcune considerazioni comparative*", in Resp. Civ. prev., 1991, 650.

PRESSACCO L., "*Nuove tecnologie e giustizia penale: il congresso annuale dell'associazione internazionale di diritto penale - gruppo italiano new technologies and criminal justice: the annual congress of the international association of penal law – italian group*", in Cassazione Penale, fasc.7, 1 luglio 2019, pag. 2718.

QUARTA E., "*Giustizia e predizione: l'algoritmo che legge il futuro*", 10 marzo 2019, Giustizia Insieme: https://www.giustiziainsieme.it/it/cultura-e-societa/600-giustizia-e-predizione-l-algoritmo-che-legge-il-futuro

QUATTROCOLO S., "*Intelligenza artificiale e giustizia: nella cornice della Carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche*", in *www.lalegislazionepenale.eu*, 18 dicembre 2018: http://www.lalegislazionepenale.eu/wp-content/uploads/2019/02/Carta-etica-LP-impaginato.pdf

QUATTROCOLO S., "*Processo penale e rivoluzione digitale: da ossimoro a endiadi?*" in rivista quadrimestrale di Media Laws, Rivista di Diritto dei Media, 3/2020, Milano: https://www.medialaws.eu/wp-content/uploads/2020/12/RDM-3-2020-Quattrocolo-121-135.pdf

QUATTROCOLO S., "*Quesiti nuovi e soluzioni antiche? Consolidati paradigmi normativi vs rischi e paure della giustizia digitale "predittiva" - New doubts and old solutions? Traditional legal concepts vs the conundrum of predictive justice*", Cassazione Penale, fasc.4, 1 aprile 2019, pag. 1748.

QUINTARELLI S., "*Intelligenza artificiale: cos'è davvero, come funziona, che effetti avrà*", Bollati Boringhieri, Gruppo Editoriale Mauri Spagnol, Torino, 2020.

RISSLAND E.L., "*Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning*", The Yale Law Journal, Vol. 99 no. 8, The Yale Law Journal Company, Inc., 1990, (pag. 1957 – 1981): https://doi.org/10.2307/796679

SALVADORI I., "*Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale*", Rivista Italiana di Diritto e Procedura Penale, fasc.1, 1 marzo 2021, pag. 83.

SANTOSUOSSO A., "*Intelligenza artificiale e diritto: perché le tecnologie di IA sono una grande opportunità per il diritto*", Scienza e Filosofia (collana diretta da A. MASSARENTI), Mondadori Education Università, Firenze, 2020.

SCHANK R.C., "*What's IA, anyway?*", in *IA Magazine*, Vol. 8 n.4, 1987, (pag. 59 – 65): https://doi.org/10.1609/aimag.v8i4.623

SCHLESINGER P., TORRENTE A., "*Manuale di diritto privato*", Ventitreesima edizione a cura di ANELLI F. e GRANELLI C., Giuffrè Editore, Milano 2017.

SEVERINO P., "*Intelligenza artificiale e diritto penale*", 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).

SIMONCINI A., "*L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*", in BioLaw Journal – Rivista di BioDiritto, n. 1/2019 - ISSN 2284-4503, (pag. 63-89).

TREVISI C., "*La regolamentazione in materia di Intelligenza artificiale, robot, automazione: a che punto siamo*", in rivista quadrimestrale di Media Laws, Law and Policy of the Media in a Comparative Perspective, Rivista di Diritto dei Media, 2/2018 (pag. 447 - 458), 25 Giugno 2018, Milano: https://www.medialaws.eu/wp-content/uploads/2019/05/RDM-2-2018.pdf

TURING A. M., "*Computing Machinery and Intelligence*," in *Mind* 59, no. 236 (pag. 433 – 460), 1950.

TUZET G., "*L'algoritmo come pastore del giudice? Diritto, tecnologie, prova scientifica*", in rivista quadrimestrale di Media Laws, Rivista di Diritto dei Media, 1/2020 (pag. 45 – 55), 16 marzo 2020, Milano: https://www.medialaws.eu/wp-content/uploads/2020/03/RDM-1-2020.pdf

UBERTIS G., "*Intelligenza artificiale, giustizia penale, controllo umano significativo*", in rivista trimestrale Diritto Penale Contemporaneo, Milano, 4/2020 (pag. 75 – 88): https://dpc-rivista-trimestrale.criminaljusticenetwork.eu/pdf/DPC_Riv_Trim_4_2020_Ubertis.pdf

VIOLI L., "*Trattamento penitenziario*" in BALDUCCI P. e MACRILLÒ A., "*Esecuzione penale e ordinamento penitenziario*", Giuffrè Francis Lefebvre S.p.a. Milano – 2020; Parte III "Il diritto penitenziario", Capitolo 1 (pag. 675 – 765).

WESTERLUND M., "*The Emergence of Deepfake Technology: A Review*", Technology Innovation Management Review, Volume 9, Issue 11 (pag. 39 – 52), November 2019: https://timreview.ca/sites/default/files/article_PDF/TIMReview_November2019%20-%20D%20-%20Final.pdf

## ***SITOGRAFIA***

ANSA REDAZIONE, "*Assistenti poco smart, Alexa propone sfida rischiosa a bimba*", Milano, 30/12/2021

ANSA, "*Tesla: Elon Musk annuncia l'arrivo del robot umanoide*", 31/08/2021: https://www.ansa.it/sito/notizie/tecnologia/hitech/2021/08/20/tesla-elon-musk-annuncia-larrivo-del-robot-umanoide_a46baca4-a623-472f-9440-2603d965fb47.html

ARTICLE19, "*Necessary & Proportionate - International Principles on the Application of Human Rights Law to Communications Surveillance – Background and Supporting International Legal Analysis*", Electronic Frontier Foundation, eff.org, Maggio 2014: https://www.ohchr.org/Documents/Issues/Privacy/ElectronicFrontierFoundation.pdf

BOWLES N., "*Google self-driving car collides with bus in California, accident report says*", in The Guardian, 1/03/2016: https://www.theguardian.com/technology/2016/feb/29/google-self-driving-car-accident-california

BUTTI L., "*Le auto guideranno da sole, ma con quali responsabilità?"*, in Il Bo Live, Università di Padova, 9 novembre 2018: https://ilbolive.unipd.it/it/news/auto-guideranno-sole-quali-responsabilita

CALIMERI F. e MARZULLO A., "*Telechirurgia, le operazioni "a distanza" saranno la norma? Ostacoli e prospettive*", in *Agendadigitale.eu*, 1 settembre 2021: https://www.agendadigitale.eu/sanita/telechirurgia-le-operazioni-a-distanza-saranno-la-norma-ostacoli-e-prospettive/

CAMPANELLI S., "*2022 anno del Tesla Bot, il robot umanoide da lavoro di Elon Musk*", Huffpost, 20/08/2021: https://www.huffingtonpost.it/entry/tesla-bot-il-robot-umanoide-metallico-di-elon-musk_it_611f881ae4b0e8ac791d153d

CAPONE E., "*Le intelligenze artificiali fra razzismo e questione etica*", in IT Italian.Tech, Parte del gruppo GEDI e la Repubblica, 06/04/2021: https://www.italian.tech/2021/04/06/news/le-intelligenze-artificiali-fra-razzismo-e-questione-etica-299491573/

CAPRINO M., "*Distrarsi alla guida diventa legale a luglio: via libera agli Adas di livello 3*", del 10 febbraio 2022, in il Sole 24 Ore: https://www.ilsole24ore.com/art/guida-autonoma-convenzione-vienna-apre-sistema-adas-AEm3x1BB

DE NARDIS L., "*L'Internet delle cose è anche l'Internet del sé*", in LUISS University Press, 30 agosto 2021: https://luissuniversitypress.it/internet-in-ogni-cosa-laura-denardis-estratto/

EDPB, European Data Protection Board, "*EDPB & EDPS call for ban on use of AI for automated recognition of human features in publicly accessible spaces, and some other uses of AI that can lead to unfair discrimination*", 21 giugno 2021: https://edpb.europa.eu/news/news/2021/edpb-edps-call-ban-use-ai-automated-recognition-human-features-publicly-accessible_en

ENCICLOPEDIA TRECCANI, definizione di "*algoritmo*": https://www.treccani.it/vocabolario/algoritmo/

ENCICLOPEDIA TRECCANI, definizione di "*dato*": https://www.treccani.it/vocabolario/dato/

HUNT E., "*Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter*", in The Guardian, 24/03/2016: https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter

IL POST, "*Facebook ha ammesso di avere sbagliato in Myanmar*", 7 novembre 2018: https://www.ilpost.it/2018/11/07/facebook-ha-ammesso-di-avere-sbagliato-in-myanmar/

ITALIANO G.F., "*Intelligenza artificiale, che errore lasciarla agli informatici"*, in *Agendadigitale.eu*, 11 giugno 2019: https://www.agendadigitale.eu/cultura-digitale/intelligenza-artificiale-che-errore-lasciarla-agli-informatici/

LANGEWIESCHE W., "*What Really Brought Down the Boeing 737 Max?*", 18 settembre 2019, aggiornato il 2 luglio 2021, in The New York Times Magazine: https://www.nytimes.com/2019/09/18/magazine/boeing-737-max-crashes.html

LEONI M., PAPARELLA M. e SOLVI S., "*Sanità 4.0, una strategia integrata e globale per la trasformazione digitale*" in *Agendadigitale.eu*, 13 febbraio 2019: https://www.agendadigitale.eu/sanita/sanita-4-0-una-strategia-integrata-e-globale-per-la-trasformazione-digitale/

LIVELLI F.M.R., "*Deepfake e revenge porn, combatterli con la cultura digitale: ecco come*", in Network Digital 360 – Cybersecurity 360, 8 febbraio 2021: https://www.cybersecurity360.it/nuove-minacce/deepfake-e-revenge-porn-combatterli-con-la-cultura-digitale-ecco-come/

LOMBARDI M., "*Come ripensare la formazione in Sanità ai tempi del digitale*" in *Agendadigitale.eu*, 20 febbraio 2017: https://www.agendadigitale.eu/infrastrutture/come-ripensare-la-formazione-in-sanita-ai-tempi-del-digitale/

MAGLIO G., "*Robotica in sanità, come vengono impiegati nell'assistenza agli anziani: vantaggi e limiti*", in *Agendadigitale.eu*, 23 aprile 2021: https://www.agendadigitale.eu/sanita/robot/

MALGIERI G. e PASQUALE F., "*L'Europa regola l'Intelligenza Artificiale ad altro rischio Lezione per gli USA*", in LUISS University Press, 2 agosto 2021: https://luissuniversitypress.it/l-europa-regola-l-intelligenza-artificiale-ad-alto-rischio/

MARINO D., MICELI A., NACCARI CARLIZZI D. e QUATTRONE G., "*Telemedicina, cos'è e come farla in Italia: tecnologie e finalità, un modello possibile*", in *Agendadigitale.eu*, 8 aprile 2020: https://www.agendadigitale.eu/sanita/telemedicina-come-farla-in-italia-le-tecnologie-le-finalita-un-modello-possibile/

MARINO N., "*Intelligenza Artificiale e medicina, serve fare formazione di medici e pazienti: ecco perché*", in *Agendadigitale.eu*, 25 marzo 2019: https://www.agendadigitale.eu/sanita/intelligenza-artificiale-e-medicina-serve-fare-formazione-di-medici-e-pazienti-ecco-perche/

META NEWSROOM "*Removing Myanmar Military Officials From Facebook*", 28 Agosto 2018, https://newsroom.fb.com/news/2018/08/removing-myanmar-officials/

MINISTERO DELLA SALUTE - Governo clinico e sicurezza delle cure, "*Sistema nazionale Linee Guida – SNLG*", 21 gennaio 2022: https://www.salute.gov.it/portale/sicurezzaCure/dettaglioContenutiSicurezzaCure.jsp?lingua=italiano&id=4835&area=qualita&menu=lineeguida

MISCHITELLI L., "*Così i robot aiutano i chirurghi a operare meglio: progressi e prospettive della tecnologia*", in *Agendadigitale.eu*, 4 giugno 2021: https://www.agendadigitale.eu/sanita/cosi-i-robot-aiutano-i-chirurghi-a-operare-meglio-progressi-e-prospettive-della-tecnologia/

MISCHITELLI L., "I pregiudizi dell'intelligenza artificiale in Sanità: perché si creano e come prevenirli", in *Agendadigitale.eu*, 23 aprile 2021: https://www.agendadigitale.eu/sanita/i-pregiudizi-dellintelligenza-artificiale-in-sanita-perche-si-creano-e-come-prevenirli/

MORUZZI M., "*Robot sanitari alla sfida autonomia: la svolta «quinta dimensione»*", in *Agendadigitale.eu*, 21 ottobre 2020: https://www.agendadigitale.eu/sanita/robot-sanitari-alla-sfida-autonomia-la-svolta-quinta-dimensione/

MOZUR P., "*A Genocide Incited on Facebook, With Posts From Myanmar's Military*", in The New York Times 16 Ottobre 2018, https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html

PASQUALE F., "*La cura degli umani*", in LUISS University Press, 3 giugno 2021: https://luissuniversitypress.it/nuove-leggi-della-robotica-frank-pasquale-estratto/

PATTARO A. F., "*Fascicolo Sanitario Elettronico, cos'è, a che serve e come attivarlo*" in *Agendadigitale.eu*, 16 settembre 2021: https://www.agendadigitale.eu/sanita/fascicolo-sanitario-elettronico-cose-e-a-che-punto-e-la-guida/

PILLON SERGIO, "*Il rischio clinico nella medicina digitale, che cos'è e perché è importante*", in *Agendadigitale.eu*, 11 aprile 2018: https://www.agendadigitale.eu/sanita/il-rischio-clinico-nella-medicina-digitale-che-cose-e-perche-e-importante/

PILLON SERGIO, "*Sanità digitale in Italia, le tendenze e le priorità del 2019*", in *Agendadigitale.eu*, 30 gennaio 2017: https://www.agendadigitale.eu/sanita/sanita-digitale-in-italia-le-tendenze-e-le-priorita-del-2019/

PORRO G., "*Il gemello digitale del corpo umano: la nuova frontiera della medicina personalizzata*" in *Agendadigitale.eu*, 15 gennaio 2022: https://www.agendadigitale.eu/sanita/il-gemello-digitale-del-corpo-umano-la-nuova-frontiera-della-medicina-personalizzata/

PRINCIPALI L. e SALERNO D., "*5G e mobilità, ecco le innovazioni per le auto (in arrivo anche in Italia)*", in *Agendadigitale.eu*, 21 novembre 2018: https://www.agendadigitale.eu/infrastrutture/5g-e-mobilita-ecco-le-innovazioni-per-le-auto-in-arrivo-anche-in-italia/

RIVA G., "*Terapie virtuali, cosa sono e quali sono i vantaggi clinici*", in *Agendadigitale.eu*, 6 ottobre 2021: https://www.agendadigitale.eu/sanita/virtual-therapeutics-cosa-sono-e-quali-sono-i-vantaggi-clinici/

ROVELLI M., "*Uber, la polizia scagiona l'auto a guida autonoma: «Incidente inevitabile*", Corriere della sera – Tecnologia/New Economy, 20 marzo 2018: https://www.corriere.it/tecnologia/economia-digitale/cards/uber-polizia-scagiona-l-auto-guida-autonoma-incidente-inevitabile/caso-uber-impatto-inevitabile_principale.shtml

RUGGIERO N., "*Auto a guida autonoma, ecco le frontiere (tra reti neurali e 5G)*", in *Agendadigitale.eu*, 20 marzo 2019: https://www.agendadigitale.eu/infrastrutture/auto-a-guida-autonoma-ecco-le-frontiere-attuali-tra-reti-neurali-e-5g/

RUGGIERO N., "*TikTok, una finestra sul futuro dei social network (e di noi stessi)*" in *Agendadigitale.eu*, 3 febbraio 2021: https://www.agendadigitale.eu/cultura-digitale/tik-tok-una-finestra-sul-futuro-dei-social-network-e-di-noi-stessi/

SANTORO E., "*L'algoritmo è un buon pediatra: così migliorano le diagnosi dell'AI*", in *Agendadigitale.eu*, 5 marzo 2019: *https://www.agendadigitale.eu/cultura-digitale/lalgoritmo-e-un-buon-pediatra-cosi-migliorano-le-diagnosi-dellai/*

SANTORO E., "*Sanità, ecco le tre innovazioni che la cambieranno quest'anno*", in *Agendadigitale.eu*, 20 febbraio 2019: https://www.agendadigitale.eu/sanita/sanita-ecco-le-tre-innovazioni-che-la-cambieranno-questanno/

SANTORO E., "*Tecnologie in Sanità, ecco tre prodotti innovativi di provata efficacia*", in *Agendadigitale.eu*, 5 febbraio 2018: https://www.agendadigitale.eu/sanita/tecnologie-sanita-tre-prodotti-innovativi-provata-efficacia/

SPARISCI D., "*Morto per un errore del pilota automatico Cos'è l'Autopilot di Tesla e cosa è successo*" Corriere della sera – Tecnologia/New Economy, 1 luglio 2016: https://www.corriere.it/tecnologia/provati-per-voi/cards/morto-un-errore-pilota-automatico-cos-l-autopilot-tesla-cosa-successo/morte-marine_principale.shtml

SU S., "*Update on Myanmar*", in Meta Newsroom, 15 Agosto 2018, https://newsroom.fb.com/news/2018/08/update-on-myanmar/

SUMAN F., "*Dilemmi morali per le auto a guida autonoma"*, in Il Bo Live, Università di Padova, 7 novembre 2018: https://ilbolive.unipd.it/it/news/dilemmi-morali-auto-guida-autonoma

USACM, Issues Statement on Algorithmic Transparency and Accountability, 12 gennaio 2017, https://www.acm.org/articles/bulletins/2017/january/usacm-statement-algorithmic-accountability

VOLZ B., "*Family calls for investigation after Tesla driver dies in Osceola crash*", in ClickOrlando.com, 26 settembre 2019: https://www.clickorlando.com/news/2019/09/26/family-calls-for-investigation-after-tesla-driver-dies-in-osceola-crash/

## ***LETTERATURA GRIGIA***

AJDER H., CAVALLI F. e PATRINI G., "*Automating Image Abuse, Deepfakes Bots on Telegram*", Sensity, Ottobre 2020 https://www.medianama.com/wp-content/uploads/Sensity-AutomatingImageAbuse.pdf

AJDER H., CAVALLI F., CULLEN e L. PATRINI G., "*The state of deepfakes landscape, threats, and impact*", Deeptrace, Settembre 2019: http://regmedia.co.uk/2019/10/08/deepfake_report.pdf

BRANCATI B. G., "*L'impatto dell'Intelligenza Artificiale (AI-Artificial Intelligence) sul ciclo di intelligence e sugli strumenti a disposizione per i pianificatori militari e le forze dell'ordine*", Lavoro di gruppo 2° sessione, a cura di B.C. RAMPONI, A.A. HUSSEIN, M. BELLADONNA, D. MARZINOTTO, T.J. MACKINTOSH, P. DELATO, M. MINENNA, P. COPPOLA, K.A. RAHMANI, H. ZAO, B.G. SCAFURI, E. ZIELLO, S. VALIO, M. DRAGONI, A. ADORNI, Centro Alti studi per la difesa, Istituto alti studi per la difesa, 71° Sessione di studio IASD, 2019-2020.

BROOKS R., BRYNJOLFSSON E., CALO R., ETZIONI O., HAGER G., HIRSCHBERG J., KALYANAKRISHNAN S., KAMAR E., KRAUS S., LEYTON-BROWN K., PARKES D., PRESS W., SAXENIAN A.L., SHAH J., STONE P., TAMBE M., TELLER A., "*Artificial intelligence and life in 2030*", Report of the 2015 Study Panel, Standford University, 2016: https://apo.org.au/sites/default/files/resource-files/2016-09/apo-nid210721.pdf

D'AGOSTINO M., DI BELLA R. e PAGANINI M., "*Intelligenza artificiale e imaging diagnostico. Implicazioni per il Tecnico sanitario di radiologia medica*", Federazione nazionale Ordini TSRM PSTRP, 8

novembre 2020: https://www.congressonazionaletsrm.it/wp-content/uploads/2020/11/Intelligenza-artificiale-e-TSRM-8-novembre-2020-FNO-TSRM-e-PSTRP.pdf

FALCINELLI D., FLOR R., MARCOLINI S., "*La giustizia penale nella reta – Le nuove sfide della società dell'informazione all'epoca di Internet*", Collana DIPLAP, Sezione Atti 2015, I Convegno Nazionale del Laboratorio Permanente di Diritto Penale Perugia, 19 settembre 2014, DIPLAP Editor (Via Fontana 28 – 20122 Milano).

FORNASARI G. e WENIN R., "*Diritto penale e modernità - Le nuove sfide fra terrorismo, sviluppo tecnologico e garanzie fondamentali*", Atti del convegno Trento, 2 e 3 ottobre 2015, quarta sessione, (pag. 309 – 439), 2017.

GALLUCCI E., "*Appunti delle lezioni di Diritto penale parte speciale a.a. 2019 – 2020*", presso l'Università Luiss Guido Carli.

HAWKING S., intervento durante la Conferenza *Zeitgeist*, Londra, maggio 2015, citazione riportata da Redazione, "*Do You Trust This Computer?"* in BASILE F., "*Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine",* in Diritto Penale Contemporaneo online, Milano, 2019: https://archiviodpc.dirittopenaleuomo.org/upload/3089-basile2019.pdf , 15 maggio 2019.

LAWLOR R. C., 1963, International Association for Artificial Intelligence, "*A Manifesto for Artificial Intelligence in the Law*", slide 12, Professor R. Susskind, 14 giugno 2017: http://www.iaail.org/?q=page/keynote-speeches-icail

MASSARO A., "*Omicidio stradale e lesioni personali stradali gravi o gravissime: da un diritto penale frammentario a un diritto penale frammentato*", in Diritto penale Contemporaneo, scritto in relazione al Convegno "Omicidio e lesioni stradali – tra esigenza di giustizia, di deterrenza e di obiettività ricostruttiva della dinamica dell'incidente stradale" svoltosi il 1 aprile 2016 presso la Sala Unità d'Italia della Corte di Appello di Roma.

MCCARTHY J., MINSKY M.L., ROCHESTER N. e SHANNON C.E., "*A proposal for the Dartmouth summer research project on artificial intelligence*", Dartmouth College, Hanover, New Hamphsire, 1955: http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

NORVIG P. e RUSSELL S. J., "*Artificial Intelligence. A Modern Approach*", 3 ed. Prentice Hall, Englewood Cliffs, N. J., 2010.

OSCE, ORGANIZATION FOR SECURITY AND CO-OPERATION IN EUROPE, *"Annual Police Experts Meeting: Artificial Intelligence and Law Enforcement: An Ally or an Adversary*?", Conferenza 23-24 September, Wien, (pag.1): https://dirittopenaleuomo.org/wp-content/uploads/2019/07/19.pdf, in rivista DPU, Diritto Penale e Uomo, Criminal Law and Human Conditions, 23 Settembre 2019: https://dirittopenaleuomo.org/segnalazioni/artificial-intelligence-and-law-enforcement-an-ally-or-an-adversary/

SABELLA P.M., "*Il delitto di diffamazione. Struttura del fatto tipico nella dimensione offline e online*", Lezione del 18 marzo 2020, Luiss Guido Carli, Dipartimento di Giurisprudenza, Cattedra di Diritto penale 2, Titolari di insegnamento – Prof.ri E. GALLUCCI E M.N. MASULLO.

SABELLA P.M., *"La manifestazione odiosa del pensiero in Internet. Responsabilità individuali e dell'Internet Service Provider*", Lezione del 23 marzo 2020, Luiss Guido Carli, Dipartimento di Giurisprudenza, Cattedra di Diritto penale 2, Titolari di insegnamento – Prof.ri E. GALLUCCI E M.N. MASULLO.

STEINHOFF J.**, "***The Automation of Automating Automation: Automation in the AI Industry*", Comunicazione al *Canadian Communication Association Annual Conference*, in *UBC* 2-6 giugno 2019.

WHO, WORLD HEALTH ORGANIZATION, "*Ethics and Governance of Artificial Intelligence for Health: WHO Guidance*", Ginevra, 2021: https://www.who.int/publications/i/item/9789240029200

## ***NORMATIVA***

"*Convenzione sulla circolazione stradale*", Conclusa a Vienna l'8 novembre 1968, Approvata dall'Assemblea federale il 15 dicembre 1978: https://fedlex.data.admin.ch/filestore/fedlex.data.admin.ch/eli/cc/1993/402_402_402/20200519/it/pdf-a/fedlex-data-admin-ch-eli-cc-1993-402_402_402-20200519-it-pdf-a.pdf

"*Convenzioni per la protezione delle vittime di guerra*", Firmate a Ginevra nel 1949: https://www.difesa.it/Il_Ministro/ONORCADUTI/Accordi_intergovernativi/Documents/Convenzione_di_Ginevra.pdf

*Carta dei Diritti Fondamentali dell'Unione Europea*, 2000: https://www.europarl.europa.eu/charter/pdf/text_it.pdf

CEDU, *Convenzione Europea dei Diritti dell'Uomo*: https://www.echr.coe.int/documents/convention_eng.pdf

*Codice della Strada*, D.Lgs. n. 285 del 30 aprile 1992 ed entrato in vigore il 1 gennaio 1993: https://www.normattiva.it/atto/caricaDettaglioAtto?atto.dataPubblicazioneGazzetta=1992-05-18&atto.codiceRedazionale=092G0306&atto.articolo.numero=0&atto.articolo.sottoArticolo=1&atto.articolo.sottoArticolo1=10&qId=8c6b76a4-7cdb-4062-834b-fcd8f66378c0&tabID=0.032171021390881016&title=lbl.dettaglioAtto

*Codice di Procedura Penale*, D.P.R. 22 settembre 1988, n. 477.

*Codice Penale*, R.D. 19 ottobre 1930, n. 1398.

COMMISSIONE EUROPEA "*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni – Digitalizzazione della giustizia*

*dell'Unione Europea, Un pacchetto di opportunità*", Bruxelles 2.12.2020, COM(2020) 710 final: https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52020DC0710&from=EN

COMMISSIONE EUROPEA PER L'EFFICIENZA DELLA GIUSTIZIA (CEPEJ), "*Carta etica per l'uso dell'intelligenza artificiale nei sistemi giudiziari e nel loro ambiente*", App. III, "Glossario", p. 47.

COMMISSIONE EUROPEA, "*Comunicazione della Commissione - Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni - L'intelligenza artificiale per l'Europa*", Bruxelles, 25.4.2018, COM(2018) 237 final.

COMMISSIONE EUROPEA, "*Comunicazione della Commissione - Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni - Piano coordinato sull'intelligenza artificiale*" Bruxelles, 7.12.2018, COM(2018) 795 final.

COMMISSIONE EUROPEA, "*Comunicazione della Commissione - Europa 2020: Una strategia per una crescita intelligente, sostenibile e inclusiva*", COM(2010) 2020 definitivo, Bruxelles, 3.3.2010: http://publications.europa.eu/resource/cellar/6a915e39-0aab-491c-8881-147ec91fe88a.0008.02/DOC_1

COMMISSIONE EUROPEA, "*Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni - Un'agenda digitale europea*", COM(2010)245, Bruxelles, 19.5.2010: https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52010DC0245&from=IT

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

CONSIGLIO D'EUROPA, "*Convenzione per la protezione dei Diritti dell'Uomo e della dignità dell'essere umano nei confronti dell'applicazioni della biologia e della medicina: Convenzione sui Diritti dell'Uomo e la biomedicina*", Oviedo, 4 aprile 1997.

CONSIGLIO D'EUROPA, "*Convenzione sulla protezione dei minori dallo sfruttamento e dagli abusi sessuali*", conclusa a Lanzarote il 25 ottobre 2007: https://fedlex.data.admin.ch/filestore/fedlex.data.admin.ch/eli/cc/2014/249/20150826/it/pdf-a/fedlex-data-admin-ch-eli-cc-2014-249-20150826-it-pdf-a.pdf

CONSIGLIO DELL'UNIONE EUROPEA, "*Piano d'azione 2019-2023 in materia di giustizia elettronica*", Bruxelles 31 ottobre 2018 http://data.consilium.europa.eu/doc/document/ST-11724-2018-REV-4/it/pdf

CONSIGLIO DELL'UNIONE EUROPEA, "*Progetto UE di strategia in materia di giustizia elettronica (2019-2023)*", Bruxelles 31 ottobre 2018: https://data.consilium.europa.eu/doc/document/ST-12794-2018-REV-3/it/pdf

*Costituzione della Repubblica Italiana*, Roma, 27 dicembre 1947 entrata in vigore il 1 gennaio 1948.

COUNCIL OF EUROPE, "*Convention on Cybercrime*", Budapest, 23.11.2001: https://rm.coe.int/1680081561

COUNCIL OF EUROPE, "*European Convention on the Legal Protection of Services based on, or consisting of, Conditional Access*", Strasbourg, 24.I.2001: https://rm.coe.int/1680080623

D.l. 161/2019 (conv. con mod. in l. 7/2020), "*Modifiche urgenti alla disciplina delle intercettazioni di conversazioni o comunicazioni*": https://www.gazzettaufficiale.it/eli/id/2019/12/31/19G00169/sg

D.l. 28/2020 (conv. con mod. in l. 70/2020), Testo del decreto-legge 30 aprile 2020, n. 28 (in Gazzetta Ufficiale - Serie generale - n. 111 del 30 aprile 2020), coordinato con la legge di conversione 25 giugno 2020, n. 70 (in questa stessa Gazzetta Ufficiale alla pag. 1), recante"*Misure urgenti per la funzionalità dei sistemi di intercettazioni di conversazioni e comunicazioni, ulteriori misure urgenti in materia di ordinamento penitenziario, nonché disposizioni integrative e di coordinamento in materia di giustizia civile, amministrativa e contabile e misure urgenti per l'introduzione del sistema di allerta Covid-19*": https://www.gazzettaufficiale.it/eli/id/2020/06/29/20A03469/sg

D.lgs. 24 febbraio 1997, n. 46 emendato col D.lgs. 25 gennaio 2010, n. 37 – Recepimento Direttiva 2007/47/CE, "*Attuazione della Direttiva 93/42/CEE del 14 giugno 1993 concernente i dispositivi medici*".

D.lgs. 51/201840; "*Attuazione della direttiva (UE) 2016/680 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, nonché alla libera circolazione di tali dati e che abroga la decisione quadro 2008/977/GAI del Consiglio*": https://www.gazzettaufficiale.it/eli/id/2018/05/24/18G00080/sg

D.lgs. 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.

D.M. Infrastrutture e trasporti 28/02/18, G.U. 18/04/18 n. 90 sulle "*Modalità attuative e strumenti operativi della sperimentazione su strada delle soluzioni di Smart Road e di guida connessa e automatica*" (18A02619): http://www.prefettura.it/FILES/AllegatiPag/1173/Decreto_Smart_Road_registrato_Prot._n._70_del_28_febbraio_2018.pdf

Decisione Quadro del Consiglio (2001/413/GAI) "*relativa alla lotta contro le frodi e le falsificazioni di mezzi di pagamento diversi dai contanti*" del 28 maggio 2001: https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32001F0413&from=IT

Decreto-legge n. 158 del 13 settembre 2012 (c.d. Decreto Balduzzi), convertito in legge n. 189/2012 l'8 novembre 2012.

Direttiva 2013/40/UE del Parlamento Europeo e del Consiglio, "*relativa agli attacchi contro i sistemi di informazione e che sostituisce la decisione quadro 2005/222/GAI del Consiglio*", del 12 agosto 2013: https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32013L0040&from=SK

Direttiva 2014/62/UE del Parlamento Europeo e del Consiglio, sulla "*protezione mediante il diritto penale dell'euro e di altre monete contro la falsificazione e che sostituisce la decisione quadro 2000/383/GAI del Consiglio*", del 15 maggio 2014: https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32014L0062&from=EN

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

Direttiva 91/250/CEE del Consiglio, del 14 maggio 1991, relativa alla "*tutela giuridica dei programmi per elaboratore*": https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:31991L0250&from=MT

*Disposizioni di attuazione al Codice di Procedura Penale*, D.lgs. 28 luglio 1989, n. 271.

EUROPEAN COMMISSION, "*Digital Justice Ministerial Forum*", online, Belgio, del 12 ottobre 2021: https://digital-justice-ministerial-forum-2021.b2match.io/home

EUROPEAN COMMISSION, "White Paper on Artificial Intelligence - A European approach to excellence and trust", Brussels, 2020: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

EUROPEAN COMMISSION'S HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, "*Ethics Guidelines for Trustworthy AI"*, 2019.

IV Convenzione dell'Aja del 1907 concernente "*le leggi e gli usi della guerra per terra*", Conclusa all'Aja il 18 ottobre 1907: https://www.difesa.it/SMD_/CASD/IM/ISSMI/Corsi/Corso_Consigliere_Giuridico/Documents/65159_convenzione4.pdf

Legge 10 ottobre 1986, n. 663, "*Modifiche alla legge sull'ordinamento penitenziario e sulla esecuzione delle misure privative e limitative della libertà.*

Legge 13 ottobre 1975, n. 654, "*Ratifica ed esecuzione della convenzione internazionale sull'eliminazione di tutte le forme di discriminazione razziale, aperta alla firma a New York il 7 marzo 1966*": https://www.gazzettaufficiale.it/eli/id/1975/12/23/075U0654/sg

Legge 18 marzo 2008, n. 48, "*Ratifica ed esecuzione della Convenzione del Consiglio d'Europa sulla criminalità informatica, fatta a Budapest il 23 novembre 2001, e norme di adeguamento dell'ordinamento interno*", pubblicata nella *Gazzetta Ufficiale* n. 80 del 4 aprile 2008 - Supplemento ordinario n. 79: https://www.parlamento.it/parlam/leggi/080481.htm

Legge 19 luglio 2019, n. 69, "*Modifiche al codice penale, al codice di procedura penale e altre disposizioni in materia di tutela delle vittime di violenza domestica e di genere*" (GU n.173 del 25-7-2019).

Legge 23 aprile 2009, n. 38, "*Conversione in legge, con modificazioni, del decreto-legge 23 febbraio 2009, n. 11, recante misure urgenti in materia di sicurezza pubblica e di contrasto alla violenza sessuale, nonché*

*in tema di atti persecutori*", pubblicata nella *Gazzetta Ufficiale* n. 95 del 24 aprile 2009: https://web.camera.it/parlam/leggi/09038l.htm

Legge 23 marzo 2016, n. 41, "*Introduzione del reato di omicidio stradale e del reato di lesioni personali stradali, nonché disposizioni di coordinamento al decreto legislativo 30 aprile 1992, n. 285, e al decreto legislativo 28 agosto 2000, n. 274 (16G00048)*", entrata in vigore il 25/03/2016 (GU. N.70 del 24/03/2016): https://www.gazzettaufficiale.it/eli/id/2016/03/24/16G00048/sg

Legge 24 febbraio 2006, n. 85, "*Modifiche al codice penale in materia di reati di opinione*", pubblicata nella *Gazzetta Ufficiale* n. 60 del 13 marzo 2006: https://web.camera.it/parlam/leggi/060851.htm

Legge 25 giugno 1993, n. 205 "*Conversione in legge, con modificazioni, del decreto-legge 26 aprile 1993, n. 122, recante misure urgenti in materia di discriminazione razziale, etnica e religiosa*", entrata in vigore il 27 giugno 1993: https://www.gazzettaufficiale.it/eli/id/1993/06/26/093G0275/sg

Legge n.24/2017 dell'8 marzo 2017: "*Disposizioni in materia di sicurezza delle cure e della persona assistita, nonché in materia di responsabilità professionale degli esercenti le professioni sanitarie*".

PARLAMENTO EUROPEO E DEL CONSIGLIO, Regolamento (UE) 2016/679 del 27 aprile 2016 relativo alla "*protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE*" (Regolamento generale sulla protezione dei dati): https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32016R0679&from=IT

PARLAMENTO EUROPEO, "*Proposta di Risoluzione del Parlamento Europeo recante le raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*", 27/01/2017: https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_IT.html?redirect

PARLAMENTO EUROPEO, Risoluzione del 16 febbraio 2017 recante "*raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*" (2015/2103(INL)), Gazzetta ufficiale dell'Unione Europea: https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=CELEX%3A52017IP0051

*Trattato sul Funzionamento dell'Unione Europea*: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:12012E/TXT:it:PDF

## ***GIURISPRUDENZA***

BRANDSTETTER V. AUSTRIA, CEDU.

DAUBERT V. MERRELL DOW PHARMACEUTICALS, INC., 113 S.Ct. 2786(1993).

MALENCHIK V. STATE, 928 N.E.2d 564, 574 (Ind. 2010).

Sentenza della Cassazione Penale, Sez. II, n. 21987 del 20 maggio 2019.

Sentenza della Cassazione Penale, Sez. III, n. 15757 del 9 aprile 2018.

Sentenza della Cassazione Penale, Sez. III, n. 22265 del 13 gennaio 2017.

Sentenza della Cassazione Penale, Sez. III, n. 36906 del 14 settembre 2015.

Sentenza della Cassazione Penale, Sez. III, n. 37581 del 7 maggio 2008.

Sentenza della Cassazione penale, Sez. IV, n. 10643 del 14 novembre 1996.

Sentenza della Cassazione penale, Sez. IV, n. 17505 del 19 marzo 2008.

Sentenza della Cassazione penale, Sez. IV, n. 29721 del 14 giugno 2017.

Sentenza della Cassazione Penale, Sez. IV, n. 43786 del 17 settembre 2010 (Cozzini).

Sentenza della Cassazione Penale, Sez. V, n. 54946 del 27 dicembre 2016.

Sentenza della Cassazione Penale, Sez. V, n. 61 del 2 gennaio 2019.

Sentenza della Cassazione Penale, Sezione III del 30 gennaio 2006, n. 3615, (c.d Caso Jolly Mediterraneo s.r.l.).

Sentenza della Cassazione Penale, Sezioni Unite, n. 2437 del 18 dicembre 2008.

Sentenza della Cassazione Penale, Sezioni Unite, n. 30328 del 11 settembre 2002 (Franzese): https://www.giurisprudenzapenale.com/wp-content/uploads/2014/01/Cass-Pen-Sez-Un-Franzese-2002.pdf

Sentenza della Cassazione Penale, Sezioni Unite, n. 38343 del 18 settembre 2014: 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

Sentenza della Cassazione Penale, Sezioni Unite, n. 8770 del 22 febbraio 2018: https://www.biodiritto.org/ocmultibinary/download/3259/31842/8/ba374071a4619dfb28edf5d0587fb316/file/cass-pen-sez-un-2018-8770.pdf

Sentenza della Corte Costituzionale n. 242 del 25 settembre 2019.

Sentenza della Corte Costituzionale n. 247, del 16 maggio 1989: https://www.giurcost.org/decisioni/1989/0247s-89.html

Sentenza della Corte Costituzionale n. 364, del 23 marzo 1988: https://www.cortecostituzionale.it/actionSchedaPronuncia.do?anno=1988&numero=364

Sentenza della Corte Costituzionale n. 42, del 13 maggio 1965: https://www.giurcost.org/decisioni/1965/0042s-65.html

Sentenza della Corte Costituzionale n. 87 del 22 giugno 1966.

Sentenza della Corte Costituzionale n.172 del 11 giugno 2014.

Sentenza della Corte Europea dei Diritti dell'Uomo, Grande Camera, caso Engel e altri c. Paesi Bassi, 8 giugno 1976.

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

STATE V. LOOMIS, 881 N.W.2d 749 (2016) 754 (USA).

## ***FILMOGRAFIA***

CAMERON J., "*Terminator*", 1984.

JONZE S., "*Her*", 2014.

KUBRICK S., "*2001: Odissea nello spazio*", 1968.

NICCOL A., "*Gattaca – La porta dell'universo*", 1997.

WACHOWSKI A. e WACHOWSKI L., "*Matrix*", serie di tre film, 1999 – 2003.