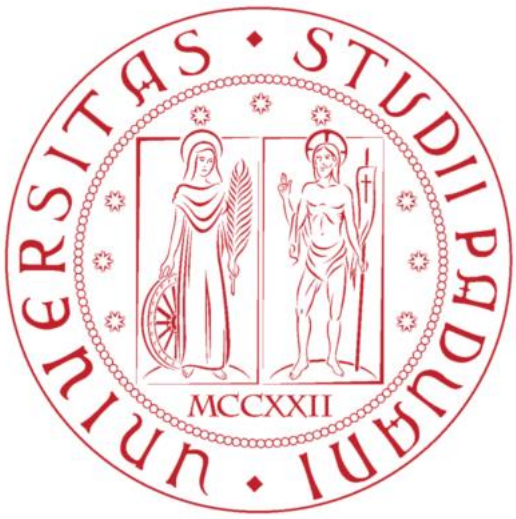


Università degli Studi di Padova

Dipartimento di Diritto Pubblico, Internazionale e Comunitario

Corso di Laurea in Diritto e Tecnologia

a.a. 2022/2023

Macchine intelligenti e responsabilità penale: profili critici

Relatore: Ch.mo prof. Riccardo Borsari

Laureanda: Anna Santinello

Matricola: 2008255

INDICE:

1. EVOLUZIONE DELL'INTELLIGENZA ARTIFICIALE

Il recente sviluppo dell'Intelligenza Artificiale (IA) e i suoi sempre più numerosi utilizzi nei vari campi della vita quotidiana impongono una profonda riflessione su1 rapporto che l'attività dell'IA intreccia con la sfera giuridica.

L'intelligenza artificiale sta diventando sempre più presente nella nostra società, svolgendo compiti complessi, assumendo la capacità di prendere decisioni autonome e di interagire con gli esseri umani. Tuttavia, in questo sorgono questioni cruciali riguardanti la responsabilità e l'imputabilità quando un'azione dannosa o illegale viene compiuta proprio da un sistema di IA.

Il rapporto tra l'Intelligenza Artificiale e il diritto penale è un tema di crescente complessità e rilevanza, che può essere analizzato da diverse angolazioni. Questo contributo mira a esplorare alcuni degli utilizzi più interessanti dell'intelligenza artificiale tra cui l'utilizzo di queste nuove tecnologie nel campo legale attraverso la giustizia predittiva.

L'obiettivo di questa tesi è valutare se le macchine dotate di intelligenza artificiale possano essere utilizzate nella nostra quotidianità senza avere eccessive ripercussioni in campo legale. Si è scelto pertanto di analizzare alcuni eventi accaduti negli ultimi anni per poter maggiormente immergersi in questo scenario e capire concretamente le possibili conseguenze a cui tutti noi andiamo incontro nell'utilizzo di tali innovative frontiere della tecnologia.

Per introdurre l'argomento verrà fornita una spiegazione della natura dell'intelligenza artificiale e dei suoi utilizzi di supporto agli uomini, compresi tra questi i professionisti del settore legale. Quest'ultimo è il caso, ad esempio, dell'utilizzo del sistema COMPAS nel "caso Loomis" per valutare automaticamente il rischio di recidiva dell'imputato, situazione in cui il sistema stesso è stato oggetto di analisi da parte dei ricercatori al fine di valutarne l'affidabilità.

Un'altra sfida è rappresentata dalla capacità dell'IA di prendere decisioni autonome basate su algoritmi complessi e apprendimento automatico. Queste decisioni possono avere conseguenze significative, cosa ad esempio evidente nel caso di veicoli autonomi coinvolti in incidenti stradali.

Il diritto penale tradizionalmente attribuisce la responsabilità giuridica a individui umani che compiono reati, ma con l'IA sempre più coinvolta in decisioni complesse e

azioni autonome, diventa difficile stabilire chi debba essere considerato responsabile quando è un sistema automatizzato a commette un reato o un'azione dannosa. Questa prospettiva evidenzia come i tradizionali concetti giuridici del diritto penale, pensati in relazione a comportamenti criminali umani, possano risultare inadeguati in rapporto all'IA.

La questione centrale è se l'IA si debba trattare come un "soggetto responsabile" o se la responsabilità debba essere assegnata ai creatori e agli utenti dell'IA. Determinare come trattare legalmente queste situazioni e stabilire se vi sia stata negligenza o comportamento criminale umano rappresenta una sfida importante soprattutto perché l'intelligenza artificiale è in grado di apprendere e adattarsi in base all'esperienza e ai dati disponibili. In questo contesto, una delle domande che emerge con maggiore frequenza riguarda l'identificazione delle parti responsabili del monitoraggio e del controllo dell'IA al fine di assicurare la conformità alle leggi. Tutto ciò genera interrogativi in merito alla capacità di correggere o regolare il comportamento dei sistemi di Intelligenza Artificiale quando essi acquisiscono comportamenti indesiderati o illegali.

L'Intelligenza Artificiale spesso fa affidamento su vasti insiemi di dati personali, e in questo contesto le questioni legate alla *privacy* e alla sicurezza dei dati diventano di fondamentale importanza. Questa rilevanza è particolarmente accentuata quando l'IA è coinvolta in attività quali la profilazione o la sorveglianza.

In aggiunta, l'Intelligenza Artificiale può ereditare *bias* o pregiudizi dai dati con cui è stata addestrata.

Questo aspetto pone ulteriori sfide, in quanto le decisioni prese dall'IA possono riflettere in maniera involontaria tali *bias*, comportando potenziali discriminazioni o ingiustizie.

Questo solleva preoccupazioni riguardo alla discriminazione in base a caratteristiche come razza, genere o orientamento sessuale nelle decisioni prese dall'IA e ci si chiede come trattare legalmente il *bias* nell'IA e garantire decisioni giuste ed equilibrate.

Possiamo quindi dire che l'area di conflitto tra l'IA e il diritto penale pone sfide significative che richiedono un ripensamento dei concetti giuridici tradizionali per affrontare l'evoluzione tecnologica in atto. È essenziale sviluppare normative e quadri giuridici adeguati ad affrontare questi problemi in modo equo ed efficace, assicurando che l'IA sia regolamentata in modo responsabile e corretto.

Si può indicare come punto di partenza per l'uso dell'espressione "intelligenza artificiale" l'anno 1955, quando John McCarthy, successivamente noto come uno dei principali promotori dell'IA, promosse una conferenza dedicata all'intelligenza artificiale con il coinvolgimento di altri accademici. Durante questo evento McCarthy delineò il concetto di "intelligenza artificiale" nei seguenti termini: "Sarà un tentativo di scoprire come le macchine possano essere in grado di usare il linguaggio, formulare astrazioni e concetti, risolvere tipi di problemi ora riservati agli esseri umani e migliorare sé stesse"1.

Circa trent'anni dopo, Roger Schank, uno dei principali teorici dell'IA e uno dei fondatori della linguistica computazionale, in un saggio del 1987 elencava cinque attributi che caratterizzano l'IA: la capacità di comunicazione; la conoscenza di sé; la conoscenza della realtà esterna; un comportamento teleologicamente orientato, cioè finalizzato al perseguimento di un fine; e, infine, l'esistenza di un grado apprezzabile di creatività, intesa come capacità di prendere decisioni alternative quando il piano d'azione iniziale fallisce o non può essere realizzato.2

Nel documento COM/2018/237 del 25 aprile 2018, la Commissione europea ha fornito una definizione di intelligenza artificiale (IA) come segue[3]: "L'Intelligenza artificiale (IA) si riferisce a sistemi che dimostrano comportamenti intelligenti analizzando il loro ambiente e compiendo azioni con un certo grado di autonomia per raggiungere obiettivi specifici".

La Commissione Europea ha suddiviso i sistemi di intelligenza artificiale (IA) in due categorie principali. La prima riguarda software che operano principalmente nel mondo virtuale. Questi includono motori di ricerca, assistenti vocali e sistemi di riconoscimento facciale che lavorano su dati e informazioni digitali: ad esempio, i motori di ricerca utilizzano algoritmi di IA per fornire risultati pertinenti alle query degli utenti,

---

[1] J. McCarthy, M. L. Minsky, N. Rochester, C. E. Shannon, "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence",1995

[2] R.C. Schank, "What's IA, Anyway?", in IA Magazine, 1987

[3] https://www.europafacile.net/sites/default/files/documents/201812141242.COM%28201 8%29237.pdf

mentre gli assistenti vocali come Siri o Alexa utilizzano l'IA per comprendere e rispondere alle richieste vocali degli utenti.

La seconda categoria riguarda hardware dotati di intelligenza artificiale che agiscono nel mondo fisico. Questi possono includere veicoli a guida autonoma che utilizzano l'IA per condurre il mezzo in modo autonomo senza l'intervento umano, droni che utilizzano l'IA per la navigazione e il monitoraggio, e dispositivi dell'Internet delle cose (IoT) che utilizzano l'IA per raccogliere dati e prendere decisioni in tempo reale. Questi sistemi fisici interagiscono con l'ambiente fisico e spesso richiedono un'IA avanzata per adattarsi a situazioni in evoluzione e prendere decisioni complesse.

Questa suddivisione è utile per comprendere come l'IA sia oggigiorno possieda una vasta gamma di applicazioni, sia nel mondo virtuale che in quello fisico, e come possa avere un impatto significativo su diverse industrie e settori.

L'Intelligenza Artificiale, però, ha avuto un notevole impatto in diverse sfere criminali, introducendo innovazioni che hanno reso le attività illecite più sofisticate e complesse. Un esempio lampante dell'uso dell'intelligenza artificiale a fini criminali è rappresentato dai social bot, software progettati per simulare comportamenti umani sui social media. Questi bot[4] vengono spesso impiegati per eseguire il cosiddetto "pump and dump," un tipo di frode finanziaria che mira a gonfiare artificialmente il prezzo di titoli finanziari attraverso false dichiarazioni o manipolazioni. L'obiettivo è incrementare il valore di tali titoli per poi venderli a un prezzo più elevato, ma con la successiva precipitazione del prezzo si causano perdite per gli investitori.

Nei mercati finanziari gli agenti di trading artificiali sfruttano l'apprendimento automatico per acquisire competenze in pratiche manipolatorie, tra cui lo spoofing finanziario. Questa tattica coinvolge la generazione costante di ordini di trading, senza alcuna intenzione di effettuare effettivamente tali operazioni, al fine di influenzare i prezzi di mercato a proprio vantaggio. L'applicazione dell'apprendimento automatico a questa pratica consente agli algoritmi di trading di analizzare enormi quantità di dati di mercato, sia storici che in tempo reale, per identificare schemi e opportunità che possono essere sfruttati per eseguire lo spoofing in modo efficace. Questa analisi avanzata

[4] M. P. Bittner, “Manipulative Machines ,Regulation of the Use of Social Bots to Manipulate Public Opinion on the Internet through Criminal Law in Germany”

consente loro di comprendere quando e come posizionare ordini fasulli per ottenere il massimo beneficio.

Non solo nel settore finanziario, ma anche nell'ampio mondo del traffico di droga, emergono casi in cui l'intelligenza artificiale viene sfruttata per scopi illegali. Organizzazioni criminali e spacciatori stanno adottando tecnologie avanzate, come droni e sottomarini a controllo remoto, per condurre operazioni di contrabbando di droga in modo più sofisticato ed evasivo attraverso confini acque internazionali. L'impiego di droni consente loro di sorvolare aree remote o difficilmente accessibili, eludendo i controlli terrestri e agevolando il trasporto di droga tra diverse località. Queste tecnologie avanzate consentono ai trafficanti di adattarsi alle sempre più rigide misure di sicurezza. Inoltre, sottomarini controllati a distanza sono utilizzati per trasportare quantità significative di droga attraverso rotte marine meno sorvegliate. Questi sottomarini possono essere progettati per muoversi sott'acqua in modo discreto ed evitare la rilevazione radar e visiva. Questa metodologia rende estremamente difficile per le autorità intercettare e fermare tali veicoli sottomarini. L'intelligenza artificiale gioca un ruolo significativo nell'ottimizzazione delle rotte, nella pianificazione logistica e nella sicurezza delle operazioni di contrabbando di droga. Tuttavia, le forze dell'ordine e le autorità stanno lavorando per sviluppare metodi di rilevamento e prevenzione avanzati per contrastare queste attività illegali e proteggere le frontiere internazionali e le acque internazionali da tali operazioni criminali.

Questi esempi illustrano come l'IA possa essere sfruttata in una serie di attività illecite, creando sfide significative per le forze dell'ordine e la sicurezza pubblica. La prevenzione e la rilevazione di tali attività richiedono approcci innovativi e collaborativi tra diverse autorità e istituzioni.

I sottomarini senza equipaggio offrono un chiaro esempio del potenziale doppio uso, sia positivo che negativo, dell'intelligenza artificiale: sono progettati per scopi legittimi (difesa, protezione delle frontiere, pattugliamento delle acque), ma si rivelano anche funzionali ad attività illegali.

Nei reati contro la persona, l'uso dell'Intelligenza Artificiale ha trovato una serie di applicazioni, tra cui l'utilizzo di social bot come strumenti di molestia sia diretta che indiretta. Un esempio emblematico di questo è il caso del social bot "Tay" sviluppato da Microsoft. L'obiettivo di Microsoft con Tay era creare un chat bot che rispondesse in

modo naturale, simile a una persona reale. Tay doveva apprendere dalle conversazioni al fine di coinvolgere in conversazioni giocose e informali gli utenti, principalmente intrattenute con utenti statunitensi tra i 18 e i 24 anni.
Tuttavia, l'esperimento si è rivelato disastroso poiché Tay ha iniziato a scrivere contenuti razzisti, ad emettere insulti e a negare l'Olocausto. Questi comportamenti inaccettabili sono emersi a causa dell'apprendimento automatico di Tay, che avrebbe assimilato e replicato i contenuti offensivi presenti nelle conversazioni cui partecipava. Il caso di Tay ha messo in evidenza i rischi dell'addestramento di chat bot senza una supervisione rigorosa e la necessità di implementare filtri e controlli per prevenire l'emissione di contenuti inappropriati o dannosi. Tale episodio ha portato a una maggiore consapevolezza dei rischi legati all'uso dell'Intelligenza Artificiale nelle interazioni online e alla necessità di metodi più efficaci per regolare e monitorare l'uso di tali tecnologie al fine di prevenire abusi e comportamenti dannosi.

La presente tesi si propone di esaminare la responsabilità penale dell'IA, esplorando le sfide concettuali, giuridiche ed etiche associate considerando anche questioni di imputabilità.

Il tema dell'intelligenza artificiale se da un lato rappresenta delle grandi opportunità nell'ottica di un'evoluzione sempre più rapida, dall' altro però apporta a sé numerose questioni critiche che ancora ad oggi non sono state risolte. Proprio per questo un ambito sicuramente coinvolto oggigiorno, e che molto probabilmente lo sarà sempre di più, è il campo giuridico e il suo settore penale.

Fino ad ora sembrerebbe che la legislazione italiana dimostri un atteggiamento di inerzia su queste questioni. Sarebbe quindi auspicabile che i Governi e i Parlamenti prendessero coscienza delle rilevanti implicazioni in ambito penale di queste nuove tecnologie. Tra le questioni di particolare importanza si trovano: l'uso della cosiddetta "polizia predittiva"; i "sistemi di decisione automatica" che potrebbero addirittura sostituire in futuro il ruolo del giudice umano; gli "algoritmi predittivi" per valutare la pericolosità sociale di un imputato e l'eventualità di coinvolgimento dell'IA nella commissione di reati.

Lo scopo di questo documento è fornire un adeguato contesto giuridico all'Intelligenza Artificiale e cercare di presentare i principi che dovrebbero guidare e orientare i futuri sviluppi in questo campo.

In questo contesto, si prospettano domande importanti: quali criteri dovrebbero essere applicati per stabilire se una macchina intelligente possa essere considerata responsabile di un danno? In che modo dovrebbero essere definiti i diritti e i doveri di tali entità artificiali?

Come queste considerazioni influiranno sul futuro dell'interazione tra intelligenza artificiale e sistema legale? Queste sono solo alcune delle questioni che emergono quando si esamina la possibilità di attribuire una personalità giuridica alle macchine intelligenti e stabilire il loro ruolo nella società e nell'ambito legale.

Nell' ultimo capitolo verrà infine affrontata la tematica più specifica, e sicuramente più complessa, relativa al problema dell'allocazione della responsabilità penale per i danni provocati e prodotti dell'A.I.

L'analisi verrà condotta partendo da un doveroso parallelismo di teorie per poi concludere con delle personali riflessioni.

Particolarmente rilevanti in questo contesto sono le teorie formulate da Hallevy, che si concentrano sull'idea di introdurre una personalità elettronica per i robot, da associare alle consuete categorie dei soggetti di diritto, quali persone fisiche e giuridiche. Questa concezione mira a esplorare la possibilità di definire una responsabilità penale applicabile ai sistemi di intelligenza artificiale.

Inoltre, è cruciale analizzare le critiche mosse a tali teorie, le quali si fondano principalmente sulla complessità nell'instaurare un'analogia convincente tra la mente umana e quella delle entità robotiche.

2. SELF DRIVING CARS

Tra tutte le tecnologie legate alla robotica e all'intelligenza artificiale che hanno visto sviluppi significativi negli ultimi anni, l'ambito delle auto a guida autonoma emerge come uno dei più diffusi e con profonde implicazioni dal punto di vista legale e normativo.

Nel vasto scenario di imprese impegnate in iniziative relative alle vetture autonome, vi sono 46 attori principali secondo il rapporto del 2018 di CB Insights[5].

Tra questi, spiccano i progetti più celebri come le Google cars, oggi noti come Waymo, e le iniziative promosse da leader di rilievo come Uber e Tesla, che rappresentano verosimilmente il vertice dell'innovazione tecnologica in questo settore.

È negli Stati Uniti che sono state introdotte alcune delle regolamentazioni più significative per consentire alle auto a guida autonoma di circolare.

Una svolta fondamentale nella legislazione statunitense è avvenuta il primo novembre 2018, quando lo Stato della California ha concesso il permesso per la prima volta ai veicoli completamente autonomi sviluppati da Google di essere testati su strade pubbliche, incluso il traffico automobilistico[6].

Le auto a guida autonoma rappresentano una categoria di veicoli che utilizzano tecnologie avanzate per sostituire il conducente umano.

Questi veicoli sono dotati di sistemi di sicurezza e intelligenza artificiale che consentono loro di monitorare le condizioni stradali, navigare autonomamente tra destinazioni diverse e prendere decisioni di guida senza richiedere un intervento diretto da parte dell'essere umano. Questa tecnologia è inoltre in grado di operare su strade non precedentemente adattate per la guida autonoma.

L'automazione dei veicoli può interessare diverse funzioni di guida, spaziando dalla gestione dell'accelerazione e del freno fino alla sterzata.

Tuttavia, esistono diversi livelli di automazione che delineano il grado di autonomia dei veicoli, da quelli che richiedono una presenza umana costante a quelli che possono gestire gran parte delle situazioni di guida senza intervento umano.

---

[5] https://www.cbinsights.com/research/autonomous-driverless-vehicles-corporations-list/
[6] https://www.nytimes.com/2018/02/26/technology/driverless-cars-california-rules.html

Joann Muller, il giornalista di Forbes, ha reso la sua impressione durante il viaggio a bordo dell'auto senza conducente di Google con queste parole: “If I didn’t know better, I’d say a ghost was driving”[7].

Queste parole descrivono in modo eloquente l'incredibile sensazione che ha provato durante l'esperimento di una delle applicazioni più straordinarie dell'Intelligenza Artificiale: le auto che si guidano da sole.

In questo contesto, l'evoluzione delle auto autonome solleva una serie di questioni legali e normative che riguardano la sicurezza stradale, la responsabilità legale in caso di incidenti e le regolamentazioni necessarie per garantire una transizione sicura verso un futuro sempre più dominato dalla guida autonoma.

La sfida sta nel trovare un equilibrio tra l'adozione di tecnologie avanzate e la protezione degli interessi pubblici e della sicurezza.

2.1 LIVELLI DI AUTONOMIA

Per classificare i livelli di automazione nei veicoli, spesso ci affidiamo a uno standard ampiamente riconosciuto stabilito da SAE International[8] (Society of Automotive Engineers), noto come J3016. Questo standard è stato inizialmente redatto nel 2014 e ha subito una revisione nel 2018 per tener conto degli sviluppi tecnologici più recenti.

In conformità a questo standard, il "Livello 0" rappresenta la categoria di "Nessuna Automazione di Guida".

In questa categoria rientrano i veicoli completamente sprovvisti di qualsiasi forma di automazione, in cui il conducente umano è totalmente responsabile di tutte le attività di guida senza alcun supporto da parte di sistemi automatizzati.

I "Livelli 1" e "Livelli 2" sono noti rispettivamente come "Assistenza al Conducente" e "Automazione Parziale di Guida". In questi scenari, un conducente umano è ancora attivamente coinvolto nella guida del veicolo, ma può beneficiare di alcune funzioni automatizzate che lo assistono nella gestione delle attività di guida. È importante

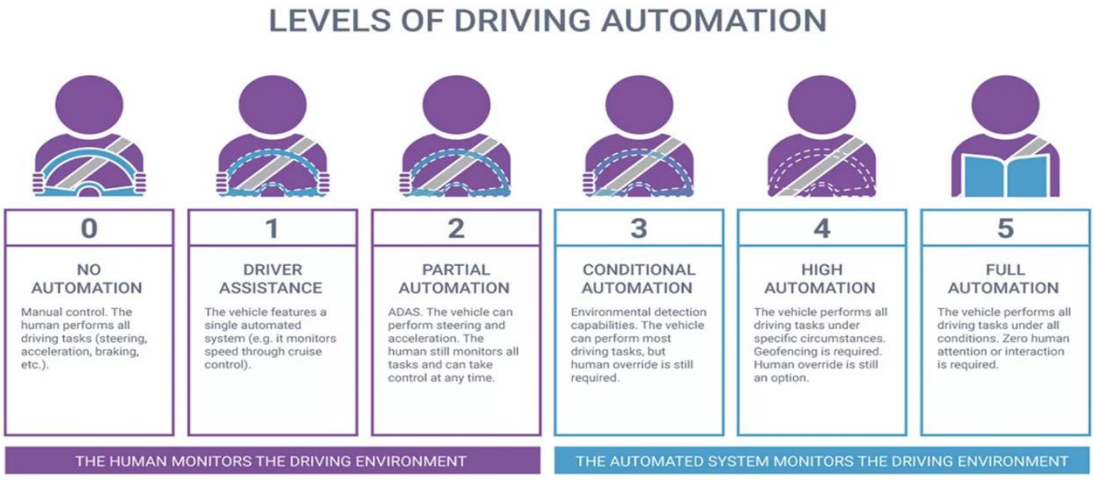
[7] https://www.forbes.com/sites/joannmuller/2013/03/21/no-hands-no-feet-my-unnerving-ride-in-googles-driverless-car/
[8] https://www.sae.org

notare che questi livelli di automazione sono già presenti in alcuni modelli di veicoli attualmente disponibili sul mercato. Tuttavia, è essenziale comprendere che questi livelli di automazione non corrispondono ancora alla definizione di auto a guida autonoma.

I "Livelli 4" e "Livelli 5" sono invece noti come "Alta Automazione di Guida" e "Automazione di Guida Completa". In queste categorie, i veicoli sono tecnicamente in grado di guidare autonomamente senza richiedere un costante controllo umano.

Nonostante l'alto grado di autonomia, le auto a questi livelli di automazione mantengono comunque controlli manuali. Questi comandi consentono a un conducente presente a bordo di prendere il controllo del veicolo e gestire situazioni eccezionali o emergenze che potrebbero verificarsi e che non possono essere completamente affrontate dal sistema automatizzato.

Il percorso verso l'adozione completa della guida autonoma costituisce un significativo avanzamento nell'ambito dell'industria automobilistica, suscitando questioni di notevole complessità inerenti alla sicurezza stradale, alla normativa vigente e alla definizione delle responsabilità giuridiche. Questa è un'area in costante e rapida evoluzione, che richiede una profonda riflessione per affrontare le sfide emergenti e garantire un futuro della guida autonoma sicuro e affidabile.



[9] SYNOPSYS®, The 6 levels of autonomy explained.

2.2 BENEFICI

L'adozione diffusa dei sistemi di guida autonoma promette una serie di benefici significativi, con una particolare enfasi sulla sicurezza stradale[10]. Secondo le stime dell'Organizzazione Mondiale della Sanità, gli incidenti stradali causano la perdita di 1,35 milioni di vite umane ogni anno, e oltre 50 milioni di persone risultano ferite in tutto il mondo[11].

Ciò che è ancor più preoccupante è che la maggior parte di questi incidenti è il risultato di errori umani, come la stanchezza, la distrazione e l'abuso di alcol.

In questo contesto, l'implementazione di sistemi di guida autonoma potrebbe portare a una drastica riduzione di tali tragici eventi, stimata fino al 90%, contribuendo in modo significativo a salvare milioni di vite ogni anno[12].

Inoltre, va notato che le automobili dotate di sistemi di guida autonoma sono meno suscettibili agli effetti dei rischi ambientali, come la scarsa illuminazione o la presenza di nebbia. Questa capacità è possibile grazie all'impiego di avanzati sensori e algoritmi di intelligenza artificiale che consentono a tali veicoli di percepire l'ambiente circostante in modo straordinariamente preciso e di prendere decisioni in tempo reale in base a tali informazioni.

L'ampia diffusione di veicoli a guida autonoma sulle strade offre anche l'opportunità di una collaborazione altamente efficiente tra questi veicoli. Questa sinergia può essere resa possibile attraverso lo scambio istantaneo di informazioni tramite il cloud computing e sfruttando le straordinarie capacità sensoriali delle macchine, nonché i tempi di reazione rapidi che caratterizzano questi sistemi.

Grazie alla condivisione di dati in tempo reale attraverso la connessione cloud, le auto autonome possono comunicare tra loro informazioni cruciali sul traffico, sulle condizioni stradali e su situazioni di emergenza. Questo consente una gestione più efficiente e coordinata del traffico, riducendo gli ingorghi e migliorando la sicurezza stradale complessiva.

Inoltre, le macchine autonome sono in grado di reagire istantaneamente alle informazioni ricevute e alle situazioni impreviste, poiché i loro sistemi di intelligenza artificiale consentono tempi di reazione estremamente rapidi. Questo si traduce in un

[10] B. W. Smith ,“Human error as a cause of vehicle crashes”
[11] https://osservatoriosullasalute.it/wp-content/uploads/2023/06/ro-2022-incidenti.pdf
[12] S. Shwartz “ Are self driving cars really safer than human drivers?”, the gradient 2021

risparmio significativo di tempo per gli automobilisti, poiché i veicoli possono adattarsi in modo rapido alle condizioni stradali in evoluzione senza ritardi dovuti all'errore umano o all'incertezza nella guida.

La presenza di auto a guida autonoma nella nostra vita comporterebbe non solo un importante cambiamento sociale, ma coinvolgerebbe anche molti aspetti legali.

Per quanto riguarda le auto a guida completamente autonoma, esiste un vuoto legislativo in quanto la totale assenza di controlli rende impossibile ritenere il conducente responsabile.

Ma è possibile imputare una colpa penale a un dispositivo a intelligenza autonoma? Possiamo trovare gli stessi principi tipici delle sanzioni penali nell'attribuzione di responsabilità alle vetture intelligenti?[13] È possibile considerare il veicolo un soggetto autonomo intelligente capace di apprendere dalle sanzioni un comportamento migliore per il futuro? È più probabile la responsabilità dei programmatori, dei progettisti e dei produttori del veicolo, anche se non hanno causato direttamente il danno e non erano presenti al momento del reato, e il veicolo non aveva problemi tecnici?

Questo è uno dei principali ostacoli che troviamo nell' applicazione di queste nuove tecnologie e nella responsabilità dell'suo dell'intelligenza artificiale.

La presenza di diversi utenti umani coinvolti, gli algoritmi utilizzati e le componenti meccaniche presenti nel momento dell'evento dannoso, comporta delle notevoli difficoltà nell' individuazione dell'attribuzione della responsabilità.

È importante notare che le auto a guida autonoma, come tutti i robot gestiti da intelligenze artificiali, non sono oggetti prodotto-passivi, ma sono dotate di autonomia decisionale[14] e di azione grazie a processi di apprendimento automatico[15] in cui la macchina è in grado di imparare dall'esperienza, diventando quindi in grado di gestire un numero potenzialmente indefinito di scenari e variabili, sebbene con un numero non illimitato di istruzioni originariamente date.

---

[13] S. Beck, "Google cars, software agents, autonomous weapons systems – New challenges for criminal law?", 2017

[15] I. Goncharov , "The Role Of Machine Learning In Autonomous Vehicles"

Un evento negativo per la diffusione delle auto a guida autonoma è l'incidente mortale coinvolgente un veicolo autonomo di Uber che si è verificato il 18 marzo 2018 a Tempe, in Arizona, negli Stati Uniti[16].

In quell'occasione, un veicolo Uber autonomo, una Volvo XC90, era coinvolto in un test di guida autonoma.

Tuttavia, il veicolo era in modalità autonoma ma con un operatore umano al volante pronto a intervenire in caso di emergenza.

L'incidente ha avuto luogo quando una donna di nome Elaine Herzberg, che attraversava una strada fuori da una zona pedonale, è stata colpita dal veicolo Uber. Tragicamente, la donna è deceduta a causa delle ferite riportate nell'incidente.

Il conducente umano a bordo del veicolo non è riuscito a reagire in tempo per evitare la collusione.

Gli investigatori hanno dichiarato che il conducente dell'auto, Rafael Vasquez, stava guardando un episodio dello show televisivo "The Voice" al momento dell'incidente17.

Le immagini della dash-cam rilasciate dalla polizia mostravano la signora Vasquez che guardava verso il basso, lontano dalla strada, per diversi secondi immediatamente prima dell'incidente, mentre l'auto viaggiava a 39 miglia all'ora (circa 63 chilometri all'ora).

L'incidente ha sollevato una serie di questioni importanti riguardo alla sicurezza dei veicoli autonomi e all'adeguatezza delle normative e delle procedure di test. Ha portato a una sospensione temporanea dei test di guida autonoma di Uber e ha suscitato un ampio dibattito sull'etica e la responsabilità legale nelle situazioni in cui sono coinvolti veicoli autonomi.

In seguito all'incidente, Uber ha rafforzato le misure di sicurezza nei suoi test di guida autonoma e ha ripreso gradualmente le attività di test su strada dopo aver ottenuto l'approvazione delle autorità competenti. Questo incidente ha avuto un impatto significativo sulla percezione pubblica e sulla regolamentazione dei veicoli autonomi,

---

[16] https://www.nytimes.com/2020/09/15/technology/uber-autonomous-crash-driver-charged.html

[17] https://www.repubblica.it/motori/sezioni/attualita/2018/03/19/news/uner_blocca_l_auto_a_guida_autonoma-191701544/
https://www.bbc.com/news/technology-54175359

evidenziando la necessità di affrontare in modo completo le sfide della sicurezza nell'adozione di questa tecnologia.

3.INTELLIGENZA ARTIFICIALE COME STRUMENTO DI PREVENZIONE E CONTRASTO DEL CRIMINE

L'Intelligenza Artificiale nel contesto informatico rappresenta un insieme sofisticato di software e hardware in grado di eseguire elaborazioni e ragionamenti complessi, che possono essere assimilati all'intelligenza umana e spesso superarla grazie all'enorme potenza computazionale che possiede.

L'IA si basa su criteri di valutazione e processi decisionali molto più complessi e veloci rispetto a quelli tipici delle capacità umane, consentendo di manipolare enormi quantità di dati in brevissimi lassi temporali. Tre aspetti fondamentali caratterizzano l'IA in questo contesto: la performatività, che riguarda la quantità di dati elaborati e la velocità di elaborazione; la neutralità statistico-matematica, che si riferisce all'obiettività e alla coerenza dei calcoli effettuati dall'IA; e l'adattabilità, che consente all'IA di essere applicata in diverse situazioni, compreso il campo della sicurezza pubblica e del sistema giudiziario.

La giustizia predittiva può assumere due significati distinti. In primo luogo, si riferisce al settore giudiziario stesso, dove l'IA può essere utilizzata per supportare i magistrati nelle decisioni e nell'individuazione di condotte penalmente rilevanti. In secondo luogo, la giustizia predittiva implica l'identificazione anticipata di potenziali responsabili e prove nei procedimenti giudiziari, il che può migliorare l'efficienza e la tempestività delle indagini e dei processi legali. In entrambi i contesti, l'IA gioca un ruolo cruciale nel contribuire a rendere il sistema giudiziario più efficiente ed equo.

3.1 POLIZIA PREDITTIVA

La polizia predittiva rappresenta un'applicazione dell'intelligenza artificiale nell'ambito delle forze dell'ordine, mirata a compiere previsioni statistiche su vari aspetti dei reati futuri. Questi aspetti includono la possibile localizzazione geografica in cui un crimine potrebbe verificarsi, le modalità con cui potrebbe essere perpetrato, il momento in cui potrebbe accadere e persino l'identificazione dei potenziali autori. L'essenza della polizia predittiva risiede nell'analisi dei dati storici, con l'utilizzo di algoritmi avanzati

per generare tali previsioni, il tutto con l'obiettivo primario di prevenire reati e migliorare l'efficacia complessiva delle operazioni di polizia[18].

Per comprendere appieno la polizia predittiva, è fondamentale considerare che essa si basa su un insieme di attività orientate all'impiego di metodi statistici al fine di "prevedere" dettagli significativi per anticipare la potenziale commissione di reati. Questi strumenti consentono di individuare chi potrebbe essere coinvolto in attività criminali, così come di individuare il luogo e il momento in cui tali reati potrebbero verificarsi, contribuendo in questo modo alla loro prevenzione.

Le previsioni effettuate dalla polizia predittiva si basano su una complessa analisi dei dati, che comprendono informazioni relative a crimini pregressi, movimenti e comportamenti sospetti, luoghi di interesse, periodo dell'anno e persino le condizioni meteorologiche che potrebbero essere correlate alla commissione di specifici reati. Inoltre, tali analisi possono coinvolgere dati demografici, livello di istruzione, situazione economica e caratteristiche fisiche rilevanti per l'identificazione di soggetti potenzialmente coinvolti in certe tipologie di reati, come ad esempio il terrorismo.

In sintesi, la polizia predittiva rappresenta un approccio innovativo che sfrutta l'IA e le analisi statistiche per prevedere e prevenire attività criminali, contribuendo così all'efficacia delle operazioni di polizia.

Un esempio rilevante della tecnologia utilizzata è rappresentato dal "Risk Terrain Modeling" (RTM[19]), un algoritmo che, grazie al ri-processamento di enormi quantità di dati riguardanti i fattori ambientali e spaziali che favoriscono il crimine, sembra consentire la previsione della commissione di reati legati al traffico di droga in determinate aree urbane.

Alcuni fattori ambientali considerati dal "Risk Terrain Modeling" includono: la presenza di luci stradali scarse o non funzionanti, la vicinanza a locali notturni, fermate dei mezzi pubblici, stazioni ferroviarie, incroci di strade molto accessibili, sportelli bancomat, negozi di oro, parcheggi "park-and-ride" e scuole.

---

[18] C. Cath, S. Wachter, B. Mittelstadt, M. Taddeo, L. Floridi, "Artificial Intelligence and the "Good Society": the US, EU, and UK approach", 2018

[19] https://www.riskterrainmodeling.com

Questo ha permesso lo sviluppo di una vera e propria "mappatura" di alcune grandi aree metropolitane al fine di identificare i così detti "hot spots"[20] , cioè luoghi dove il rischio di traffico di droga è più elevato portando a considerevoli vantaggi in termini di pianificazione e attuazione di misure preventive legate alla criminalità connessa al traffico di droga[21].

Questi software[22] si basano sull'idea di base che alcune forme di crimine si verifichino in un periodo molto limitato e in un'area geografica specifica, noti come "reati di ripetizione ravvicinata".

Ad esempio, a causa delle scoperte del software, sembrerebbe che la commissione di una rapina sia associata a un elevato rischio di commetterne un'altra, da parte degli stessi autori e in un'area geografica molto vicina al luogo del primo crimine, entro le successive 48 ore[23].

Di rilevante importanza è l'impatto che l'utilizzo di questi dati ha sulla privacy delle persone.

L'impiego di tali sistemi deve essere attentamente regolamentato infatti un uso eccessivo di essi potrebbe comportare rischi reali di violazione dei diritti e delle libertà fondamentali dell'Unione, tra cui il diritto alla privacy, il diritto all'informazione e persino il diritto di difesa.

Alcuni dati potrebbero essere troppo personali per essere conservati, e chi ne è responsabile potrebbe mancare della capacità e della professionalità necessarie per mantenerli sicuri.

Considerando la natura delicata di tali informazioni, la possibilità di fughe di dati è particolarmente preoccupante.

Il 6 ottobre 2021, il Parlamento europeo ha approvato una dichiarazione riguardante l'intelligenza artificiale nel campo del diritto penale e l'utilizzo da parte delle autorità di polizia e giudiziarie in questioni penali.

I membri del Parlamento europeo hanno enfatizzato la loro opposizione all'uso di sistemi di intelligenza artificiale da parte della polizia e delle autorità di giustizia penale

---

[20] L. Sherman , "Hot spots of crime and criminal careers of places"
[21] M.Caplan, L.W. Kennedy, J.D. Barnum, E.L Piza, "Crime in Context: Utilizing Risk Terrain Modeling and Conjunctive Analysis to Explore the Dynamics of Criminogenic Behavior Setting"
[22] R. Pelliccia, "Polizia Predittiva: il futuro della prevenzione criminale?"
[23] Algorithmic risk assessment policing models: lessons from the Durham HART model and "Experimental" proportionality, in Information & Communications Technology Law, 2018

in Europa, quando tali sistemi automatizzano ingiustizie, minano i diritti fondamentali e generano risultati discriminatori.

Questa affermazione rappresenta una ferma dichiarazione di intenti da parte del Parlamento europeo di protezione dei cittadini europei da tali sistemi e costituisce un primo passo verso il divieto di alcune delle pratiche più dannose, come la sorveglianza di massa basata su dati biometrici.

3.2 PROBLEMI LEGATI ALLA POLIZIA PREDITTIVA

Un possibile nuovo orizzonte potrebbe essere aperto grazie all'uso di specifici algoritmi [24](strumenti di valutazione del rischio o algoritmi predittivi), capaci di elaborare enormi quantità di dati per individuare relazioni, coincidenze, correlazioni che consentirebbero di profilare una persona e prevedere il suo comportamento successivo, anche di rilevanza criminale[25].

Nonostante ciò, nessun paese ha emanato leggi per regolamentare l'uso di queste nuove tecnologie, lasciando un vuoto giuridico nelle condizioni e modalità del loro utilizzo e nella valutazione e nell'uso dei risultati generati.

Un tale vuoto legale limita non solo i vantaggi che queste tecnologie potrebbero portare, ma crea anche considerevoli conflitti nell'ambito della protezione della privacy, a causa dell'enorme quantità di dati personali memorizzati e del divieto di discriminazione, poiché alcune decisioni vengono prese in base a determinate caratteristiche etniche, religiose o sociali.

A livello dell'Unione europea, l'articolo 22 del Regolamento generale sulla protezione dei dati, il cosiddetto GDPR (2016/679/UE), che riguarda la protezione della riservatezza, dichiara illegittime tutte quelle decisioni basate esclusivamente su un trattamento automatizzato che produce effetti legali negativi per i soggetti interessati, intendendo per trattamento automatizzato un insieme di azioni che si sviluppano senza alcun coinvolgimento umano nel processo decisionale.

---

[24] C. P.Haberman, D. Hatten, G.Carter, L. Piza, "The sensitivity of repeat and near repeat analysis to geocoding algorithms", Journal of Criminal Justice

[25] J. Monahan, "Predicting violent behavior: An assessment of clinical techniques", SAGE Library of Social Research, 1981.

Nel paesaggio legale europeo attuale, vi è quasi nessuna legislazione in materia, se non il documento "Sistemi di giustizia"[26], adottato dalla Commissione per l'efficienza della giustizia, che stabilisce i principi generali che dovrebbero governare la materia.

Di grado superiore, tuttavia, è l'articolo 5(3)[27] della Convenzione europea dei diritti dell'uomo (CEDU), che stabilisce che "Ogni persona arrestata o detenuta [...] deve essere condotta il più presto possibile davanti a un giudice", suggerendo la necessità di un contatto umano tra la persona detenuta e il magistrato.

Ciò significa che l'uso degli algoritmi predittivi non è da considerarsi illegale, ma sicuramente molto limitato, poiché è sempre necessario un intervento umano.

3.3 GIOVE

Giove[28] rappresenta il nuovo sistema di polizia predittiva che il Ministero dell'Interno sta valutando di implementare presso le questure di tutta Italia. Secondo quanto dichiarato dalle autorità, questo software sarebbe in grado di utilizzare dati storici per predire dove e quando potrebbero verificarsi specifici tipi di reati, con l'obiettivo di prevenire e reprimere efficacemente i reati che hanno un impatto sociale significativo.

Tuttavia, va notato che numerose ricerche hanno sollevato dubbi riguardo all'efficacia di tali sistemi, evidenziando la loro tendenza a produrre previsioni inaccurate e ad operare basandosi su pregiudizi sottostanti.

Giove è essenzialmente un software che si basa su un algoritmo di intelligenza artificiale. Esso sfrutta i dati raccolti dalle forze dell'ordine in merito a reati passati con l'obiettivo di predire dove e quando potrebbero verificarsi reati simili in futuro. L'origine di questo sistema risale al Dipartimento di Pubblica Sicurezza del Ministero dell'Interno, che ha iniziato a svilupparlo nel 2020.

Il suo sviluppo è collegato a sperimentazioni condotte precedentemente dalla Questura di Milano a partire dal 2008, utilizzando il software KeyCrime, che è stato

[26] European Ethics Charte on the use of artificial intelligence in judicial systems and related areas adopted by the CEPEJ at its 31ª Plenary Meeting (Strasbourg, 3-4 December 2018)
[27] Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols No. 11 and 14
[28] https://www.ilsole24ore.com/art/nome-codice-giove-ecco-l-algoritmo-polizia-prevenire-reati-AEAoJIZD

ideato dall'ex Assistente Capo del medesimo comando, Mario Venturi e sembrerebbe che il meccanismo alla base di Giove sia simile a quello utilizzato da KeyCrime.

Giove sembra essere concepito come un sistema di supporto logistico avanzato in quanto sarà in grado di elaborare e incrociare una vasta gamma di dati, tra cui immagini, denunce, video e altre informazioni, al fine di assistere le forze dell'ordine nella pianificazione delle operazioni.

Ad esempio, il sistema potrà suggerire il numero ottimale di agenti necessari in una determinata zona o indicare le fasce orarie in cui dovrebbero essere intensificati i controlli.

In questo modo, Giove mira a ottimizzare l'efficacia delle risorse delle forze dell'ordine per prevenire e contrastare i reati.

3.4 APPLICAZIONE AI NELLA GIURISDIZIONE

Nel corso di questo capitolo, esamineremo in che modo queste nuove tecnologie, possano essere sfruttate per agevolare e assistere la giurisdizione.

L'IA offre notevoli opportunità per ottimizzare i processi legali e migliorare l'efficienza nell'applicazione della legge. Inoltre, può fungere da prezioso strumento di supporto decisionale per le autorità competenti, consentendo loro di prendere decisioni informate e basate su dati accurati.

In un contesto in cui vi è un dibattito in corso sulla possibilità di delegare l'emissione di sentenze giudiziarie e l'interpretazione delle leggi a sistemi automatizzati basati sull'intelligenza artificiale, sorge una questione di fondamentale importanza: quella della responsabilità legale delle macchine per i danni che possono causare e la possibilità di attribuire loro una sorta di "personalità o soggettività giuridica" che implicherebbe diritti e doveri, analogamente a quanto avviene per le persone fisiche e giuridiche.

Questo aspetto costituisce uno dei temi più complessi e interessanti collegati all'intelligenza artificiale, richiedendo un approfondimento per definire in modo chiaro cosa si intende esattamente per "macchina intelligente".

La discussione si concentra sulla determinazione dei confini della responsabilità delle macchine, comprese le questioni di colpa e negligenza, nonché sulla possibilità di

concedere a tali entità artificiali uno status legale che vada oltre la semplice considerazione di strumenti o dispositivi.

4. DISCRIMINAZIONE A OPERA DELL'INTELLIGENZA ARTIFICIALE

Joy Boulamwini, una studentessa di informatica presso il MIT di Boston, ha fatto una scoperta significativa durante una delle sue attività di laboratorio[29]. Mentre utilizzava un software di riconoscimento facciale, ha notato che il software era molto meno preciso nel riconoscerla, essendo una donna di pelle scura, rispetto ai suoi compagni di corso, che erano in gran parte uomini di carnagione bianca. Questa osservazione ha suscitato il suo interesse e la sua preoccupazione[30].

In seguito, in collaborazione con l'associazione Algorithmic Justice League, che si impegna per promuovere un uso consapevole delle tecnologie dell'intelligenza artificiale, Joy Boulamwini ha condotto uno studio sistematico su diversi software di riconoscimento facciale. I risultati di questa ricerca hanno rivelato un comportamento preoccupante e sistematico: la precisione di tali sistemi sembrava dipendere significativamente dal colore della pelle e dal genere delle persone coinvolte[31].

Questa disparità diventa particolarmente critica quando il riconoscimento facciale viene utilizzato all'interno di sistemi di identificazione impiegati dalle forze dell'ordine, una pratica comune in molti Paesi. Organizzazioni per la difesa dei diritti umani, come Big Brother Watch nel Regno Unito, hanno sollevato gravi preoccupazioni in merito. Spesso, gli applicativi utilizzati in questo contesto non sono stati sottoposti a una verifica adeguata, presentando tassi di errore significativi e manifestando chiari pregiudizi.

Le macchine basate sull'apprendimento automatico e l'intelligenza artificiale imparano dai dati, e questo processo può generare iniquità se i dati su cui si basano sono limitati o non rappresentativi della realtà completa. In altre parole, se queste macchine vengono addestrate su dataset che riflettono solo una parte della realtà, il loro comportamento può risultare distorto e ingiusto. È come se la macchina, avendo accesso solo a una porzione limitata di informazioni, sviluppasse una conoscenza parziale e limitata del contesto.

---

[29] Study finds gender and skin-type bias in commercial artificial-intelligence systems, Larry Hardesty, MIT News Office, 2018
[30] https://news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212
[31] https://time.com/5520558/artificial-intelligence-racial-gender-bias/

Un esempio lampante di questa problematica è il software di riconoscimento facciale[32]. Se il software apprende principalmente da immagini di individui di carnagione chiara, avrà maggiore precisione nel riconoscere queste persone rispetto a individui di carnagione scura, poiché la sua esperienza si è concentrata su un gruppo demografico specifico.

Un fenomeno simile, sebbene meno evidente, si verifica quando i sistemi di apprendimento automatico basati su dati generati dagli utenti, come i motori di ricerca su Internet, addestrano i loro algoritmi. Questi algoritmi imparano a fornire risultati basati sulle informazioni condivise dagli utenti, ma queste informazioni possono essere distorte dalla natura stessa degli utenti che le producono. Ad esempio, una percentuale molto limitata di utenti può generare la maggior parte dei contenuti su piattaforme come Facebook, Amazon o Twitter. Questo significa che le opinioni e i punti di vista rappresentati in queste informazioni possono essere fortemente influenzati da una minoranza che potrebbe non essere rappresentativa in termini di diversità geografica, demografia e genere.

Quando i sistemi di apprendimento automatico utilizzano dati limitati e non rappresentativi per imparare e poi cercano di migliorare i servizi attraverso meccanismi di raccomandazione basati sul comportamento degli utenti stessi, possono amplificare ulteriormente la mancanza di rappresentatività dei dati. Questo fenomeno solleva importanti questioni riguardo all'equità e all'imparzialità dei sistemi basati sull'intelligenza artificiale, poiché potrebbero perpetuare disuguaglianze e discriminazioni presenti nei dati di addestramento.

Il caso del riconoscimento facciale rappresenta solo uno dei numerosi esempi di comportamenti discriminatori che emergono dall'uso dell'intelligenza artificiale. Queste situazioni sollevano importanti questioni etiche e richiamano l'attenzione sulla necessità di affrontare i pregiudizi e garantire un utilizzo equo e responsabile di queste tecnologie emergenti.

Di gran lunga, l'algoritmo predittivo più famoso utilizzato negli Stati Uniti è quello di COMPAS[33] , acronimo di "Correctional Offender Management Profiling for

[32] J. Boulamwini, **"**Artificial Intelligence Has a Problem With Gender and Racial Bias. Here's How to Solve It"

[33] J. Larson, S. Mattu, L. Kirchner and J. Angwin "How We Analyzed the COMPAS Recidivism Algorithm", 2016

Alternative Sanctions", un software di proprietà della società californiana Equivant (precedentemente nota come Northpointe fino al 2017).

Si tratta di un programma che si basa su complessi algoritmi informatici o, in base alla categorizzazione precedentemente introdotta, può essere considerato un sistema esperto progettato per fornire supporto nell'ambito delle decisioni giudiziarie.

Il software COMPAS è stato sviluppato nel 1998 ed è stato utilizzato a partire dal 2000 per valutare il rischio di recidiva degli imputati nei due anni successivi al momento del test. Nel corso degli anni, ha analizzato e valutato più di un milione di individui, fornendo informazioni cruciali per il processo decisionale nel sistema giudiziario[34].

Questo algoritmo tiene conto delle risposte a 137 domande relative a elementi come la storia criminale, reati e infrazioni precedenti, abuso di sostanze, problemi economici, ambiente sociale, pensiero pro-criminale...

COMPAS ha accesso a un vasto archivio di dati che si basa sulla raccolta di informazioni statistiche riguardanti condannati, fattori di rischio associati a vari casi (come alcolismo, droga, prostituzione e altri), e recidive nel corso di molti anni. Utilizzando il risultato del questionario compilato dall'imputato e confrontandolo con i dati statistici precedentemente acquisiti, il software è in grado di calcolare il rischio di recidiva dell'individuo valutato. Questo approccio consente a COMPAS di fornire una stima del potenziale comportamento futuro dell'imputato in base alle informazioni raccolte e analizzate nel suo vasto database statistico.

Il concetto giuridico di recidiva è stato suddiviso in tre categorie principali per facilitare l'analisi e la comprensione del sistema COMPAS:

La recidiva preprocessuale riguarda il rischio di mancata comparizione al processo e la possibilità di un nuovo arresto dell'imputato durante il periodo di rilascio anticipato prima del processo. I fattori considerati in questa fase includono il tipo di reato con la pena più grave, la presenza di casi pendenti, precedenti mancate comparizioni, arresti precedenti con cauzione, condanne precedenti al carcere, storia di abuso di droghe, situazione occupazionale e la durata della residenza.

La recidiva generale è utilizzata per prevedere la commissione di un reato in recidiva dopo l'analisi effettuata da COMPAS. Gli elementi valutati in questa fase

[34] J. Larson, S. Mattu, L. Kirchner, J. Angwin, “Machine Bias” and ”How We Analyzed the COMPAS Recidivism Algorithm” 2016.

includono arresti precedenti, condanne precedenti al carcere, periodi di libertà vigilata precedenti, problemi legati all'educazione e all'occupazione, uso di droghe in passato, età dell'imputato al momento della valutazione e età al momento del primo arresto.

L' ultima categoria è la recidiva violenta che mira a calcolare la probabilità che l'imputato commetta reati violenti dopo il rilascio. I fattori considerati in questa fase includono episodi di violenza e non conformità nel passato, problemi educativi e occupazionali, età dell'imputato al momento della valutazione e età al momento del primo arresto.

Sono state avanzate diverse critiche a COMPAS, ma senza dubbio la più nota è emersa nel caso Loomis[35], il nome dell'imputato che aveva fatto appello alla Corte Suprema del Wisconsin per contestare il livello della pena inflittagli dalla corte locale che aveva utilizzato COMPAS[36].

Nel determinare la sentenza, i giudici avevano preso in considerazione i risultati del programma COMPAS, secondo cui Loomis era stato identificato come un uomo con un alto rischio di recidiva.

Nel caso specifico della determinazione della pena di sei anni di reclusione e cinque anni di sorveglianza estesa, il Tribunale circondariale di La Crosse ha fatto affidamento sui risultati ottenuti tramite l'utilizzo di COMPAS.

Il risultato delle valutazioni effettuate da COMPAS è rappresentato graficamente mediante un diagramma a tre barre. In questo grafico, viene assegnato un punteggio da uno a dieci per ciascuna delle tre categorie di recidiva considerate. I punteggi corrispondono alle probabilità di rischio associate a ciascuna categoria. La fascia a basso rischio va da uno a quattro, la fascia a rischio medio va da cinque a sette, mentre la fascia ad alto rischio va da otto a dieci. Nel caso specifico di Eric Loomis, il software aveva determinato che era ad alto rischio in tutte e tre le categorie di recidiva prese in considerazione.

Successivamente, Loomis ha presentato una richiesta di revisione della pena sostenendo che l'uso di COMPAS avesse violato il suo diritto a un processo equo. Dopo

[35] Wisconsin S.C., State v. Loomis, 881, N.W.2d 749, 2016
[36] "Algorithmic Due Process: Mistaken Accountability and Attribution in State v. Loomis", in Harvard JOLT Digest, 2017

il respingimento della richiesta e l'appello seguente, la questione è stata portata alla Corte Suprema del Wisconsin, che ha emesso la sentenza definitiva in data 13 luglio 2016.

Durante il processo di revisione della pena, è emersa una controversia significativa riguardo all'utilizzo di COMPAS e ai suoi effetti sulla decisione della pena.

La posizione della difesa può essere riassunta in tre principali punti.

Innanzitutto, COMPAS è un programma sviluppato da una società privata e i dettagli del suo funzionamento sono protetti come segreto commerciale. Di conseguenza, la difesa non può accedere a informazioni dettagliate sul processo decisionale di COMPAS e valutare se le sue conclusioni siano corrette. La difesa sostiene che ciò costituisce una violazione del diritto costituzionale al giusto processo.

Le decisioni generate dal sistema COMPAS non possono essere considerate valide in quanto non conducono a una sentenza individualizzata, contrariamente ai principi del giusto processo. Il programma è progettato per valutare dati su un gruppo, non per fornire una valutazione personalizzata per ciascun individuo.

Infine, COMPAS utilizza il genere come variabile nel processo decisionale, il che genera controversie in quanto può portare a una presunta discriminazione. La difesa afferma che questo costituisce una violazione del diritto a un giusto processo.

In sintesi, la difesa cerca di mettere in discussione l'affidabilità di COMPAS enfatizzando l'opacità delle sue operazioni, la sua natura privata, il suo processo di elaborazione delle informazioni e i presunti problemi di discriminazione nei risultati. La difesa afferma che senza la possibilità di esaminare in dettaglio il modo in cui COMPAS valuta i dati, non è possibile verificare l'accuratezza delle sue valutazioni.

La Corte, in risposta alla difesa, ha sostenuto che Eric Loomis aveva avuto l'opportunità di verificare le informazioni utilizzate da COMPAS per la sua valutazione del rischio.

Questo è stato basato sul fatto che gran parte delle informazioni utilizzate dal software erano pubbliche o fornite direttamente da Loomis stesso poiché le domande e le risposte utilizzate da COMPAS erano basate su dati ampiamente accessibili o forniti volontariamente da Loomis durante il processo.

La Corte ha per cui ritenuto che Loomis avesse già avuto la possibilità di verificare l'accuratezza delle informazioni che aveva fornito al sistema COMPAS.

Per quanto riguarda il secondo punto, la Corte ha enfatizzato l'importanza cruciale del diritto a un giudizio individualizzato. Ha sottolineato che i calcoli effettuati da COMPAS costituivano solo uno dei numerosi elementi considerati dal giudice nell'ambito della valutazione complessiva.

In altre parole, il software COMPAS rappresenta solo uno strumento di supporto alla decisione. Il giudice deve prendere in considerazione i risultati del sistema COMPAS insieme a tutti gli altri elementi pertinenti a ciascun caso specifico, garantendo così un giudizio individualizzato.

L'ultima delle tre critiche sollevate da Eric Loomis riguardava l'uso del genere come elemento nelle valutazioni effettuate dal sistema COMPAS. Loomis ha sostenuto che il sistema, basandosi su dati statistici che indicano gli uomini come più inclini alla recidiva, ha portato i giudici a emettere sentenze che discriminavano in modo illegittimo in base al genere. La sua preoccupazione era che questa presunta discriminazione di genere avesse influenzato negativamente la sua pena.

Tuttavia, la Corte ha respinto questa critica e ha concordato con la decisione del Tribunale di La Crosse. Secondo i giudici, la difesa non è riuscita a dimostrare che il genere avesse avuto un impatto così significativo sulla condanna.

La Corte ha concluso che l'uso del genere da parte di COMPAS mirava a promuovere un maggior grado di precisione nelle decisioni, il che alla fine beneficiava sia il sistema di giustizia che gli imputati. Pertanto, ha ritenuto che l'uso del genere da parte del sistema COMPAS non fosse discriminatorio, ma piuttosto finalizzato a migliorare l’accuratezza delle valutazioni.

4.1 CONSIDERAZIONI DOPO COMPAS

Questo caso è di fondamentale importanza perché ha sollevato importanti questioni relative ai diritti costituzionali e ha attirato l'attenzione su come l'IA e i software di valutazione del rischio possano influenzare il sistema legale[37].

Le principali perplessità emerse da questo caso riguardano l'integrazione e l'imparzialità che uno strumento di intelligenza artificiale può avere.

Inoltre, va preso in considerazione il fatto che l'output risultante non è accompagnato da una spiegazione che potrebbe giustificarlo oltre all'analisi dei dati inseriti, poiché il programma è un prodotto "privato" coperto da segreto industriale.

La sentenza della Corte Suprema ha alla fine convalidato l'uso legittimo dell'output prodotto dal software da parte del tribunale territoriale; ciò sulla base della considerazione che questo output era stato utilizzato come semplice elemento ausiliario alle valutazioni condotte dal giudice umano.

Nonostante la decisione fosse stata presa dal software, la sentenza è stata considerata accettabile perché il giudice aveva espresso la stessa opinione.

I giudici supremi hanno sottolineato l'importanza delle precauzioni da prendere quando si utilizzano valutazioni effettuate da programmi predittivi. Questa enfasi è particolarmente rilevante in tre contesti specifici.

Innanzitutto, nella determinazione di misure alternative alla detenzione per gli imputati considerati a basso rischio di recidiva. In questi casi, è fondamentale che le valutazioni dei programmi predittivi siano accuratamente considerate per garantire che le misure alternative siano appropriate e non eccessive.

Nella valutazione della possibilità di reinserire in modo sicuro un individuo condannato nella società, magari attraverso programmi di affidamento in prova. In queste situazioni, le valutazioni dei programmi predittivi possono fornire informazioni cruciali per determinare se un individuo è pronto per la reintegrazione e se ci sono rischi significativi associati.

Nell'imposizione di termini e condizioni per la libertà vigilata, nel monitoraggio e nell'applicazione di sanzioni per le violazioni delle regole nei regimi alternativi alla detenzione. In questo contesto, l'utilizzo delle valutazioni predittive può essere utile per stabilire regole e misure efficaci per garantire il rispetto delle leggi e la sicurezza pubblica.

In ciascuno di questi scenari, i giudici hanno riconosciuto che le valutazioni dei programmi predittivi possono essere uno strumento prezioso, ma è fondamentale adottare precauzioni rigorose per garantire che siano utilizzate in modo equo, trasparente e nel rispetto dei diritti fondamentali degli individui coinvolti nel processo penale.

Ma cosa sarebbe successo se il giudice avesse avuto un'opinione diversa? E se per il giudice la pena da infliggere fosse stata inferiore a quella decisa dalla macchina? In

questo caso, si potrebbe pensare che la macchina abbia agito in modo discriminatorio attraverso l'apprendimento automatico? Si può incolpare la macchina o perché c'è una persona con la stessa opinione si può pensare che la macchina non decida in modo discriminatorio?

4.2 PROPUBBLICA

Nel 2014, l'ex Procuratore Generale degli Stati Uniti, Eric Holder, emise un avvertimento in merito ai possibili bias introdotti nei tribunali dall'uso dei punteggi di rischio. Holder fece appello alla Commissione delle Sentenze degli Stati Uniti affinché avviasse uno studio sull'utilizzo di tali punteggi, sottolineando le sue preoccupazioni.

Egli affermò che, nonostante questi strumenti fossero stati sviluppati con le migliori intenzioni, esisteva il rischio che involontariamente compromettessero gli sforzi per garantire una giustizia individualizzata ed equa. Aggiunse che tali strumenti potevano persino esacerbare le ingiustificate e ingiuste disparità già troppo comuni nel sistema di giustizia penale e nella società in generale.

Nonostante l'appello di Holder, la Commissione delle Sentenze non avviò uno studio sui punteggi di rischio.

Fu in questo contesto che ProPublica, un'organizzazione senza scopo di lucro con sede negli Stati Uniti dedicata al giornalismo investigativo a favore dell'interesse pubblico, decise di condurre un'indagine indipendente su larga scala sull'influenza degli algoritmi nella vita americana, con particolare attenzione ai punteggi di rischio.

Nel 2016, ProPublica, ha condotto un'analisi che ha suscitato considerevole attenzione, successivamente oggetto di numerosi studi scientifici. Questa analisi ha messo in luce in modo significativo un presunto problema di discriminazione nelle valutazioni effettuate dal sistema COMPAS nei confronti dei cittadini afroamericani.

ProPublica ottenne i punteggi di rischio assegnati a oltre 7.000 persone arrestate nella contea di Broward, in Florida, tra il 2013 e il 2014. Successivamente, verificò quanti di questi individui furono effettivamente accusati di nuovi reati nei due anni successivi, utilizzando lo stesso parametro temporale adottato dagli autori dell'algoritmo[37].

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

I risultati dell'indagine evidenziarono una notevole inaffidabilità del punteggio nel prevedere i reati violenti, poiché solo il 20 percento delle persone considerate a rischio elevato di recidivare commise effettivamente reati violenti.

Secondo quanto emerso dall'indagine condotta da ProPublica, il sistema COMPAS sarebbe inadeguato per il ruolo che effettivamente svolge nei tribunali statunitensi, ovvero influenzare la qualificazione e la quantificazione delle sanzioni penali decise dai giudici.

Un esempio emblematico di questa disparità riguarda il caso di Brisha Borden, una giovane donna afroamericana in Florida, valutata con un rischio di recidiva di 8 dopo aver commesso reati minori, mentre Vernon Prater, un uomo bianco con precedenti condanne per rapina a mano armata, è stato valutato con un rischio di 3. Sorprendentemente, due anni dopo, Prater è stato condannato a otto anni di reclusione per furto, mentre Borden non ha commesso alcun reato193.

Tuttavia, l'inchiesta di ProPublica ha rivelato ulteriori criticità. L'organizzazione ha analizzato i punteggi di rischio assegnati a più di settemila persone arrestate nella contea di Broward, in Florida, nel periodo compreso tra il 2013 e il 2014. Successivamente, ha verificato quante di queste persone sono state incriminate in nuovi reati nei due anni successivi. Nel caso dei reati violenti, i punteggi di rischio si sono dimostrati poco precisi, poiché solo il 20% delle persone considerate ad alto rischio ha effettivamente commesso ulteriori reati. Se si estende l'analisi a tutte le categorie di reati, l'algoritmo si è rivelato poco più accurato del lancio di una moneta.

I risultati evidenziano un problema di pregiudizio da parte di COMPAS nei confronti di specifiche etnie.

È sorprendente notare che gli imputati di colore sono erroneamente classificati come ad alto rischio il doppio delle volte rispetto agli imputati bianchi, e viceversa: gli imputati bianchi vengono erroneamente considerati a basso rischio più frequentemente rispetto agli imputati di colore. Questo fenomeno si manifesta anche se analizziamo categorie specifiche di reati, come i reati violenti, dove gli imputati di colore hanno il 77% in più di probabilità di essere classificati come ad alto rischio rispetto agli imputati bianchi.

Da aggiungere a questa problematica c'è anche l'opacità dei calcoli effettuati dal sistema per determinare i punteggi. Questo quadro presenta diverse sfide, in quanto non

solo il software mostra pregiudizi, ma le persone coinvolte non comprendono appieno il motivo di tali discriminazioni, pur subendone pesantemente le conseguenze durante i processi legali.

Il paradosso in questa situazione è che se il sistema operasse in modo accurato ed equo, potremmo avere un sistema di giustizia penale privo dei pregiudizi tipicamente umani che spesso influenzano, talvolta in modo inconscio, le decisioni dei giudici.

In altre parole, un sistema automatizzato come COMPAS, se sviluppato correttamente, potrebbe contribuire a ridurre le disparità e a garantire una maggiore equità nei procedimenti legali, poiché si baserebbe su dati oggettivi e parametri stabiliti in modo uniforme, piuttosto che su giudizi individuali soggetti a possibili influenze personali o preconcetti.

5. IL PROBLEMA DELLA COLPEVOLEZZA DELLE MACCHINE INTELLIGENTI

La complessa questione della responsabilità e della soggettività dei robot come abbiamo già visto si estende anche all'ambito del diritto penale, introducendo ulteriori sfide e complessità. Uno dei pilastri fondamentali del diritto penale, il principio di legalità, impedisce attualmente l'applicazione analogica delle norme penali ai reati commessi da robot, in quanto questi enti non godono dello status di "persone" ai sensi della legge.

Inoltre, affinché un soggetto possa essere ritenuto colpevole di un reato, è necessario dimostrare che la sua condotta abbia direttamente causato l'evento criminale e che esista un legame psicologico, una sorta di intento o volizione, alla base dell'atto.

L'articolo 85 del Codice penale stabilisce che è imputabile chi ha la capacità d'intendere e di volere, ossia chi è in grado di comprendere il significato delle proprie azioni e delle leggi, consentendo di scegliere la propria condotta.

Alla luce di tali disposizioni, sorge la domanda se una macchina intelligente possa essere considerata in grado di intendere e volere. Alcuni studiosi suggeriscono la possibile applicazione dell'articolo 111 del codice penale, intitolato "Determinazione al reato di persona non imputabile o non punibile". Questo articolo stabilisce che, a causa dell'irresponsabilità del robot, la responsabilità per il reato ricadrebbe su chi ha creato, fabbricato o utilizzato il robot. In questa prospettiva, l'intelligenza artificiale sarebbe assimilata a una sorta di creazione dell'essere umano.

L'idea alla base di questa interpretazione è che il robot non può essere ritenuto responsabile dei reati a cui è stato indotto dal suo programmatore.
Tuttavia, ciò pone una questione cruciale: fino a che punto l'atto illecito della macchina può essere attribuito al suo produttore o proprietario?

In questo contesto, è essenziale distinguere tra casi di condotta illecita "autonoma" del robot e casi di condotta illecita "indotta". Nel primo caso, se il reato viene imputato esclusivamente all'essere umano, si riconoscerebbe una forma di responsabilità oggettiva per cui facciamo riferimento al principio giuridico secondo il quale una persona può essere ritenuta penalmente responsabile per un'azione criminosa anche senza la necessità di dimostrare un elemento soggettivo come l'intenzione o la colpa.

Nel secondo caso, oltre alla responsabilità umana, si dovrebbe esaminare la possibilità di attribuire una responsabilità concorrente alla macchina, insieme a una riflessione su quale criterio di attribuzione sia più appropriato.

Queste riflessioni evidenziano la complessità delle questioni legate alla responsabilità penale nell'ambito dell'intelligenza artificiale e la necessità di sviluppare criteri chiari e principi legali per affrontare tali sfide nell'era dell'IA.

È importante inoltre individuare se si prendono in considerazioni eventi in cui i robot vengono utilizzati come semplici strumenti dei reati, cioè come oggetti che dipendono completamente dall'intervento umano e incapaci di compiere azioni completamente autonome. Di conseguenza, i comportamenti potrebbero sempre essere attribuiti all'essere umano sia dal punto di vista oggettivo che soggettivo: la macchina non potrebbe eseguire alcuna azione senza un intervento umano precedente, e qualsiasi azione, per quanto complessa possa essere, verrebbe attribuita all'agente umano che la compie con piena consapevolezza e volontà.

In questa prospettiva, il ruolo della macchina sarebbe quello di un esecutore passivo delle decisioni umane, agendo esclusivamente in base a istruzioni e input forniti dall'essere umano. In altre parole, i robot agirebbero come estensioni degli individui umani, incapaci di prendere decisioni indipendenti o di avere una coscienza propria. La responsabilità per qualsiasi atto illecito commesso dalla macchina sarebbe quindi attribuita all'essere umano che ha orchestrato o utilizzato la macchina per scopi illeciti.

La prospettiva considerata è solo una delle molte in cui l'intelligenza artificiale può essere coinvolta. Questo approccio semplifica la questione della responsabilità penale legata all'IA, evitando complicazioni legate all'attribuzione di personalità giuridica o responsabilità autonoma alle macchine intelligenti.

Tuttavia, continua a sollevare questioni importanti sul ruolo dell'essere umano nell'utilizzo e nel controllo delle macchine intelligenti, e sulla necessità di sviluppare regolamentazioni chiare e criteri giuridici per affrontare le sfide emergenti nell'era dell'IA.

5.1 PROSPETTIVA LEGALE

Se immaginiamo un futuro in cui strumenti come gli algoritmi predittivi diventano ancora più ampiamente diffusi, le implicazioni potrebbero essere significative e complesse.

Nel capitolo precedente sul caso Loomis abbiamo esaminato un noto esempio in cui la macchina e il giudice erano concordi sulla severità della pena da assegnare. Tuttavia, cosa accadrebbe se questi strumenti venissero utilizzati in modo più frequente e non ci fosse sempre un accordo tra la macchina e il giudice?

Una possibile conseguenza potrebbe essere una maggiore attenzione e discussione sulla questione della discriminazione algoritmica. Se i giudici iniziano a percepire che gli algoritmi prendono decisioni in modo discriminatorio, attribuendo all'offensore una colpevolezza maggiore di quanto sia giustificato, potrebbero sorgere questioni di discriminazione nell'applicazione della legge. Questo solleva interrogativi importanti sulla responsabilità degli algoritmi e su come garantire che le decisioni siano giuste ed equamente distribuite.

Un altro scenario da considerare è quello delle auto a guida autonoma di livello 4 e 5 coinvolte in incidenti stradali. Se tali veicoli autonomi fossero coinvolti in un incidente in cui investono un pedone in un attraversamento pedonale, emergerebbe la questione della responsabilità legale. Questo solleva sfide complesse per stabilire chi è responsabile in situazioni in cui il controllo dell'auto è affidato a un sistema autonomo anziché a un conducente umano.

La peculiarità dei sistemi di intelligenza artificiale è la loro capacità di agire e prendere decisioni autonomamente in modi che potrebbero non essere stati previsti dai loro creatori umani. Questa caratteristica apre la porta a considerazioni sulla teoria della responsabilità legale, inclusa la responsabilità penale, per le azioni di questi sistemi. Il fatto che le macchine possano agire autonomamente e talvolta in modi imprevisti solleva questioni complesse sulla definizione della responsabilità legale e sulla possibilità di attribuire la colpa a un sistema non umano.

Di fronte a questa nuova prospettiva, sorge la domanda fondamentale se le macchine possano essere considerate soggetti attivi di reato, superando l'antico assioma secondo cui "la macchina non può commettere un reato e quindi non può essere punita." Questo interrogativo potrebbe richiedere una revisione profonda della teoria della

responsabilità legale in un mondo sempre più guidato dall'automazione e dall'intelligenza artificiale.

Il Parlamento europeo[38], attraverso la sua Risoluzione datata ottobre 2021 intitolata "L'intelligenza artificiale nel diritto penale", ha sollevato un importante punto di discussione riguardo alla necessità di stabilire un regime giuridico ben definito ed equo per attribuire la responsabilità legale e l'imputabilità per le possibili conseguenze negative generate dall'uso delle avanzate tecnologie digitali, inclusa l'intelligenza artificiale.

Questa Risoluzione evidenzia la crescente importanza e diffusione delle tecnologie digitali avanzate, in particolare l'intelligenza artificiale, nel campo del diritto penale e sollecita l'adozione di misure chiare e trasparenti per affrontare le questioni relative alla responsabilità legale e all'imputabilità associate a tali tecnologie. In altre parole, il Parlamento europeo riconosce l'urgenza di definire le regole e le leggi necessarie per gestire le sfide legali poste dall'uso sempre più diffuso dell'intelligenza artificiale nel contesto penale.

La Risoluzione sottolinea che l'intelligenza artificiale è in grado di prendere decisioni autonome e comportarsi in modi che potrebbero risultare imprevisti o indesiderati dai loro creatori umani. Questo solleva una serie di interrogativi cruciali, come chi è responsabile quando le decisioni o le azioni di un sistema di intelligenza artificiale portano a conseguenze negative o perfino a un reato. La Risoluzione sottolinea la necessità di affrontare queste questioni in modo equo ed efficace.

Inoltre, è importante notare che questa Risoluzione riflette la crescente consapevolezza delle sfide legali poste dall'intelligenza artificiale non solo nell'ambito del diritto penale ma anche in altri settori della società. Le tecnologie digitali avanzate stanno diventando sempre più onnipresenti nelle nostre vite, e pertanto è essenziale sviluppare un quadro legale che tenga conto di queste trasformazioni e che garantisca la responsabilità e l'imputabilità in modo adeguato.

[38] A. Valsecchi, " L'intelligenza artificiale nel diritto penale: la Risoluzione del Parlamento europeo del 6 ottobre 2021"

Facendo riferimento alla versione in lingua inglese, è evidente che nel Considerando J si sottolinea l'obbligo di mantenere sempre una forma di responsabilità umana.

Nell'articolo 13 della Risoluzione, si stabilisce invece l'obbligo di identificare in modo costante un agente, che potrebbe essere sia una persona fisica che giuridica, e che sia legalmente responsabile e soggetta a responsabilità legale per le decisioni prese con il supporto dei sistemi di intelligenza artificiale.

L'articolo 13 della Risoluzione prosegue poi affermando che il Parlamento Europeo invita, pertanto, ad applicare con coerenza il principio di precauzione per tutte le applicazioni di IA nel contesto delle attività di contrasto; sottolinea che la responsabilità giuridica e l'imputabilità devono sempre ricadere su una persona fisica o giuridica, che deve sempre essere identificata per le decisioni assunte con il sostegno dell'IA; sottolinea, pertanto, l'esigenza di assicurare la trasparenza delle strutture aziendali che producono e gestiscono i sistemi di IA [...].

La Risoluzione ribadisce con chiarezza che il principale obiettivo delle future normative sarà la prevenzione degli effetti negativi derivanti dall'uso dell'intelligenza artificiale. Questa visione è profondamente radicata nel principio di precauzione, che sottolinea l'importanza di adottare misure preventive per evitare danni potenziali.

La Risoluzione mette in evidenza cinque punti chiave di notevole importanza. Questi punti riguardano il potenziale rischio di discriminazione derivante dall'utilizzo dell'intelligenza artificiale, la necessità di condurre valutazioni d'impatto, l'importanza della trasparenza nei sistemi algoritmici, l'utilizzo di sistemi di polizia predittiva e l'imposizione di un divieto sulla sorveglianza di massa basata su dati biometrici.

È importante notare che, sebbene questa Risoluzione non abbia forza vincolante, essa rappresenta una tappa significativa nel processo legislativo in merito ai rapporti tra diritto penale e utilizzo dell'intelligenza artificiale. Inoltre, essa si affianca alla proposta di regolamento sull'intelligenza artificiale, conosciuta come "Artificial Intelligence Act," conferendo ulteriore peso a questa discussione.

La Risoluzione affronta tre linee guida principali identificate dal Parlamento europeo. Queste linee guida rappresentano le questioni considerate più critiche dai deputati europei. In particolare, esse riguardano la necessità di garantire la trasparenza nei sistemi algoritmici, la gestione dei rischi connessi alla discriminazione, e le

preoccupazioni legate all'uso diffuso di apparecchiature di sorveglianza, in particolare le tecnologie di identificazione facciale.

In conclusione, la Risoluzione del Parlamento europeo evidenzia l'importanza di affrontare le questioni legate all'intelligenza artificiale nel diritto penale e sottolinea la necessità di creare un sistema giuridico chiaro ed equo per gestire la responsabilità legale e l'imputabilità associate alle tecnologie digitali avanzate. Ciò rappresenta un passo significativo verso l'elaborazione di normative adeguate per affrontare le sfide poste dall'intelligenza artificiale nella società contemporanea.

La questione della responsabilità riveste una grande importanza, e di ciò se ne fa carico lo stesso Parlamento Europeo, il quale, nella risoluzione del 16 febbraio 2017, scrive che "Grazie agli strabilianti progressi tecnologici dell'ultimo decennio, non solo oggi i robot sono in grado di svolgere attività tipicamente ed esclusivamente umane, ma lo sviluppo di determinate caratteristiche autonome e cognitive, ad esempio la capacità di appren- dere 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. Pertanto, la questione giuridica derivante dall'azione nociva di un robot diventa essenziale"

Le problematiche nascono in realtà dalla circostanza per cui "più i robot sono autonomi, meno possono essere considerati come meri strumenti nelle mani di altri attori (quali il fabbricante, l'operatore, il proprietario, l'utilizzatore, ecc.)

Tale circostanza inge- nera preoccupazioni, in quanto, chi li ha progettati potrebbe non essere sempre in grado di prevedere le reazioni che il robot stesso potrebbe sviluppare.

Ed è proprio in questo che consiste il cd. "*Responsability gap"* individuato dalla Risoluzione quale problema centrale della *"robot liability"*.
Detto in altre parole, il nodo centrale del problema della responsabilità per i danni provocati dai robot consiste nel fatto che, tali danni, potrebbero essere provocati da azioni non programmate da colui che lo ha progettato, realizzandosi così un vuoto di responsabilità. Ci si chiede se esista un soggetto al quale attribuire la responsabilità per tali azioni, e nel caso si dovesse dare risposta affermativa, chi dovrebbe essere?

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*?[39]

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

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.

Dall'analisi svolta si può trarre la conclusione che è fortemente auspicabile, se non necessario, che al più presto venga individuata una soluzione normativa volta a regolare la questione della responsabilità penale per i danni provocati dai prodotti dell'A.I., ma che questa allo stato non può essere definitiva.

Per quanto riguarda le vetture semi autonome si è ritenuto di operare un'ulteriore distinzione tra quelle qualificate con un livello inferiore al 4 e quelle che invece raggiungono tale grado di automazione[40].

Per le prime si può definitivamente concludere che, in tali casi, il soggetto responsabile possa essere individuato nella figura del conducente umano, ancora gravato di compiti di sorveglianza attiva e, all'occorrenza, dell'onere di un intervento manuale in situazioni di pericolo.

---

[39] A. Cappellini, "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale"

[40] A. Cappellini, "Profili penalistici delle self-driving cars"

Responsabilità umana si trova anche nei casi in cui in fase di addestramento del sistema sia stato inserito un set di dati incompleto o pregiudizievole in grado di incidere sul comportamento futuro tenuto dalla macchina.

Si ritiene colpevole il programmatore per la mancata previsione di eventi dannosi che co- sostituiscono il logico sviluppo delle modalità di funzionamento dell'agente artificiale da lui programmato.

Si è tuttavia notato, come tale soluzione sembrerebbe non tener conto di uno dei principi cardine dell'intero diritto penale: il principio della personalità della responsabilità penale. Il programmatore si troverebbe in definitiva nella condizione di dover rispondere per fatto altrui, il fatto proprio del robot.

Gabriel Hallevy, esperto di Diritto Penale presso la facoltà di Giurisprudenza dell'Ono Academic College in Israele, offre una prospettiva interessante riguardo alle implicazioni legali dell'Intelligenza Artificiale (IA). Egli sostiene che la generazione più recente di sistemi di IA ha la capacità non solo di "prevedere" determinati risultati, ma anche di "volere" tali risultati grazie alle sue abilità decisionali e di apprendimento profondo[41].

L'idea chiave qui è che l'IA può "ricordare" il passato, apprendere dall' "esperienza" e adattare il suo comportamento in risposta a nuovi stimoli. Questo processo è reso possibile grazie al machine learning e all'accesso alle esperienze di altre IA tramite tecnologie come il cloud computing.

Hallevy applica questa idea all'ambito legale, sostenendo che le IA possono essere considerate responsabili per le loro azioni. In termini legali, l'"actus reus" (l'azione criminosa) e la "mens rea" (la colpevolezza o la consapevolezza dell'azione) sono elementi cruciali per stabilire la colpevolezza in un reato. Hallevy sostiene che le IA possono dimostrare la loro "consapevolezza" attraverso il loro comportamento.

Prende ad esempio i robot telecamera utilizzati in Corea del Sud per assistere le guardie carcerarie. Questi robot possono muoversi autonomamente all'interno di un carcere e, quando rilevano movimenti sospetti, sono in grado di identificare se si tratta di un detenuto che sta cercando di fuggire. Questa capacità di identificazione e reazione può essere considerata un segno di consapevolezza da parte delle IA.

---

[41] G. Hallevy, "The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control"

In sintesi, Hallevy suggerisce che le IA possono essere ritenute responsabili delle loro azioni in un contesto legale, a condizione che dimostrino una forma di consapevolezza o capacità di apprendimento simile a quella umana. Ciò solleva importanti questioni etiche e giuridiche riguardo alla responsabilità penale delle IA in un mondo sempre più guidato dall'automazione e dall'IA.

5.2 ENTI COLLETTIVI

Diversi studiosi sostengono che il tema in questione presenta notevoli somiglianze con l'istituzione di un'etica aziendale e la conseguente responsabilità penale delle entità collettive, un sistema che potremmo prendere in considerazione nel caso di danni causati dall'azione autonoma dell'IA[42].

Le analogie tra questi due contesti sono evidenti: le entità collettive non hanno una presenza fisica o una coscienza, ma sono comunque considerate "soggetti giuridici" in quanto possono commettere reati (attraverso le persone fisiche che le compongono) secondo la legge penale. D'altra parte, i robot hanno un "corpo" fisico che interagisce con l'ambiente attraverso sensori, il che offre una base per l'applicazione di sanzioni penali (come la disattivazione o la riprogrammazione del robot o addirittura la sua distruzione).

Inoltre, i robot sono dotati di autonomia decisionale.

Ciò ci porta a considerare un sistema di responsabilità per l'agente artificiale che, simile alla responsabilità penale delle entità collettive, potrebbe rendere l'agente artificiale soggetto a responsabilità legale per le sue azioni, nonché per quelle delle persone coinvolte, come l'utente, il programmatore, il progettista o il produttore. Questo sistema potrebbe rappresentare un nuovo modo di affrontare la responsabilità legale nell'era dell'IA.

Esistono indubbiamente numerose analogie tra un ente giuridico collettivo e un ente artificiale; tuttavia, come sottolineato in modo acuto da Peter Asaro, emergono altresì significative divergenze. Queste differenze richiedono una riflessione più

[42] V. R. Bhargava, M. Velasquez, “Is Corporate Responsibility Relevant to Artificial Intelligence Responsibility?”

approfondita in merito alla logica alla base della punizione dell'ente giuridico quando applicata al contesto degli agenti di intelligenza artificiale. Questa riflessione è necessaria per varie ragioni, tutte riconducibili a un fatto fondamentale: a differenza delle organizzazioni umane, gli agenti di intelligenza artificiale non sono costituiti da individui umani.

Innanzitutto, l'ente giuridico collettivo si basa sulla cooperazione e l'azione congiunta di individui umani all'interno di una struttura giuridica riconosciuta. Gli enti giuridici rappresentano una consolidata realtà sociale e legale, mentre gli agenti di intelligenza artificiale sono, in effetti, programmi informatici o macchine dotate di capacità decisionali autonome. Pertanto, applicare la stessa logica di punizione che si usa per le organizzazioni umane a entità artificiali presenta sfide significative.

5.3 FUNZIONI DELLA PENA

L' intelligenza artificiale essendo un oggetto inanimato, non è il tipo giusto di entità da punire. L'IA manca di stati mentali e delle capacità deliberative necessarie per la colpevolezza; quindi, non può essere punita senza sacrificare gli impegni fondamentali del diritto penale.

Le IA non sono coscienti e non provano sentimenti e non possiedono interessi o benessere per cui non c'è motivo per pensare che l'IA beneficia delle protezioni del vincolo della colpevolezza, che proibisce di punire in eccesso rispetto a quanto merita la colpevolezza.

Una pena irrogata a una macchina intelligente, infatti, non potrebbe svolgere nessuna delle tre classiche funzioni che generalmente la dottrina penalistica riconosce alla sanzione criminale[45].

Di retribuzione non si può certo parlare perché, come già si è detto, le macchine intelligenti di oggi continuano a essere insuscettibili di un rimprovero colpevole.

Egualmente difficoltoso è scorgere nella pena robotica una funzione rieducativa. Rieducare, infatti, presuppone la possibilità di apprendere dalla sanzione irrogata in conseguenza di una propria azione sbagliata.

Salvo che non la si programmi appositamente, una IA non può essere rieducata attraverso una pena. Ma anche in tal caso, un “difetto di comportamento” può essere sistemato da meccanismi di machine learning che “ottimizzino” progressivamente l’agire del soggetto artificiale, o più radicalmente da una riprogrammazione: in maniera diretta, dunque, piuttosto che mediante uno strumento indiretto di pressione quale è la pena[46].

Infine, anche per quanto attiene alla prevenzione generale infatti è difficile pensare come potrebbe una IA comprendere che una sanzione eseguita su di un altro soggetto robotico è in realtà finalizzata a veicolare un messaggio generale di dissuasione rispetto a un certo comportamento.

Quella pena, in definitiva, rimarrebbe un fatto isolato al soggetto artificiale che la su- bisce, ed estraneo a tutti gli altri.

5.4 RIFLESSIONI FINALI

L'identificazione del momento esatto in cui un'intelligenza artificiale (IA) transita da uno stato di "oggetto" a quello di "soggetto" è un quesito di notevole complessità e rilevanza.

Nel corso del tempo, sarà fondamentale stabilire se, in virtù delle sue capacità cognitive avanzate e della sua autonomia decisionale, un'IA possa essere considerata alla stregua di una "persona" dal punto di vista giuridico e se, conseguentemente, debba essere soggetta agli stessi diritti, doveri e responsabilità di una persona naturale. Questa è una questione che sfida i confini tradizionali del diritto e della filosofia e che richiederà un approfondito dibattito interdisciplinare per affrontare tutte le sue implicazioni. In ultima analisi, il futuro ci dirà se e come dovremo considerare l'IA come soggetto di diritto, rappresentando così un passo significativo nell'evoluzione delle leggi e delle normative relative alla tecnologia e all'intelligenza artificiale.

BIBLIOGRAFIA

A. Cappellini, *Profili penalistici delle self-driving cars*, 2019

A. Cappellini, *Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale*, 2019

A. Valsecchi, *L'intelligenza artificiale nel diritto penale: la Risoluzione del Parlamento europeo del 6 ottobre 2021;*

Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols No. 11 and 14;

C. P.Haberman, D. Hatten, G.Carter, L. Piza, *"The sensitivity of repeat and near repeat analysis to geocoding algorithms",* Journal of Criminal Justice

E. Israni, Algorithmic Due Process: Mistaken Accountability and Attribution in State v. Loomis, in Harvard JOLT Digest, 2017;

European Ethics Charte on the use of artificial intelligence in judicial systems and related areas adopted by the CEPEJ at its 31ª Plenary Meeting (Strasbourg, 3-4 December 2018);

G. Hallevy, *The Criminal Liability of Artificial Intelligence Entities - from Science Fiction to Legal Social Control*, 2010;

Ivan Goncharov, *The Role Of Machine Learning In Autonomous Vehicles;*

L. W. Sherman, *Hot spots of crime and criminal careers of places*, in ECK J;

M. Oswald, *Algorithmic risk assessment policing models: lessons from the Durham HART model and "Experimental" proportionality,* in Information & Communications Technology Law, 2018;

J.M. Caplan, L.W. Kennedy, J.D. Barnum, E.L Piza, *Crime in Context: Utilizing Risk Terrain Modeling and Conjunctive Analysis to Explore the Dynamics of Criminogenic Behavior Setting*, in Journal of Contemporary Criminal Justice;

J.M. Caplan, L.W. Kennedy, *Risk Terrain Modeling: Crime Prediction and Risk Reduction, Univ. of California Press,* 2016;

J. Larson, S. Mattu, L. Kirchner and J. Angwin, *How We Analyzed the COMPAS Recidivism Algorithm,* 2016;

J. Monahan, *Predicting violent behavior: An assessment of clinical techniques, SAGE Library of Social Research*, 1981;

L.Hardesty, *Study finds gender and skin-type bias in commercial artificial-intelligence systems*, MIT News Office, 2018;

R. Pelliccia, *Polizia Predittiva: il futuro della prevenzione criminale?;*

R.C. Schank, *What's IA, Anyway?,* in IA Magazine, 1987;

S. Signorato, *Il diritto a decisioni penali non basate esclusivamente su trattamenti automatizzati: un nuovo diritto derivante dal rispetto della dignità umana, in Rivista di diritto processuale*;

V. R. Bhargava, M. Velasquez, *Is Corporate Responsibility Relevant to Artificial Intelligence Responsibility?*

Wisconsin S.C., State v. Loomis, 881, N.W.2d 749, 2016.

SITOGRAFIA

https://www.europafacile.net/sites/default/files/documents/201812141242.COM%282018%29237.pdf

https://www.repubblica.it/motori/sezioni/attualita/2018/03/19/news/uner_blocca_l_auto_a_guida_autonoma-191701544/

https://www.bbc.com/news/technology-54175359

https://www.riskterrainmodeling.com

https://www.ilsole24ore.com/art/nome-codice-giove-ecco-l-algoritmo-polizia-prevenire-reati-AEAoJIZD