

*T4F Series*  
PAPER N. 7  
a.a. 2023/2024

**Responsabilità delle driverless  
cars:  
il dilemma etico e le risposte  
giuridiche**

CHIARA BIFOLCHETTI, MARIA ELISA L.  
CHERCHI, FEDERICO FOMMEI

Trento BiLaw Selected Student Papers

I paper sono stati selezionati a conclusione del corso *Diritto e Intelligenza Artificiale* a.a. 2023-2024, organizzato all'interno della Cattedra Jean Monnet "T4F – TrAIning 4 Future. Artificial Intelligence and EU Law", coordinato presso l'Università di Trento dal docente Carlo Casonato.

# Responsabilità delle driverless cars: il dilemma etico e le risposte giuridiche

Chiara Bifolchetti, Maria Elisa L. Cherchi, Federico Fommei\*

ABSTRACT: L'articolo affronta il tema della responsabilità per i danni causati dalle auto a guida autonoma, esplorando le questioni etiche sollevate dall'uso dell'Intelligenza Artificiale (IA) nei veicoli e provando a fornire una risposta di carattere giuridico che ne tenga conto. La prima sezione, partendo dal rinomato "Trolley Problem", analizza se e come le decisioni morali possano essere programmate nei sistemi di IA. Si discute l'introduzione di una "Manopola Etica" come possibile soluzione per bilanciare il benessere dei passeggeri con quello degli altri utenti della strada. La seconda sezione esamina in senso lato il tema della responsabilità per le decisioni dell'IA, riportando alcune delle teorie discusse in ambito etico e giuridico. La terza sezione si concentra sul caso dell'incidente di Tempe del 2018, suggerendo soluzioni per l'attribuzione della responsabilità nelle tecnologie emergenti.

PAROLE CHIAVE: Responsabilità civile; Responsabilità etica; Intelligenza artificiale; Moralità; Auto a guida autonoma.

ACCOUNTABILITY OF AUTONOMOUS VEHICLES: THE ETHICAL DILEMMA AND LEGAL ANSWERS

ABSTRACT: The article addresses the issue of liability for damages caused by self-driving cars, exploring the ethical issues raised by the use of Artificial Intelligence (AI) in vehicles and trying to provide a legal response that takes them into account. The first section, starting with the renowned "Trolley Problem," analyses how moral decisions can be programmed into AI systems. It discusses the introduction of an "Ethics Knob" as a possible solution to balance the welfare of passengers with that of other road users. The second section looks broadly at the issue of responsibility for AI decisions, reporting on some of the theories discussed in ethics and law. The third section focuses on the case of the 2018 Tempe accident, suggesting solutions for assigning liability in emerging technologies.

KEYWORDS: Civil Liability; Moral Responsibility; Artificial intelligence; Morality; Self-driving cars.

SOMMARIO: 1. Le auto a guida autonoma – 2. I livelli di automazione – 3. Etica e auto a guida autonoma – 3.1 Il dilemma etico del trolley problem – 3.2 Differenze e scelte etiche: mancanza di indirizzi univoci – 4. La responsabilità dell'Intelligenza Artificiale a cavallo tra l'etica e il diritto – 4.1 La Responsabilità elettronica – 4.2 Il vuoto di responsabilità e la Responsabilità Diffusa – 4.3 Il ritorno all'idea di *Strict Liability* – 5. Il caso studio: l'incidente di Tempe del 2018 – 5.1 La responsabilità per i danni cagionati dalle driverless cars – 5.2 La responsabilità del produttore – 6. Conclusioni.

## 1. Le auto a guida autonoma

L'evoluzione tecnologica è un fenomeno trasversale che coinvolge ogni settore del vivere civile, impattando sulla nostra quotidianità. La circolazione certamente non fa eccezione, infatti anche le automobili sono sempre più automatizzate. Le auto a guida autonoma, conosciute anche con l'acronimo inglese CAV (*connected and autonomous car*), funzionano attraverso una combinazione di intelligenza artificiale, sensori

---

\* *Studenti dell'Università di Trento. Chiara Bifolchetti ha curato il paragrafo 4, Maria Elisa L. Cherchi il paragrafo 5 e Federico Fommei il paragrafo 3. Le conclusioni sono frutto del lavoro congiunto degli autori. Email: [chiara.bifolchetti@studenti.unitn.it](mailto:chiara.bifolchetti@studenti.unitn.it); [maria.cherchi@studenti.unitn.it](mailto:maria.cherchi@studenti.unitn.it); [federico.fommei@studenti.unitn.it](mailto:federico.fommei@studenti.unitn.it).*

e radar che consente l'interazione con la realtà circostante<sup>1</sup>. Questi veicoli incorporano una tecnologia di connessione V2X (*vehicle-to-everything*), avendo la capacità di interagire sia con gli altri veicoli (V2V, *vehicle-to-vehicle*) sia con le infrastrutture (V2I, *vehicle-to-infrastructure*)<sup>2</sup>.

I vantaggi derivanti dall'utilizzo di tali nuovi veicoli sono molteplici. Innanzitutto, la guida autonoma consente una maggiore fluidificazione del traffico, ovvero uno sfruttamento più efficiente delle strade, garantendo un maggiore risparmio di tempo, considerata anche la possibilità di impiegare più utilmente il tempo di viaggio<sup>3</sup>. In secondo luogo, si stima che la diffusione delle auto autonome comporti un aumento di produttività nel settore *automotive*, creando nuovi posti d'impiego, e più in generale dovrebbe avere un effetto positivo sull'economia dei Paesi produttori<sup>4</sup>. Inoltre, le *driverless car* potrebbero avere un ruolo fondamentale nella realizzazione di una mobilità più inclusiva, consentendo la circolazione anche a soggetti anziani o diversamente abili, normalmente incapaci di condurre un veicolo<sup>5</sup>. Infine, i CAV dovrebbero essere in grado di diminuire gli incidenti stradali: i dati rivelano che la maggior parte degli incidenti sono causati da fatti dell'uomo, come eccessiva velocità, guida in condizioni alterate, mancato utilizzo di sistemi di sicurezza come caschi e cinture e guida disattenta<sup>6</sup>. Si stima quindi che l'eliminazione del fattore umano dalla guida diminuirà gli incidenti del 90%<sup>7</sup>. Tuttavia, è necessario fare una precisazione. Le auto a guida autonoma vengono programmate *in primis* i cd. *jaywalkers*, i pedoni che attraversano al di fuori delle

---

<sup>1</sup> B. KUIPERS ET AL., *Shakey: From Conception to History*, in *AI Magazine*, 1, 2017, pp. 89-90, disponibile al link <https://doi.org/10.1609/aimag.v38i1.2716> (ultima consultazione 28/10/2024).

<sup>2</sup> COMMISSIONE EUROPEA, *Report GEAR 2030 Strategy 2015-2017 - Comparative analysis of the competitive position of the EU automotive industry and the impact of the introduction of autonomous vehicles*, Bruxelles, 2017, disponibile al link <https://downloads.gws-os.com/ET0417523ENN.en.pdf> (ultima consultazione 17/10/2024). Cfr. anche S. MEENAKSHI AMMAL, M. KATHIRESH, R. NEELAVENI ARTIFICIAL, *Intelligence and Sensor Technology in the Automotive Industry: An Overview*, in M. KATHIRESH, R. NEELAVENI (a cura di), *Automotive Embedded Systems*, Coimbatore, Tamil Nadu, 2021, p. 145.

<sup>3</sup> <https://www.avvens.com/it-it/blog/futuro-della-mobilita/self-driving-cars/> (ultima consultazione 17/10/2024).

<sup>4</sup> Cfr. WORLD ECONOMIC FORUM, *Digital transformation: Reinventing the wheel: digital transformation in the automotive industry*, Cologne, 2016; A.I. CESARANO, *Autonomous driving: an exploratory investigation on public perception*, in *European Journal of Privacy Law & Technology*, 2, 2020, p. 265.

<sup>5</sup> Cfr. S. PETTIGREW, *Why public health should embrace the autonomous car*, in *Australian and New Zealand Journal of Public Health*, 41(1), 2017, p. 6; A. BERTOLINI, E. PALMERINI, *Regulating robotics: A challenge for Europe*, in *EU Parliament, Workshop on Upcoming issues of EU law for the IURI Committee*, Bruxelles, 2014, p. 186; S. VANTIN, *Il diritto antidiscriminatorio nell'era digitale. Potenzialità e rischi per le persone, la Pubblica Amministrazione, le imprese*, Milano, 2021.

<sup>6</sup> Report realizzato dall'organizzazione mondiale della sanità: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries> (ultima consultazione 17/10/2024).

<sup>7</sup> Stime della National Highway Traffic Safety Administration, agenzia governativa statunitense facente parte del dipartimento dei trasporti: disponibili ai link <https://trendsformative.com/it/i-vantaggi-di-passare-dalla-guida-assistita-alla-guida-autonoma/> (ultima consultazione 17/10/2024); e <https://libroverde.autostrade.it/pillole/le-tecnologie-che-allontanano-il-rischio-di-incidenti/> (ultima consultazione 17/10/2024). In tal senso, anche il Rapporto sulla sicurezza stradale 2023 di Dekra, società tedesca che si occupa di sicurezza automobilistica, consultabile al sito <https://dekraproduct-media.e-spirit.cloud/fb2198ff-7a46-490d-8c58-de62fd7caa44/media/dekra-vsr-2023-it-0921-web.pdf> (ultima consultazione 17/10/2024).

Per approfondire, L. YE, T. YAMAMOTO, *Evaluating the impact of connected and autonomous vehicles on traffic safety*, in *Physica A*, 526, 2019; A. BROGGI, A., ZELINSKY, Ü. ÖZGÜNER, C. LAUGIER, *Intelligent Vehicles*, in B. SICILIANO, O. KHATIB, *Springer Handbook of Robotics*, Berlino, 2016, pp. 1627-1656; M.C. GAETA, *Automazione e responsabilità civile automobilistica*, in *Responsabilità civile e previdenza*, 5, 2016, p. 1718 ss.

strisce pedonali. È fondamentale, pertanto, valutare il fenomeno della cd. pericolosità asimmetrica che consiste nel rendere un'attività complessivamente più sicura ma più rischiosa per alcuni soggetti<sup>8</sup>.

## 2. I livelli di automazione

Per comprendere chi possa essere ritenuto responsabile dei danni causati dalle auto autonome è preliminare analizzare il funzionamento delle stesse. Quello di veicolo automatizzato è infatti un concetto generico che ricomprende sia le auto dotate di sistemi ausiliari alla guida umana sia le auto integralmente autonome. La classificazione realizzata dalla SAE (Society of Automotive Engineers International)<sup>9</sup>, individua sei livelli di progressiva automazione, fissando nel passaggio dal livello 2 (*partial automation*) al livello 3 (*conditional automation*) il discrimine tra la guida umana assistita e guida autonoma. I veicoli di livello 2, già ampiamente diffusi, sono capaci di svolgere in autonomia diverse manovre ma necessitano comunque della costante attenzione del guidatore<sup>10</sup>. La tecnologia di livello 3 consente di affidare la guida al sistema intelligente in certe condizioni di traffico e ambientali ma richiede comunque che il guidatore sia sempre in grado di riprendere il controllo della macchina. Infatti, il persistente obbligo in capo al conducente di riprendere il controllo del veicolo quando necessario comporta che, ai fini della responsabilità, la vera distinzione risieda nel passaggio dal livello 3 al livello 4, in cui tutti i soggetti a bordo sono egualmente qualificabili come passeggeri e i comandi di guida, sterzo e pedali, potrebbero non essere affatto installati. La nostra trattazione si propone di affrontare la questione della responsabilità per i danni causati dall'Intelligenza Artificiale (IA) che governa queste automobili da un punto di vista non solo giuridico, ma anche etico. Ci interroghiamo, innanzitutto, sulla capacità delle auto a guida autonoma di prendere decisioni moralmente rilevanti, e poi sulla possibilità di queste (e di altri sistemi di IA) di essere ritenute responsabili per i danni derivati dalle azioni conseguenza dei propri outcomes. Viene poi analizzato l'incidente avvenuto a Tempe, in Arizona, nel 2018. Attraverso questo caso di studio si analizzano quali siano le possibili applicazioni delle norme di responsabilità relative alle figure del conducente, del proprietario e del produttore ai danni causati dalle auto autonome.

---

<sup>8</sup> U. RUFFOLO, *Responsabilità da produzione e gestione di A.I. self-learning*, in P. PERLINGIERI, S. GIOVA, I. PRISCO, *Rapporti civilistici e intelligenze artificiali: attività e responsabilità: atti del 15° Convegno nazionale 14-15-16 maggio 2020*, Napoli, 2020, p. 244.

<sup>9</sup> Si tratta della classificazione SAE J3016 disponibile al sito <https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic> (ultima consultazione 30/09/2024). Nonostante la SAE sia una società privata, la sua classificazione, fatta propria anche dalla US National Highway Traffic Safety Administration, è riconosciuta e utilizzata a livello mondiale. Per un'analisi tecnica dei livelli SAE, cfr. B.C. ZANCHIN, R. ADAMSHUK, M.M. SANTOS, K.S. COLLAZOS, *On the Instrumentation and Classification of Autonomous Cars, Conference Paper*, in *IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Banff Center, Banff, Canada, 5-8 ottobre 2017, disponibile al sito: [www.researchgate.net/publication/320907470](http://www.researchgate.net/publication/320907470) *On the Instrumentation and Classification of Autonomous Cars* (ultima consultazione 17/10/2024); J.M. ANDERSON, N. KALRA, K.D. STANEY, P. SORENSEN, C. SAMARAS, O.A. OLUWATOLA, *Autonomous Vehicle Technology. A guide for Policymakers*, Santa Monica, 2014.

<sup>10</sup> Rientrano nel livello 2 di automazione, ad esempio, i sistemi di parcheggio automatico o i sistemi di *cruise control* adattivo combinato al *lane centering* che mantengono il veicolo in carreggiata a velocità costante salvo rilevare la presenza di altre macchine innanzi, adattando automaticamente la velocità al veicolo di fronte.

### 3. Etica e auto a guida autonoma

La sempre più elevata spinta verso la creazione di autoveicoli a guida completamente automatizzata ha portato alla creazione di automobili sempre più efficienti, sicure e affidabili. Le recenti dichiarazioni del celebre magnate americano Elon Musk circa l'imminente implementazione di sistemi di guida *driverless* di livello di autonomia 4 e 5 è un segnale della sempre maggiore importanza posta dal mercato e dalla stessa società a questo settore tecnologico<sup>11</sup>. A tale irrefrenabile sviluppo si correla una sempre maggiore produzione normativa che, in linea di principio, dovrebbe porre rimedio ai numerosi, ed irrisolti, dubbi che anche tali innovativi sistemi fanno emergere e che si accompagnano a questioni etiche di carattere tradizionale<sup>12</sup>. Se da una parte la riflessione e il dibattito legati ai veicoli a guida autonoma hanno avuto come punto focale le sfide di carattere normativo, tecnico e commerciale, dall'altra si è assistito ad un disinteresse agli altrettanto annosi dilemmi di carattere etico, solo recentemente e lentamente riscoperti<sup>13</sup>. È innegabile, difatti, che lo sviluppo di sistemi di guida sempre più tecnologicamente sofisticati e autonomi porti con sé numerose implicazioni di carattere etico e morale.

Prime fra tutte emergono le questioni legate alla possibilità che tali automezzi si trovino coinvolti in situazioni potenzialmente fonte di pericolo, come in caso di incidenti, e alle conseguenti problematiche di responsabilità. Non va dimenticato, infatti, che i sinistri automobilistici registrati solo in Italia nel 2023 ammontano a 44.778 e non si trova motivo per dubitare che tali numeri possano interessare, soprattutto in un futuro prossimo, un numero crescente di veicoli a guida autonoma o parzialmente autonoma<sup>14</sup>. Ciò spinge, per di più, a svolgere riflessioni più approfondite legate ai processi decisionali che la macchina che si venga a trovare in tali situazioni estreme e di non facile risoluzione, in cui una decisione celere può determinare la vita o meno dei pedoni e del conducente stesso, deve necessariamente compiere. In relazione a ciò, ci si chiede se sia configurabile una qualche forma di moralità in capo a tali strumenti e se questa influenzi le decisioni prese dalla *driverless car* in sostituzione del conducente umano.

#### 3.1. Il dilemma etico del trolley problem

La risposta a tali interrogativi è tutt'altro che semplice e richiede un'analisi approfondita che coinvolge riflessioni su questioni e dilemmi di natura filosofica ed etica. A tal fine risulta utile, se non necessario, fare riferimento al celebre dilemma morale noto come *Trolley problem* o, in italiano, Problema del carrello<sup>15</sup>. Il quesito in oggetto mette in scena un cortocircuito concettuale che nasce dalla compresenza e indecidibilità di due alternative contrapposte che creano un conflitto mentale. Tale simpatico quanto inusuale rompicapo trae origine dal pensiero della filosofa Philippa Foot e dal suo dilemma dello scambio (1967), così

---

<sup>11</sup> <https://insideevs.it/news/680097/tesla-guida-autonoma-livello-4/> (ultima consultazione 17/10/2024).

<sup>12</sup> Regolamento (UE) n. 2024/1689 del Parlamento e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale e modifica alcuni atti legislativi dell'unione (regolamento sull'intelligenza artificiale).

<sup>13</sup> M. BALISTRERI, *Macchine senza guidatore: considerazioni morali*, in *Mondo digitale*, novembre 2019.

<sup>14</sup> Statistiche e dati disponibili presso: [https://www.ansa.it/canale\\_motori/notizie/mondo\\_motori/2024/01/05/meno-vittime-della-strada-nel-2023-calo-del-17\\_fb878ed5-183d-4505-bd77-f5262c8181e4.html](https://www.ansa.it/canale_motori/notizie/mondo_motori/2024/01/05/meno-vittime-della-strada-nel-2023-calo-del-17_fb878ed5-183d-4505-bd77-f5262c8181e4.html) (ultima consultazione 17/10/2024).

<sup>15</sup> G. BOTTINI, *L'auto a guida autonoma e i dilemmi etici*, in *officeadvice.it*, 8 febbraio 2021, disponibile online al seguente link: <https://officeadvice.it/novita-giuridiche/lauto-a-guida-autonoma-e-i-dilemmi-etici/> (ultima consultazione 17/10/2024).

formulato: un vagone ferroviario è fuori controllo e la locomotiva si sta dirigendo verso cinque persone ignare del pericolo che si trovano sui binari; queste sono impossibilitate ad abbandonare i binari in tempo<sup>16</sup>. Nonostante ciò, sorge una possibilità ulteriore: il percorso dei binari presenta una deviazione e quindi, azionando una leva, è possibile stornare la corsa della locomotiva e salvare i cinque malcapitati; tuttavia, un'altra persona si trova lungo il tracciato alternativo e, questa, morirebbe sicuramente se venisse azionata la leva dello scambio. L'alternativa, dunque, è duplice: azionare la leva e deviare il vagone uccidendo una sola persona ma lasciandone in vita cinque, oppure, astenersi da tale azione, lasciando i poveri cinque malcapitati in balia del loro inesorabile destino. La quasi totalità delle persone risponderebbe che, in tale caso, sembrerebbe moralmente ed eticamente doveroso azionare la leva e salvare le cinque persone a scapito del soggetto presente sul binario secondario. In realtà, è stato evidenziato come la maggioranza dei soggetti se posti concretamente nella medesima ideale situazione rimarrebbe paralizzata dal terrore, portando ad un epilogo decisamente ben diverso da quello teoricamente prefigurato<sup>17</sup>.

All'elaborazione della filosofa inglese, risponde la filosofa americana Judith Jarvis Thomson la quale riformula il dilemma della collega configurando il "Dilemma dell'uomo grasso"<sup>18</sup>. Lunghi da considerazioni di carattere medico, la studiosa fondamentale prevede che, in sostituzione del binario alternativo e della leva dello scambio, vicino al binario su cui si sta svolgendo la folle corsa del vagone si trovi un uomo di corporatura robusta e che l'unico modo per impedire la morte dei cinque soggetti sia quello di spingere il corpulento malcapitato contro il mezzo affinché, con il suo peso, riesca a bloccare la corsa. Naturalmente, opposto sarebbe il responso se si sottoponessero ad esame i medesimi soggetti che nel caso precedente avrebbero tirato la leva. Alcuni potrebbero obiettare che nel primo esempio il soggetto compie un'azione, quella di spingere la leva, che inconsapevolmente e indipendentemente dalla sua volontà porta un uomo a morire; mentre, nel secondo esempio, sia deliberata la scelta di spingere un uomo sui binari portandolo alla morte. Tali considerazioni, sicuramente corrette, esulano però dall'ambito d'indagine di questa breve riflessione. Risulta interessante notare che il rapporto costi e benefici in entrambi gli esperimenti mentali risulta il medesimo, un uomo muore e cinque vengono salvati; ciò nonostante, la soluzione alternativa del primo dilemma è generalmente considerata moralmente accettabile mentre quella del secondo assolutamente non eticamente sostenibile. A questo punto è necessario chiedersi come risponderebbe un algoritmo e, soprattutto, una *driverless car* ai quesiti sopra proposti.

Secondo Alan Turing, antesignano delle moderne innovazioni tecnologiche, i calcolatori digitali erano e sono dotati di una particolare forma di intelligenza di tipo computazionale e razionale che «comprende tutti quei processi decisionali per i quali la miglior descrizione e spiegazione è quella che fa ricorso a un sistema computazionale di tipo classico o standard, cioè assimilabile a una macchina di Turing»<sup>19</sup>.

È evidente, infatti, come una macchina non sia in grado di competere con le emozioni e le sensazioni che caratterizzano gli organismi pensanti, ciò nonostante, potrebbe essere in grado di imitare i processi decisionali di questi ultimi al fine di elaborare la soluzione più socialmente accettabile. A tal riguardo si

---

<sup>16</sup> P. R. FOOT, *The Problem of Abortion and the Doctrine of the Double Effect*, in *Oxford Review*, 5, 1967, pp. 5-15.

<sup>17</sup> C. MARCHI, *Uccidere uno per salvarne cinque: il trolley problem e le auto*, in *everyeye.it*, 4 giugno 2021, disponibile online al seguente link: <https://tech.everyeye.it/articoli/speciale-uccidere-salvarne-cinque-trolley-problem-auto-53233.html> (ultima consultazione 17/10/2024).

<sup>18</sup> J. J. THOMSON, *Turning the Trolley*, in *Philosophy & Public Affairs*, 4, 2008, pp. 359-374.

<sup>19</sup> E. SIGNORINI, *Cervelli e carrelli: formalizzare la moralità nelle auto a guida autonoma*, in *Rivista internazionale di filosofia e psicologia*, 1, 2019, pp. 78-93.

sottolinea, all'opposto, come tali sistemi sarebbero dotati di una intelligenza incompleta che, in quanto priva della componente più propriamente umana e istintiva, non sarebbe funzionale alla risoluzione di questioni moralmente complesse<sup>20</sup>. Va rilevato, inoltre, che gli esseri umani necessariamente incorporano nei loro giudizi morali tendenze egoistiche e, posti di fronte ai dilemmi morali, spesso sono indirizzati ad agire dalle circostanze in cui si trovano oltre che dai convincimenti personali e dai valori propri di ciascun individuo. I sistemi di intelligenza artificiale, invece, mediante il processo del machine learning, possono meramente seguire rigidamente la logica del programma ad esse incorporato. Tutte queste considerazioni calate nel concreto stanno a significare che l'IA, in sé e per sé considerata, non compie scelte di carattere etico o morale ma è ontologicamente e tecnologicamente portata alla massimizzazione dei propri obiettivi, ovvero degli scopi che un determinato sistema persegue e per cui viene programmato.

Dunque, mentre la maggior parte degli esseri umani probabilmente si rifiuterebbe di spingere l'uomo sovrappeso sui binari, un sistema di intelligenza artificiale, programmato per massimizzare il benessere collettivo, eseguirebbe tale azione in conformità con i suoi obiettivi<sup>21</sup>. Infatti, l'IA, seguendo rigorosamente la logica del suo programma, continuerebbe a scegliere l'opzione che comporta il minor danno complessivo, indipendentemente dalle implicazioni etiche o emotive che influenzerebbero il giudizio umano. Declinando la questione rispetto alle auto a guida autonoma, è immaginabile che l'obiettivo principale per cui l'intero sistema sia programmato e studiato sia quello di garantire la sicurezza del conducente, quella del mondo circostante e l'integrità del veicolo. Provando a esemplificare facendo riferimento ai dilemmi di cui sopra, dunque, si può ipotizzare che, se il sistema intelligente di una *driverless car* fosse progettato per massimizzare la sicurezza, il singolo in entrambi i casi sarebbe considerato sacrificabile in funzione della sicurezza, non solo dei cinque soggetti, ma anche del conducente stesso. D'altronde, come dichiarato nel 2016, e subito ritrattato per ovvii motivi, da Christoph von Hugo, direttore della Mercedes-Benz per i sistemi di assistenza al conducente e di sicurezza attiva, intervistato dalla rivista *Car and Drive*: «Se sai che puoi salvare almeno una persona, almeno salva quella. Salva quella nell'auto. Se tutto quello che sai con certezza è che puoi evitare una morte, allora quella è la tua priorità»<sup>22</sup>. A tal riguardo, un interessante studio pubblicato dalla rivista *Science* ha chiesto a un gruppo di persone di valutare su una scala da 0 a 100 quanto fossero d'accordo sul fatto che le macchine a guida autonoma debbano essere programmate per salvare il maggior numero di vite<sup>23</sup>. Come ipotizzabile i punteggi erano risultati molto elevati. Nonostante ciò, gli stessi intervistati si erano dichiarati contrari all'idea di acquistare tali veicoli, poiché, pur essendo progettati per salvaguardare il maggior numero di vite, avrebbero inevitabilmente aumentato il rischio per quella del conducente. Si può infatti affermare che le macchine non prendano decisioni prettamente morali bensì utilitaristiche, legate al contesto di programmazione e alle scelte svolte "a monte" dal programmatore umano<sup>24</sup>. Questo naturalmente genera numerosi problemi sintetizzabili, facendo riferimento agli esperimenti di cui sopra, nel fatto che casi apparentemente così simili, letti in modo asettico e meccanico dal sistema informatico, in realtà sono profondamente diversi sul piano etico.

---

<sup>20</sup> *Ivi*, p. 85.

<sup>21</sup> A. LORENTE, *Driven towards a moral crash*, in *Rivista internazionale di filosofia e psicologia*, 11, 2020, pp. 223-237.

<sup>22</sup> <https://attivissimo.blogspot.com/2016/10/se-unauto-autonoma-deve-scegliere-chi.html> (ultima consultazione 17/10/2024).

<sup>23</sup> J.-F. BONNEFON, A. SHARIF, I. RAHWAN, *The social dilemma of autonomous vehicles*, in *Science*, 24 giugno 2016, disponibile al link <https://www.science.org/doi/10.1126/science.aaf2654> (ultima consultazione 17/10/2024).

<sup>24</sup> C. MARCHI, *op. cit.*

Una possibile soluzione a tale cortocircuito potrebbe essere quella di programmare tali soluzioni in anticipo, assumendosi la responsabilità di tali scelte, pensando a delle regole generalizzate che vadano ad orientare il lavoro dei creatori di tali strumenti e il funzionamento stesso di tali macchine in funzione della tutela, *in primis*, dei diritti fondamentali protetti dalle carte costituzionali<sup>25</sup>. In realtà, idee innovative che potranno in un futuro prossimo condizionare l'evoluzione tecnica e il veloce sviluppo tecnologico in tale ambito, sono emerse per dare una possibile soluzione al problema fin qui considerato.

Si fa riferimento, in particolare, alla cosiddetta "*Ethical knob*" o, in italiano, "Manopola etica", espressione coniata dai ricercatori del dipartimento di scienze giuridiche dell'università di Bologna, G. Contissa, F. Lagioia e G. Sartor, in collaborazione con la *European University Institute* di Firenze<sup>26</sup>. Si tratta, come facilmente intuibile dal nome, di un vero e proprio interruttore installato nelle auto a guida autonoma dotate di maggiore autonomia (dunque riferibile al livello 4 e 5 della scala graduata) che dovrebbe permettere al conducente di scegliere, in base all'interesse che ritiene maggiormente meritevole di tutela, tra tre distinte modalità: altruista, egoista, imparziale.

Sulla scorta di tale scelta il veicolo autonomo varierà la propria condotta in caso di sinistro stradale. Più precisamente, nel primo caso, la vettura cercherà di limitare i danni all'esterno mettendo, dunque, in secondo piano l'incolumità delle persone a bordo; nel secondo, il mezzo assicurerà la protezione dei passeggeri in via prioritaria; infine, nella terza modalità, la vettura cercherà sempre il male minore applicando, come già osservato, un principio utilitaristico. Dunque, il punto centrale sul quale si concentra la scelta del conducente è quello di stabilire chi tutelare e a quale livello di rischio sottoporre sé stesso e i terzi. Questa proposta, sebbene abbia carattere prettamente tecnico, potrebbe rappresentare una soluzione determinante al fine della risoluzione del problema qui affrontato. Va rilevato, inoltre, come la sua implementazione non sarebbe priva di interessanti ripercussioni sia sul piano etico che su quello giuridico. Infatti, è stato argutamente evidenziato che: «L'ammissione della Manopola Etica da parte degli ordinamenti consentirebbe di lasciare la responsabilità in ambito individuale, nel caso in cui la scelta sia totalmente in capo ai singoli, ferma restando la possibilità per le autorità statali di fissare dei limiti entro i quali si potrà legittimamente privilegiare sé stessi. Il divieto del Knob sottrarrebbe la scelta etica alla responsabilità individuale, trasferendola all'ordinamento. Ad esso spetterebbe quindi scegliere se adottare un atteggiamento imparziale, privilegiare le vite dei passeggeri o quelle dei pedoni»<sup>27</sup>. Infine, va considerato il contributo al dibattito fornito dell'Università di Stanford, la quale ha pubblicato sulla rivista *Science Robotics* una ricerca relativa all'autonomia degli autoveicoli che ha evidenziato come i modelli di guida autonoma siano sempre più in grado di fare fronte alla gestione degli imprevisti e alle situazioni di pericolo<sup>28</sup>. In particolare, è stato sottolineato come tali sistemi basati sull'IA apprendano dalle esperienze passate al fine rendere più sicura la guida in caso di imprevisti e condizioni meteo non ottimali. Secondo

---

<sup>25</sup> K. KIRKPATRICK, *The moral challenges of driverless cars*, in *Communications of the ACM*, 8, 2015, pp. 19-20.

<sup>26</sup> G. CONTISSA, F. LAGIOIA, G. SARTOR, *The Ethical Knob*, in *Artificial Intelligence and Law*, 25, 2017, pp. 365-378.

<sup>27</sup> G. BOTTINI, *op. cit.*

<sup>28</sup> R. MAURICI, *Etica della guida autonoma*, in *Istituto Italiano di Tecnologia*, 3 maggio 2019, disponibile online al seguente link: <https://opentalk.iit.it/etica-della-guida-autonoma/> (ultima consultazione 17/10/2024).

tale ricerca questi nuovi modelli di guida autonoma dovrebbero, attraverso la raccolta di dati e la loro combinazione, raggiungere il livello di un pilota esperto<sup>29</sup>.

### 3.2 Differenze e scelte etiche: mancanza di indirizzi univoci

Se i problemi fin qui emersi non trovano ancora una soluzione chiara e definitiva, a complicare ulteriormente il quadro complessivo contribuisce anche il fatto che, in realtà, etica e moralità sono fortemente influenzate da fattori culturali e sociali. Dunque, risulta arduo e complesso appellarsi ad una nozione contenutisticamente univoca di etica, in quanto essa risulta difficilmente formalizzabile in un unico *modus decidendi*, inscrivibile poi nel procedimento decisorio di un autoveicolo a guida autonoma.

A tal riguardo uno studio condotto dalla rivista *Nature*, *The Moral Machine Experiment*, oltre a confermare condivise tendenze di carattere generale (quali la preferenza a risparmiare il maggior numero di vite, dare priorità agli esseri umani rispetto agli animali, etc...), ha fatto emergere differenze regionali legate alle culture di appartenenza<sup>30</sup>. Sono stati a tal fine individuati tre grandi blocchi regionali che, pur con differenze tra i paesi compresi in ciascuna classificazione, attribuiscono un peso morale diverso alla decisione in funzione della categoria di appartenenza della vittima<sup>31</sup>. In relazione a ciò, si osserva che nel cosiddetto blocco occidentale la rilevanza ricoperta dalla apparenza della vittima a particolari categorie è minima, salvo le considerazioni di carattere generale di cui sopra. Nel blocco orientale, invece, si tende a preferire una lettura che risparmi da situazioni rischiose in maniera prioritaria i pedoni e chi non si trovasse nel torto secondo il codice della strada. Va però notato come in tali paesi venga riconosciuta una valenza marginale alle categorie dei giovani e, in particolare, delle donne. Infine, nel blocco meridionale queste due categorie specifiche trovano un rilievo primario<sup>32</sup>. I ricercatori hanno enfatizzato, inoltre, come le stesse disuguaglianze economiche giochino un ruolo fondamentale in tale ambito. Nei paesi economicamente svantaggiati, infatti, si osserva una maggiore rilevanza attribuita allo status sociale delle vittime, il che conduce a una preferenza per il sacrificio di individui appartenenti a classi sociali meno privilegiate<sup>33</sup>.

Considerazioni di questo tenore, spingono a ritenere che, logicamente, se lo stesso essere umano non possiede una moralità universalmente condivisa, pensare ad un concetto di etica univoco e uniforme nel campo dei sistemi che utilizzano l'intelligenza artificiale e, in particolare, rispetto alle stesse *driverless car*, è alquanto utopico e irrealistico. Se lo stesso processo decisionale degli uomini è caratterizzato da una forte componente non solo individuale ma anche sociale, nasce spontaneo l'interrogativo circa come declinare queste differenze che, come visto, incidono in modo diretto e irreversibile sui diritti e le libertà dei cittadini, nelle scelte e nelle decisioni svolte dai sistemi che utilizzano l'intelligenza artificiale. Andrebbero forse,

---

<sup>29</sup> N.A. SPIELBERG, M. BROWN, N.R. KAPANIA, J.C. KEGELMAN, J. C. GERDES, *Neural network vehicle models for high-performance automated driving*, in *Science Robotics*, 4, 2019. Disponibile online al seguente link: <https://www.science.org/doi/10.1126/scirobotics.aaw1975> (ultima consultazione 17/10/2024).

<sup>30</sup> E. AWAD, S. DSOUZA, R. KIM, J. SCHULZ, J. HENRICH, A. SHARIFF, J. BONNEFON, I. RAHWAN, *The Moral Machine Experiment*, in *Nature*, 563, 2018, pp. 59–64, disponibile online al seguente link: <https://www.nature.com/articles/s41586-018-0637-6> (ultima consultazione 17/10/2024).

<sup>31</sup> *Ivi*, pp. 60-61. A tal riguardo: il primo blocco (blocco occidentale) contiene il Nord America e gran parte dei paesi Europei; il secondo blocco (blocco orientale) contiene i paesi del Medio Oriente, dell'estremo oriente e del sud est asiatico, il terzo blocco (blocco meridionale) contiene i paesi dell'America Latina.

<sup>32</sup> *Ivi*, p. 63.

<sup>33</sup> *Ivi*, pp. 62-63.

innanzitutto, ricercati punti di convergenza sulle scelte di fondo, per quanto rimarrebbe assai «complesso declinare prospettive etiche all'interno dei processi decisionali automatici di una driverless car»<sup>34</sup>.

#### 4. La responsabilità dell'Intelligenza Artificiale a cavallo tra l'etica e il diritto

Tra le sfide che emergono in questa era sempre più digitalizzata c'è quella dell'adattamento delle tradizionali categorie di responsabilità agli esiti e alle azioni poste in essere da sistemi di IA, come quelli delle CAV. Mentre le macchine diventano sempre più sofisticate e capaci di prendere decisioni senza intervento umano diretto, è possibile considerare l'intelligenza artificiale alla stregua di un agente morale? Dall'analisi svolta è emerso che l'agire dell'Intelligenza Artificiale è guidato da un'ottica prettamente utilitaristica, con il rischio di divergenze dipendenti anche da fattori culturali. Allora chi dovrebbe sopportare la responsabilità per le scelte compiute da un sistema di IA, nel nostro caso una CAV, e per i danni che da queste possono essere derivati? Prima di formulare una risposta giuridica, elaborando una nostra tesi, abbiamo voluto esplorare questo tema dal punto di vista etico. Abbiamo approfondito, a fronte di un dibattito ancora aperto, alcune tra le principali teorie che si concentrano sull'attribuzione di responsabilità per le decisioni e le azioni dei sistemi di Intelligenza Artificiale (IA). Una prima trattazione in merito all'attribuzione di responsabilità si deve ad Aristotele, il quale la associava al soddisfacimento di due condizioni: una di volontarietà e una epistemica.

Nell'Etica Nicomachea, il filosofo greco spiega che l'azione moralmente valutabile deve essere un'azione volontaria: «Siccome è involontario ciò che avviene per forza o per ignoranza, il volontario ci sembrerà essere ciò il cui principio è in chi agisce, quando costui conosca i singoli aspetti nei quali l'azione si verifica<sup>35</sup>». Quindi, per poter essere considerato responsabile il soggetto deve aver agito (i.e. causato l'azione) ed essere stato consapevole, "non ignorante", di quanto accaduto. A partire da questo, M. Coeckelbergh<sup>36</sup> ha giustamente sostenuto che, per quanto si possa riconoscere che l'impatto della tecnologia sul vivere umano vada oltre quello di un mero strumento, e che le azioni di questa possano assumere una rilevanza morale, l'AI Agent non potrà mai incontrare queste due condizioni, in quanto non può essere cosciente di ciò che fa, né è in grado di agire liberamente. La volontarietà dell'azione è infatti connessa al controllo sulle proprie azioni, definibile come «consapevole e lucida capacità di scelta degli agenti<sup>37</sup>». Anche se una tecnologia di intelligenza artificiale si dimostra capace di decidere e agire sulla base di quella decisione (i.e. agency), non si può riscontrare in essa una capacità di autodeterminazione, né (certamente) consapevolezza di quanto sta facendo. La macchina agisce sulla base di un *training* (addestramento), ovvero il processo mediante il quale un modello AI viene addestrato su un insieme di dati (ad esempio, immagini, testi o numeri), imparando a riconoscere schemi e a prendere decisioni. Questo avviene sempre secondo un set predeterminato di regole, cioè un insieme di istruzioni e limiti stabiliti dagli sviluppatori che guida il comportamento dell'AI e ne definisce l'operatività<sup>38</sup>. Nell'ambito della tecnologia

---

<sup>34</sup> R. MAURICI, *Etica della guida autonoma*, in *Istituto Italiano di Tecnologia*, 3 maggio 2019. Disponibile online al seguente link: <https://opentalk.iit.it/etica-della-guida-autonoma/> (ultima consultazione 17/10/2024).

<sup>35</sup> ARISTOTELE, *Etica Nicomachea*, libro III, 1111a3.

<sup>36</sup> M. COECKELBERG, *Artificial intelligence, responsibility attribution, and a relational justification of explainability*, in *Science and engineering ethics*, 26, 2020.

<sup>37</sup> P. DONINI, *Aristotele, la giustizia e la responsabilità*, Trieste, 2008.

<sup>38</sup> P. TRAVERSO, *Breve introduzione tecnica all'Intelligenza Artificiale*, in *DPCE online*, 1, 2022, pp. 155-167.

delle auto a guida autonoma, abbiamo visto nella parte precedente quanto focale sia la possibilità per il progettista o per l'utilizzatore finale (nel caso dell'*ethical knob*) di compiere quelle valutazioni morali che guideranno l'azione dell'auto in situazioni di emergenza. Come già evidenziato, mentre gli esseri umani tendono a incorporare nei loro giudizi morali tendenze egoistiche e sono influenzati dalle circostanze e dai loro valori personali, le macchine seguono rigidamente la logica del programma ad esse incorporato tramite il processo di *machine learning*. Questa logica, solitamente orientata a garantire la sicurezza del conducente, del mondo circostante e l'integrità del veicolo, riflette un approccio prevalentemente utilitaristico, piuttosto che morale, ancorato alle scelte compiute a monte dai programmatori umani.

Escludere che le azioni delle CAV possano essere trattate alla stregua di quelle di un agente morale si traduce, in ambito giuridico, nel rigetto dell'idea di una responsabilità cosiddetta "elettronica", legata al riconoscimento di una personalità elettronica dell'entità robotica. Vedremo come questa posizione, sostenuta tra l'altro dal Parlamento Europeo, è stata fortemente criticata dalla dottrina. Questo porta a concludere che solo gli esseri umani possono essere responsabili per le sue azioni. Ma chi, se e come sia responsabile per l'utilizzo dell'IA è ancora una domanda aperta. Asserita l'esistenza di un "vuoto di responsabilità", approfondiremo tre teorie che riportano al centro l'essere umano e la sua responsabilità per le azioni poste in essere dalla macchina, intesa come uno strumento.

#### 4.1 La Responsabilità elettronica

Di questa prospettiva si è fatto portatore il Parlamento Europeo, in una risoluzione del 2017. Occupandosi di individuare una forma di responsabilità civile per i danni causati dai robot suggeriva «l'istituzione di uno status giuridico specifico per i robot nel lungo termine, di modo che almeno i robot autonomi più sofisticati possano essere considerati come persone elettroniche responsabili di risarcire qualsiasi danno da loro causato, nonché eventualmente il riconoscimento della personalità elettronica dei robot che prendono decisioni autonome o che interagiscono in modo indipendente con terzi<sup>39</sup>». A partire da questa affermazione possiamo ricavare due possibili modelli, che vengono descritti da S. Fanetti<sup>40</sup> che riprende l'analisi realizzata da A. Bertolini<sup>41</sup> per il Parlamento Europeo.

*In primis* viene delineata una forma peculiare di personalità giuridica, una pura personalità elettronica. Il dispositivo di IA dovrebbe dotarsi di un patrimonio con il quale far fronte *direttamente* al risarcimento di eventuali danni provocati nel suo operare, anche nella forma di fondo assicurativo obbligatorio. La sola attribuzione della personalità senza un patrimonio sarebbe controproducente, poiché lascerebbe scoperti eventuali creditori<sup>42</sup>. Fanetti riporta due argomenti a favore di questo modello. Il primo: considerare il dispositivo IA soggetto di diritto incentiverebbe, secondo l'autore, il progresso tecnologico, liberando produttori e sviluppatori dai rischi (come accade per le società a responsabilità limitata) e allo stesso modo

---

<sup>39</sup> Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica (2015/2103(INL)), cit., par. 59, lett. f).

<sup>40</sup> S. FANETTI, *Intelligenza artificiale, personalità e responsabilità extracontrattuale*, in *Isaidat Law Review*, 2, 2022, p. 5 ss.

<sup>41</sup> A. BERTOLINI, *Artificial Intelligence and civil liability*, European Parliament Legal Affairs, Policy Department for Citizens' Rights and Constitutional Affairs Directorate-General for Internal Policies, PE 621.926 – luglio 2020.

<sup>42</sup> S. FANETTI, *op. cit.*, p. 23.

gli utilizzatori di tali strumenti tecnologici, favorendo il mercato dell'innovazione<sup>43</sup>. Chi scrive non si sente, tuttavia, di condividere quanto sostenuto dall'autore. Innanzitutto, le grandi compagnie High-tech stanno acquisendo un potere enorme di incidere sui cittadini, un potere che storicamente era stato riservato allo Stato, e che in quanto tale ha bisogno di essere regolato<sup>44</sup>. Chi scrive ritiene che nascondersi dietro allo scudo di una personalità fittizia impedirebbe ai controllori dei sistemi di IA di sopportare i rischi ricollegati alla loro attività, che è tanto pervasiva nella vita dei consumatori, quanto sfuggente per i regolatori. Per quanto riguarda gli sviluppatori, come illustreremo più avanti, si potrebbe esonerarli da qualsiasi responsabilità che non dipenda da difetti di fabbrica, per salvaguardare la ricerca e l'avanzamento tecnologico. Infine, l'utilizzo di alcune IA genera un pericolo per i consociati, ed è per questo che utenti come coloro che viaggiano con una *driverless-car*, devono sopportarne i rischi. Anche questo punto verrà approfondito nella continuazione. Queste considerazioni appartengono anche alla dottrina, difatti Ugo Ruffolo sottolinea che «responsabilizzare la sola macchina, significherebbe restringere la tutela patrimoniale dei soggetti lesi ad un solo patrimonio; quindi, potrebbe essere un modo per sfuggire a tali obblighi da parte del proprietario, del gestore o del programmatore»<sup>45</sup>.

Il secondo vantaggio dell'istituzione di una responsabilità elettronica secondo Fanetti, è che faciliterebbe la soddisfazione dell'interesse risarcitorio<sup>46</sup>, in quanto il danneggiato potrebbe individuare chiaramente il soggetto responsabile. Per quanto il ragionamento possa sembrare lineare, si pongono comunque delle problematiche. Afferma Ruffolo che una «responsabilità diretta, e con patrimonio proprio, dell'A.I. non sembra, a ben vedere, indispensabile. Essa non accresce, e potrebbe anzi ridurre, la tutela risarcitoria dei soggetti lesi, maggiormente tutelati da una responsabilizzazione di chi dovrebbe essere chiamato a rispondere per la macchina intelligente, sia esso il suo produttore, o proprietario, o "custode", o gestore. Responsabilizzare la macchina, ed essa sola, comporterebbe, di fatto, la limitazione della responsabilità patrimoniale ad uno specifico (e più ridotto) patrimonio di rischio»<sup>47</sup>. È opportuno chiedersi chi sarebbe a pagare per questo patrimonio di rischio, o per un fondo assicurativo. Inoltre, bisogna considerare che la sindacabilità delle scelte adottate dagli algoritmi ai fini dell'imputazione della responsabilità, va a braccetto con la questione della conoscibilità e della comprensibilità delle scelte stesse. Questione complicata dal problema della *black box* e dalla mancanza di nozioni per comprendere le dinamiche che presiedono ai processi decisionali seguiti<sup>48</sup>. Gli obblighi di trasparenza imposti dall'AI Act<sup>49</sup> e dal Consiglio di Stato<sup>50</sup> sono il

<sup>43</sup> S. FANETTI, *op. cit.*, p. 24.

<sup>44</sup> Si veda in proposito C. CASONATO, *L'intelligenza artificiale e il diritto pubblico comparato ed europeo*, in *DPCE online*, 1, 2022, pp. 169-179.

<sup>45</sup> U. RUFFOLO, *Il problema della "personalità elettronica"*, in *Journal of Ethics and Legal Technologies*, 2(1), 2020.

<sup>46</sup> S. FANETTI, *op. cit.*, p. 24.

<sup>47</sup> U. RUFFOLO, *op. cit.*

<sup>48</sup> A. ANTONINA ET AL., *La persona elettronica: verso un tertium genus di soggetto?*, in F. BILOTTA, F. RAIMONDI (a cura di), *Il soggetto di diritto storia ed evoluzione di un concetto nel diritto*, Napoli, 2020.

<sup>49</sup> L'art. 50 dell'AI Act prevede obblighi di trasparenza per determinati sistemi di IA che si sostanziano in dovere dei fornitori di informare gli utenti di star interagendo con un interlocutore artificiale, o che un'immagine è stata generata artificialmente. Più specifici doveri di trasparenza si individuano per i sistemi ad alto rischio, nell'art. 10 e nell'art. 13, relativi alla raccolta e all'utilizzo dei dati e ai fini di una comprensibilità del funzionamento del sistema. Cfr. anche <https://www.ai4business.it/intelligenza-artificiale/artificial-intelligence-act-come-garantire-la-trasparenza/> (ultima consultazione 17/10/2024).

<sup>50</sup> Cons. Stato, sez. VI, 8 aprile 2019 n. 2270. Secondo la Corte «il meccanismo attraverso il quale si concretizza la decisione robotizzata (ovvero l'algoritmo) deve essere "conoscibile", secondo una declinazione rafforzata del principio

primo passo per il riconoscimento e la risoluzione di una mancanza che potrà essere colmata, tuttavia, solo con un impegno nel garantire la comprensibilità dal punto di vista tecnico, ed un “investimento” in competenze di “AI literacy” da parte degli operatori.

Il secondo modello proposto dagli autori si riferisce alla seconda parte della citazione del Parlamento Europeo menzionata sopra: «riconoscimento della personalità elettronica dei robot che prendono decisioni autonome o che interagiscono in modo indipendente con terzi»<sup>51</sup>. Dunque, si potrebbe pensare di “personificare” (ossia attribuire i diritti e doveri della persona umana) la macchina in relazione al suo grado di intelligenza. Sostiene Ruffolo che tale personificazione dell’IA «sarebbe strumentale alla esigenza di tutelare quella entità quando assunta, nella sensibilità e coscienza sociale, ad una qualche dignità di “essere”». Così liberata «dallo status di mera cosa» la macchina «potrà essere presa in considerazione come soggetto cui imputare dirette responsabilità. Alla capacità di decisione autonoma e interazione comunicativa con gli umani, quale preconditione per l’attribuzione di libertà, diritti e doveri, dovrebbe corrispondere l’attitudine a comportamento “morale”»<sup>52</sup>. Escludiamo fermamente la possibilità di attribuire all’IA una personalità a immagine e somiglianza di quella di una persona fisica, in quanto non è possibile equiparare l’intelligenza umana a quella artificiale. Quest’ultima forma di “intelligenza” non segue i modelli di ragionamento tradizionalmente umani: pur arrivando a un risultato quanto più simile a quello umano, segue modelli matematici probabilistici di difficile comprensione. Non possiamo nemmeno accettare l’idea di una “macchina morale inconsapevole”, a detta di Ruffolo accettabile come agente etico. L’autore argomenta che la macchina sarebbe in grado di decidere scegliendo tra valori, apprendendo le regole sia direttamente, per effetto delle istruzioni impartite dallo sviluppatore, sia per inferenza, grazie ad una sua capacità di imparare che non dipende necessariamente dalla coscienza dei contenuti. Questo per noi è inaccettabile, visto che abbiamo considerato la consapevolezza una delle condizioni alla base dell’attribuzione di responsabilità<sup>53</sup>. Le IA possono essere considerate agenti morali nel senso che operano all’interno di un framework etico e legale definito, ma la loro responsabilità morale è relativa più all’adesione a regole e direttive predefinite che ad una capacità intrinseca di discernimento morale, che rimane una prerogativa umana. Fanetti<sup>54</sup> considera pericolosa la sopravvalutazione delle capacità dei robot anche più avanzati, che porta a considerarli non più come strumenti, ma a proiettare su di essi attributi umani. Ciò senza che possiedano altre tipiche qualità degli esseri umani, prima tra tutte la volontarietà. È anche per modellare questa volontarietà che esiste la legge, e l’effetto deterrente della sanzione, che verrebbe meno con l’equiparazione della persona elettronica.

---

di trasparenza, che implica anche quello della piena conoscibilità di una regola espressa in un linguaggio differente da quello giuridico».

<sup>51</sup> *Ibidem*.

<sup>52</sup> U. RUFFOLO, *op. cit.*, p. 85.

<sup>53</sup> *Ibid.* «Il dubbio è duplice: da un lato, se possa essere macchina morale una macchina inconsapevole; dall’altro, se possa essere possibile alla macchina una scelta fra valori». Le IA possono essere considerate agenti morali nel senso che operano all’interno di un framework etico e legale definito, ma la loro “responsabilità morale” è più una questione di adesione a regole e direttive predefinite che di capacità intrinseca di discernimento morale, che rimane una prerogativa umana.

<sup>54</sup> S. FANETTI, *op. cit.*

## 4.2 Il vuoto di responsabilità e la responsabilità diffusa

Nel momento in cui escludiamo che possa esistere una responsabilità *dell'intelligenza* artificiale, la conclusione più ovvia è che solo gli esseri umani possano essere responsabili per le sue azioni. Ma chi, se e come sia responsabile per l'utilizzo dell'IA rimane una domanda aperta.

Nel 2004 Matthias ha introdotto il concetto del "vuoto di responsabilità" o "*responsibility gap*"<sup>55</sup>. In base al comune senso di giustizia, gli umani sono ritenuti responsabili solo ove abbiano il controllo delle proprie azioni e delle conseguenze da esse derivabili, proprio come sosteneva Aristotele. Così l'agente sarà ritenuto responsabile solo ove abbia avuto consapevolezza dei fatti intorno alla sua azione, e sulla base di questi abbia potuto liberamente decidere come agire, di fronte a un set di alternative. C'è un *operator* del sistema di IA che acquisisce il controllo (condizione necessaria di responsabilità) dello stesso quando riceve dal *manufacturer* il manuale di istruzioni che precisa l'insieme di "trasformazioni" che il *device* subisce nel suo normale operare. In questo modo l'*operator* può gestire il *device* in maniera prevedibile e a lui viene ascritta la responsabilità del funzionamento della macchina, finché questa opera secondo quanto specificato dal *manufacturer*<sup>56</sup>. Laddove l'*operator* detenesse un controllo ridotto o mancasse totalmente di controllo, non potrebbe ritenersi responsabile. Ma a quel punto a chi imputare la responsabilità di un eventuale danno? Nel momento in cui accertiamo che la macchina non è responsabile delle conseguenze delle sue azioni, si crea un vuoto di responsabilità. Se la guida autonoma sottrae il controllo del veicolo al guidatore (*operator* dell'IA), chi deve ritenersi responsabile di possibili incidenti? Non tutti condividono l'idea di un vuoto di responsabilità, o meglio parte della dottrina ritiene di poterlo colmare, essendo sempre possibile identificare un essere umano che possa essere considerato responsabile dell'output della macchina, a prescindere da quanto il sistema sia indipendente o capace di auto-modificarsi<sup>57</sup>. «*Harm caused by even fully autonomous technologies is generally reducible to risks attributable to natural persons or existing categories of legal persons, and where this is not the case, new laws directed at individuals are a better response than creating a new category of legal person*»<sup>58</sup>. Pur apparendo astrattamente sempre possibile, nel pratico può non essere facile individuare un solo individuo responsabile, se consideriamo il vasto numero attori coinvolti tra design, sviluppo, distribuzione, utilizzo di un prodotto digitale, che in ultima analisi contribuiscono all'azione della macchina. Tra l'altro attori anche non necessariamente

<sup>55</sup> A. MATTHIAS, *The responsibility gap: Ascribing responsibility for the actions of learning automata*, in *Ethics and information technology*, 6, 2004.

Si veda anche J. DANAHER, *Robots, law and the retribution gap*, in *Ethics and Information Technology*, 18, 4, 2016 «How exactly do the gaps arise? The answer is that the gaps arise when the moral/legal tests for determining who should bear the burden fail to align with the reality of who is causing the injurious outcome».

<sup>56</sup> A. MATTHIAS, *op. cit.*, p. 175 «In case the machine does not operate according to the manufacturer's specification (that is, in case it has a flaw in its construction), we ascribe the responsibility to the manufacturer of the machine instead of the operator», è il caso della nostra responsabilità da prodotto difettoso.

<sup>57</sup> S. FANETTI, *op. cit.*

<sup>58</sup> EXPERT GROUP ON LIABILITY AND NEW TECHNOLOGIES – NEW TECHNOLOGIES FORMATION, *Liability for Artificial Intelligence and other emerging digital technologies*, 2019, disponibile al link [https://www.europarl.europa.eu/meetdocs/2014\\_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report\\_EN.pdf](https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf) (ultima consultazione 17/10/2024). Traduzione in lingua italiana: «I danni causati anche da tecnologie completamente autonome sono generalmente riducibili ai rischi attribuibili alle persone fisiche o alle categorie di persone giuridiche esistenti e, in caso contrario, nuove leggi dirette agli individui sono una risposta migliore rispetto alla creazione di una nuova categoria di persone giuridiche».

umani<sup>59</sup>. A tal proposito si parla di un “*Problem of many hands*”, un fenomeno per cui, data la complessità della situazione e il numero di attori coinvolti, risulta estremamente difficile, se non impossibile, ritenere qualcuno responsabile<sup>60</sup>. L’esistenza di questo problema è stata indagata in vari ambiti, come quello delle organizzazioni complesse in campo governativo<sup>61</sup>, della salute o del cambiamento climatico<sup>62</sup>, e certamente si adatta alla intricata ragnatela dell’ambito dell’innovazione. Nissembaum ne parla come una delle barriere che impediscono l’attribuzione di responsabilità in una società computerizzata<sup>63</sup>: «*Where a mishap is the work of “many hands,” it may not be obvious who is to blame because frequently its most salient and immediate causal antecedents do not converge with its locus of decision making*<sup>64</sup>.» In conseguenza di ciò la connessione tra un determinato *outcome* e il suo responsabile appare oscurata, per questo l’autrice parla di “*obscuring of accountability*”<sup>65</sup>. Il problema viene analizzato nell’ambito del *machine learning*, dove opera su quattro livelli:

1. Il software viene prodotto in contesto corporativo, dove le decisioni di sviluppo non vengono prese da un solo attore;
2. In questo contesto numerosi gruppi di ingegneri e scienziati si occupano di singoli segmenti o moduli del sistema, all’interno della cosiddetta *ML pipeline*. Spesso si servono del lavoro svolto da altri e reso disponibile come libreria open-source software<sup>66</sup>;
3. Molti sistemi vengono sviluppati proprio a partire da altri già esistenti, magari prodotti in larga scala e già allenati per essere usati in maniera generalizzata. Questi ultimi *Multi-purpose toolkits* potrebbero a loro volta essere danneggiati o mal adattarsi allo specifico obiettivo per cui vengono impiegati<sup>67</sup>;

---

<sup>59</sup> Si veda anche M. COECKELBERGH, *op. cit.*: «But there are not only many (human) hands; there are also what one could call “many things”: many different technologies. In AI process and history, various software is involved but also more literally various things, material technological artefacts: things that are relevant since they causally contribute to the technological action, and that may have some degree of agency. There are many interconnected elements. For example, a malfunctioning sensor interacting with the software of an airplane may causally contribute to its crash; it is then important to find out how the technological system as a whole is structured and who is responsible for the development, use, and maintenance of its parts (for example a sensor) and the interaction between the parts».

<sup>60</sup> I. VAN DE POEL ET AL., *The problem of many hands: Climate change as an example*, in *Science and engineering ethics*, 18, 2012.

<sup>61</sup> D.F. THOMPSON, *Responsibility for Failures of Government: The Problem of Many Hands*, in *American review of public administration*, 3, 2014, pp. 259-273.

<sup>62</sup> I. VAN DE POEL ET AL., *op. cit.*

<sup>63</sup> H. NISSEBAUM, *Accountability in a Computerized Society*, in *Science and Engineering Ethics*, 2, 1996. Si veda anche A. FEDER COOPER ET AL., *Accountability in an algorithmic society: relationality, responsibility, and robustness in machine learning*, in *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 2022. Gli autori hanno ripreso l’articolo di Nissembaum, riadattando le sue idee agli sviluppi storici avvenuti fin ora.

<sup>64</sup> H. NISSEBAUM, *op. cit.*: «Quando un incidente è opera di “molte mani”, può non essere ovvio di chi sia la colpa, perché spesso gli antecedenti causali più salienti e immediati non convergono con il luogo in cui è avvenuto il processo decisionale». L’autrice aggiunge: «The conditions for blame, therefore, are not clearly satisfied in a way normally satisfied when a single individual is held blameworthy for a harm».

<sup>65</sup> H. NISSEBAUM, *op. cit.*, p. 29.

<sup>66</sup> A. FEDER COOPER ET AL., *Accountability in an algorithmic society: relationality, responsibility, and robustness in machine learning*, in *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 2022.

<sup>67</sup> *Ibidem*.

4. Potrebbero intercorrere incompatibilità nell'operare di un software su un determinato hardware, e così diventa complicato distinguere tra la colpa della macchina e quella del sistema che vi opera all'interno.

Non bisogna pensare che la via d'uscita sia "se tutti sono responsabili, allora nessuno lo è". Questa concezione è errata e potenzialmente ingiusta, in quanto non solo trasferirebbe il peso del danno esclusivamente sulla vittima, ma potrebbe altresì consentire che le entità corporative sfruttino deliberatamente la complessità a discapito delle vittime danneggiate<sup>68</sup>. Cambiare il modo di pensare la responsabilità può essere una via, suggerita in un articolo di Luciano Floridi e Mariarosaria Taddeo nel 2016: «The effects of decisions or actions based on AI are often the result of countless interactions among many actors, including designers, developers, users, software, and hardware. This is known as distributed agency. With distributed agency comes distributed responsibility<sup>69</sup>». Tale risposta al "problema delle molte mani", invita ad uscire dalla tradizionale struttura etica, costruita intorno alla responsabilità individuale dell'essere umano, con l'obiettivo di allocare ad esso una punizione o un premio rispetto sulle sue azioni o intenzioni. Questo tipo di struttura non è adatta ad affrontare un quadro di responsabilità distribuita. Il modello proposto, invece, separa la responsabilità di un agente dalla sua volontà di agire in un certo modo, o dalla sua abilità di controllarne le conseguenze, e così considera responsabili tutti gli agenti operanti nel sistema distributivo dell'IA, distribuendo la responsabilità morale tra designers, regolatori e utilizzatori: «*In doing so, the model plays a central role in preventing evil and fostering good, because it nudges all involved agents to adopt responsible behaviors*»<sup>70</sup>. Basti approfondisce questa teoria inquadrando la responsabilità distribuita in un contesto sociale nel quale interagiscono comunicativamente «agenti comunicativi consapevoli, gli umani, e agenti comunicativi inconsapevoli, le macchine, sempre di più interconnessi e interdipendenti gli uni dagli altri»<sup>71</sup> e da questa interazione deriva la responsabilità morale delle decisioni e azioni. Perciò è cruciale instaurare un dialogo continuo tra progettisti e sviluppatori di IA, utilizzatori ed esperti di etica, per far sì che l'evoluzione delle tecnologie IA sia guidata non solo dall'innovazione tecnologica, ma anche da principi morali solidi. Dobbiamo rilevare come non ci troviamo d'accordo con questo sistema di attribuzione, che di per sé non illustra come ripartire la responsabilità tra gli attori, ove alcuni siano più responsabili di altri. Responsabilità distribuita non significa responsabilità egualmente distribuita. Per meglio comprendere, possiamo richiamarci all'incidente avvenuto in Arizona nel 2018<sup>72</sup> e ampiamente discusso nel prossimo paragrafo, dove una macchina a guida semi-autonoma (di livello 3) ha ucciso un pedone. Gli algoritmi governanti il software della macchina non avevano riconosciuto la persona che stava attraversando, poiché conduceva una bicicletta a mano fuori dalle strisce pedonali. A chi attribuire la responsabilità? Agli sviluppatori del software? All'atteggiamento troppo avanguardistico di Uber, che aveva deciso di impiegare quel tipo di veicolo, nonostante fosse ancora in fase di *testing*? Al pedone stesso per non aver rispettato il codice della strada, oppure al conducente dell'auto che avrebbe potuto riprendere il controllo di quest'ultima e rallentare ma era disattento? Si potrebbe anche individuare

---

<sup>68</sup> H. NISSEBAUM, *op. cit.*

<sup>69</sup> M. TADDEO, L. FLORIDI, *How AI can be a force for good*, in *Science*, 361, 2018, pp. 751-752.

<sup>70</sup> M. TADDEO, L. FLORIDI, *op. cit.*, pp. 751-752.

<sup>71</sup> G. BASTI, *La sfida etica dell'intelligenza artificiale e il ruolo della filosofia*, in *Aquinas*, 2, 2022, p. 299 ss.

<sup>72</sup> Si veda *infra* par. 4.3.

un parziale responsabilità dell'allora Governatore dell'Arizona, che aveva dato licenza alla circolazione e al *testing* di veicoli autonomi nello stato senza tenere conto, nella regolamentazione, di adeguate garanzie.

#### 4.3 La prospettiva relazionale e sociale di responsabilità condivisa

Sempre restando in un contesto di percezione collettiva della responsabilità, merita di essere menzionato l'approccio di carattere relazionale che si rifà alla dottrina di Bovens e Coeckelberg. Alle due condizioni aristoteliche della responsabilità sopra illustrate, Coeckelberg ne aggiunge una terza relazionale. La responsabilità va intesa come la capacità di poter rispondere, rendere conto a qualcuno (*responsibility as answerability, responsibility-to* piuttosto che *responsibility-for*). Bisogna tenere conto della relazione tra l'agente e l'altra persona: vi sarà responsabilità solo ove la relazione con l'altro sia percepita come moralmente rilevante. «*With regard to ascribing and exercising responsibility, it is important to see if all these conditions are met, that is, all conditions for both kinds of responsibility: the agent must have control over the action and know what she is doing, and the conditions should be such that there are others, or another, that are perceived, appear as morally relevant, and as standing in relation to her. These conditions ensure that she not only should be but also can act morally responsibly*<sup>73</sup>» Per esempio, nel traffico, un guidatore sarà responsabile se in controllo, quindi alla guida del veicolo, conscio dell'ambiente circostante e del funzionamento dell'auto, e capace di percepire una relazione con gli altri esseri umani (nell'auto e nello spazio circostante). Due sono i problemi che si pongono nel momento in cui introduciamo un'auto *self-driving*. I guidatori sono sostituiti da un oggetto che è percepito come capace di determinarsi in base a una sua *agency*, e così diventa più difficile sentirsi responsabili nei loro confronti. Non sono più considerati esseri umani, rispetto ai quali può sussistere una relazione moralmente rilevante<sup>74</sup>. Se immaginiamo che nel veicolo sia installato un *ethical-knob*, ciò potrebbe portare a preferire un'auto che in situazioni di emergenza preferirà un comportamento più egoistico, piuttosto che altruistico. Oltre a ciò, permane un problema di carattere epistemico: la difficoltà di comprendere e di spiegare all'altro il modo di funzionamento dell'auto stessa. La spiegabilità è importante per la condizione di conoscenza (sapere cosa si sta facendo, come e perché per essere considerato responsabile), ma anche in un'ottica di *answerability*<sup>75</sup> se pensiamo a quello che Coeckelbergh chiama *responsibility patient*<sup>76</sup>. Questi potrebbe essere il pedone investito dalla CAV, o un suo parente. «L'etica dell'IA, quindi, dovrebbe favorire lo sviluppo di un'IA che supporti la responsabilità degli agenti dell'IA (utenti, sviluppatori)» che dovrebbero «essere in grado di assumersi la responsabilità di ciò che fanno con l'IA e dovrebbero essere responsabili nel senso di rispondere alle persone interessate»<sup>77</sup>. Anche Bovens concepisce un tipo di responsabilità relazionale che si

---

<sup>73</sup> M. COECKELBERG, *Responsibility and the moral phenomenology of using self-driving cars*, in *Applied Artificial Intelligence*, 8, 2016, p. 751.

<sup>74</sup> M. COECKELBERG, *Artificial intelligence, responsibility attribution, and a relational justification of explainability*, in *Science and Engineering Ethics*, 26, 2020.

<sup>75</sup> M. COECKELBERG, *Artificial intelligence, responsibility attribution, and a relational justification of explainability*, cit., p. 10.

<sup>76</sup> M. COECKELBERG, *Artificial intelligence, responsibility attribution, and a relational justification of explainability*, cit., p. 13.

<sup>77</sup> M. COECKELBERG, *Artificial intelligence, responsibility attribution, and a relational justification of explainability*, cit. «There is the question who is responsible for something (responsibility attribution) but also the question who is

basa proprio su questo: il dovere di un *responsible actor* di spiegare e giustificare le proprie azioni ad un determinato *forum*<sup>78</sup>. Per quanto la relazione tra *forum* e *responsible actor* abbia un carattere più istituzionale e rigido (poiché i diversi *forum* di riferimento hanno origine da relazioni umane di tipo gerarchico o comunque individuate), si può fare un parallelo tra il *responsibility patient* e il *responsible* di Coeckelbergh per sottolineare il fatto che entrambi gli autori danno grande valore alla trasparenza. Secondo Bovens la trasparenza è necessaria al *forum* per poter ottenere informazioni in base alle quali esercita il suo potere di interrogazione sul *responsible actor*, e realizzare l'attribuzione di responsabilità<sup>79</sup>. Per Coeckelbergh «l'obiettivo primario è la spiegabilità come responsabilità da parte dell'uomo che usa e sviluppa l'IA», allora l'innovazione dovrebbe procedere nel senso di garantire questo tipo di responsabilità relazionale, che in diritto può tradursi in un vero e proprio diritto alla spiegazione, non solo all'informazione<sup>80</sup>.

#### 4.4 Il ritorno all'idea di Strict Liability

Come risulta evidente al termine di questa seppur breve disamina, non vi è nella dottrina una visione unitaria sul tema della responsabilità dell'IA, a partire dalla sua stessa considerazione come agente morale. Abbiamo visto che chi vi ricollega una tale capacità spinge per il riconoscimento di una personalità legale ad hoc. Secondo altri la totale assenza di essa rende impossibile individuare un responsabile per le conseguenze di ogni *outcome*, lasciandoci in una situazione di vuoto, da cui altri ancora propongono di uscire proprio abbandonando la prospettiva individualistica di responsabilità, per scoprirne una nuova di carattere collettivo e diffuso. Questa analisi ha contribuito a mettere in luce la rilevanza per i giuristi dell'interrogarsi sugli aspetti etici delle questioni che si apprestano a regolare, spesso sottovalutata. Proprio nel campo dell'Intelligenza Artificiale, dove molti aspetti sono ancora inesplorati, non si può prescindere dalla considerazione etica del tema, che abbiamo visto porsi a base per una conclusione giuridica consapevole. A dimostrazione di ciò concludiamo questa parte esponendo la teoria che riteniamo più confacente alla nostra "risposta" giuridica al dilemma, ossia quella che prende in considerazione singoli soggetti cui attribuire la responsabilità del verificarsi di incidenti. Per farlo ci basiamo principalmente sull'elaborazione di Hevelke e Nida-Rümelin proprio sui veicoli autonomi<sup>81</sup>.

In attesa di un intervento del legislatore in materia, che potrebbe anche portare a rielaborare la concezione della responsabilità per l'IA, il criterio che nell'imminenza può soccorrere è quello dell'analogia. Infatti, parte della dottrina etica ha voluto richiamare l'approccio giuridico del sistema della c.d. "*strict liability*",

---

responsible to whom. Responsibility is not only about doing something and knowing what you're doing; it also means answerability. It is also a relational and communicative, perhaps even dialogical matter».

<sup>78</sup> M. BOVENS, *Analysing and assessing accountability: A conceptual framework*, in *European law journal*, 4, 2007, p. 450. L'autore definisce il *forum* come il secondo componente della relazione di responsabilità, insieme all'attore. In questa visione l'attore ha degli obblighi nei confronti del *forum*, riconosciuto come "significant other". L'autore fornisce una definizione di responsabilità in senso relazionale: «Accountability is a relationship between an actor and a forum, in which the actor has an obligation to explain and to justify his or her conduct, the forum can pose questions and pass judgement, and the actor may face consequences», p. 450.

<sup>79</sup> M. BOVENS, *op. cit.*

<sup>80</sup> M. COECKELBERG, *Artificial intelligence, responsibility attribution, and a relational justification of explainability*, cit.

<sup>81</sup> A. HEVELKE, J. NIDA-RÜMELIN, *Responsibility for crashes of autonomous vehicles: An ethical analysis*, in *Science and engineering ethics*, 21, 2015, pp. 619-630.

ritenendo che dietro ogni decisione e azione di un ente artificiale si trovi sempre un essere umano responsabile, e laddove questi non sia immediatamente individuabile, ci si potrebbe richiamare alla responsabilità oggettiva, legata alla posizione di controllo o di vantaggio della figura astrattamente considerata<sup>82</sup>. Tale criterio rappresenterebbe anche una risposta al problema delle molte mani, perché sposterebbe il peso del danno dalla vittima a coloro che si trovavano nella posizione migliore per prevenirlo e mitigarlo<sup>83</sup>. L'analisi di Hevelke e Nida-Rümelin si concentra, allora, sui singoli individui coinvolti, a partire dall'esclusione della responsabilità del c.d. "manufacturer": «If in the case of crashes involving autonomous vehicles the main responsibility were to be that of the manufacturers, the liability burden on the manufacturer may be prohibitive of further development»<sup>84</sup>. Si parte da una domanda: si dovrebbe immaginare una forma di responsabilità per i veicoli autonomi tale da promuovere il loro continuo sviluppo e miglioramento? Se pensiamo vi siano buone ragioni morali (come la possibilità di guida di persone disabili e anziani, e la diminuzione degli incidenti stradali<sup>85</sup>) che permettano di ritenere desiderabile la produzione di queste automobili, allora lo Stato dovrebbe essere moralmente obbligato a tracciare una forma di responsabilità extracontrattuale che non ostacoli lo sviluppo di questa tecnologia. Un'altra prospettiva considera anche l'aumentare delle controversie che coinvolgerebbero i produttori, difficili da sostenere per i piccoli centri di ricerca, che si vedrebbero tagliati fuori dalla "technology race"<sup>86</sup>.

Questo ragionamento può ben essere generalizzato: abbiamo sostenuto precedentemente che per tutelare la figura dello sviluppatore di intelligenza artificiale dai rischi di possibili danni non è necessario introdurre una forma di responsabilità della macchina, ma semplicemente escludere quella propria di chi lavora alla programmazione, in considerazione di un interesse superiore all'innovazione tecnologica. Ovviamente permane una responsabilità dei programmatori in caso di difetto di produzione, nel momento in cui l'operator non riesce a far funzionare la macchina come da istruzioni del *manufacturer*. Infine, si arriva alla responsabilità dell'utilizzatore del sistema di IA, il "guidatore" dell'auto a guida autonoma. Se da una parte

---

<sup>82</sup> Peralto S. FANETTI, *op. cit.* sottolinea che anche di fronte ad eventuali vuoti normativi, l'attuale regime della responsabilità, potrebbe essere ben adattato per colmarli andando ad attribuire la responsabilità all'operatore nella posizione migliore per far fronte al rischio, «con, ad esempio, la predisposizione di un adeguato sistema di responsabilità 'da posizione' (F. CAROCCIA, *Soggettività giuridica dei robot?*, in G. ALPA (a cura di), *Diritto e intelligenza artificiale*, Roma, 2020). In questo senso, l'idea in sostanza è che la responsabilità dovrebbe essere attribuita all'operatore che è nella posizione migliore per far fronte a qualsiasi potenziale rischio (COMMISSIONE EUROPEA, *Libro bianco sull'intelligenza artificiale*, cit., p. 25; G. SINHA, R. DUNBAR, *Artificial Intelligence and its Regulation in the European Union*, in D.M. BIELICKI (a cura di), *Regulating Artificial Intelligence in Industry*, Abingdon-New York, p. 10)».

<sup>83</sup> A. F. COOPER ET AL., *op. cit.*

<sup>84</sup> A. HEVELKE, J. NIDA-RÜMELIN, *op. cit.*

<sup>85</sup> A. HEVELKE, J. NIDA-RÜMELIN, *op. cit.* «The consequentialist tries to reduce the diversity of morally relevant factors including duties, rights, principles etc. to one single principle: one should maximize the good. There are many reasons to reject such a reductionist scheme. (Nida-Rümelin 1995) However, a teleological rationale can still provide powerful moral reasons, especially when the good to be promoted is as important as in this case. Death and injury caused by accidents are an evil, and the protection of its citizens from such harm is a central task of the state. If the introduction of autonomous vehicles might reduce the yearly toll in death and injury exacted by road traffic even by a small degree, that would constitute a powerful moral reason in favour of promoting their development— which includes trying to design car manufacturers tort liability in such a way that the development and improvement of autonomous vehicles will be helped along. This depends, however, on there not being a group of people whose risks of injury are bound to be raised by the introduction of autonomous cars. If there was to be such a group, this could pose a major ethical problem for the introduction of autonomous cars».

<sup>86</sup> M. UZAIR, *Who is liable when a driverless car crashes?*, in *World Electric Vehicle Journal*, 2, 2021.

può concepirsi l'esistenza di un dovere di intervenire per evitare incidenti come base dell'attribuzione di responsabilità, con l'aumentare del livello di autonomia (quando raggiungiamo il quinto o il sesto livello) la possibilità di un affrettato intervento del guidatore potrebbe diminuire, se non anche portare a più incidenti. Infatti, anche ammettendo che una forma estrema di intervento fosse possibile, si può credere che l'utilizzo di una *self-driving car* porti ad una diminuzione globale della concentrazione sull'ambiente del traffico e della capacità di reazione<sup>87</sup>. Se dunque un intervento efficace non è sicuro, non si può rimproverare ad una persona di non aver agito<sup>88</sup>, e così arriviamo ad elaborare la responsabilità del guidatore come una forma di *strict liability* fondata piuttosto sul pericolo che l'introduzione di un tale veicolo all'interno della vita sociale può comportare. La *ratio* sarebbe proprio che il soggetto si è assunto il rischio di usare quel veicolo, essendo conscio e accettando il fatto che possa causare incidenti: «*Using a car poses a risk—for the person himself and for others. The more we use cars (especially where it is not necessary), the more we put others at risk—even if we do our best to drive safely*»<sup>89</sup>. Anche se l'utilizzo di veicoli automatizzati ridurrà il rischio di incidenti, residuerà un numero significativo di danneggiati. Per questo non appare implausibile considerare l'utente almeno in parte responsabile degli incidenti che possono essere causati dal suo veicolo. In fondo i proprietari sono coloro che detengono il controllo finale della macchina, per cui la responsabilizzazione è un bene da un punto di vista sociale.

Critiche dal punto di vista morale a questa teoria muovono dal fatto che la responsabilità del guidatore troverebbe una base soltanto ad un caso fortuito<sup>90</sup>. Anche il guidatore più coscienzioso può trovarsi coinvolto in un incidente attribuibile all'azione della sua auto a guida autonoma, e ritenuto responsabile nonostante non abbia avuto alcuna colpa. Ciò ci mostra come una prospettiva può essere considerata adeguata a bilanciare gli interessi in gioco, ma ingiusta da un punto di vista morale. È importante continuare ad alimentare la discussione per arrivare alla soluzione globalmente migliore.

Se moralmente viene criticata l'ingiustizia di addebitare un obbligo risarcitorio qualcuno senza che sia dimostrata la sua colpa, d'altra parte questa eventualità fa sì che la persona interessata dimostri una maggiore cura verso quel bene pericoloso<sup>91</sup>, magari informandosi attentamente sul funzionamento, prestando maggiore attenzione alla strada, e prendendo ogni accorgimento tecnico. Inoltre, si può ritenere che tale tipo di responsabilità degli utilizzatori porterebbe ad un maggior impegno per la sicurezza anche dei produttori. Concludiamo precisando che la formula presentata a proposito delle CAVs, potrebbe non essere ideale per la generalità dei sistemi di IA. Pensiamo, ad esempio, al campo medico: il dottore che lavora presso l'ospedale è l'utilizzatore del dispositivo IA, ma non si trova nella stessa posizione del proprietario, e perciò non può esservi parificato. Altrettanto si potrebbe pensare per particolari sistemi di IA ad alto rischio. È probabile che come si è già fatto per altri requisiti nell'AI Act, quello nei confronti della responsabilità si caratterizzi come un approccio "*tailored*" allo specifico tipo di IA di cui si tratta.

---

<sup>87</sup> A. HEVELKE, J. NIDA- RÜMELIN, *op. cit.*

<sup>88</sup> A. HEVELKE, J. NIDA- RÜMELIN, *op. cit.* «It is wrong to publicly accuse someone as having caused the death of another person with his inattention if he never had a real chance to do what was supposedly his "duty"».

<sup>89</sup> A. HEVELKE, J. NIDA- RÜMELIN, *op. cit.*

<sup>90</sup> Si veda in proposito A. KAUPPINEN, *Who Should Bear the Risk When Self-Driving Vehicles Crash?*, in *Journal of Applied Philosophy*, 4, 2021; M. UZAIR, *Who is liable when a driverless car crashes?*, cit.; J. DANAHER, *Robots, law and the retribution gap*, cit.

<sup>91</sup> H. NISSEBAUM, *op. cit.*

## 5. Il caso studio: l'incidente a Tempe del 2018

Analizzare un incidente realmente verificatosi negli Stati Uniti, stato con cui risulta particolarmente utile la comparazione considerata la maggiore diffusione dei veicoli autonomi, ci consente di esaminare le regole di responsabilità nel loro concreto funzionamento comprendendole al meglio. Il 18 marzo 2018, Elaine Herzberg veniva investita, mentre attraversava la strada di notte con la bici a mano al di fuori delle strisce pedonali, da un'auto di Uber che circolava condotta dal sistema di guida autonoma, inserito, al momento dell'incidente, da circa 20 minuti. Il sistema non arrestava il veicolo, non percependo la presenza del pedone o, secondo talune ricostruzioni, avendolo erroneamente classificato come veicolo, capace di uscire dalla traiettoria prima del suo passaggio<sup>92</sup>. I problemi da analizzare sono molteplici. Il possibile malfunzionamento del sistema, potenzialmente viziato, apre lo scenario a una responsabilità del produttore anche se nella decisione della Corte statunitense non se ne trova traccia<sup>93</sup>. Inoltre, non può essere trascurato il problema, già accennato, della difficoltà per i sistemi autonomi di funzionare correttamente al di fuori del loro *envelope* ottimale: il rispetto del Codice della Strada da parte di tutti gli utenti è un elemento essenziale affinché le *driverless car* possano effettivamente ridurre gli incidenti e non provocarne di diversi. Infine, un elemento centrale da considerare è il ruolo ricoperto dalla conducente: dal video registrato dalla telecamera interna all'abitacolo risulta chiaramente come questa non stava prestando attenzione alla strada, bensì guardando in basso, probabilmente il cellulare, tanto che le mani non risultavano nemmeno posizionate sul volante. Il caso giudiziario si conclude con la sentenza di patteggiamento con cui la conducente accetta la condanna per omicidio colposo<sup>94</sup>. La scelta difensiva è stata probabilmente corretta, anticipando un esito già scritto: è innegabile che la conducente, distraendosi, non avesse rispettato l'obbligo di vigilanza sulla guida autonoma, versando in stato di colpa. Come anticipato nell'analisi dei livelli di automazione, infatti, i veicoli oggi in circolazione appartenenti al secondo e al terzo livello richiedono una costante supervisione umana poiché, se le circostanze lo richiedono, il conducente dev'essere in grado di riprendere immediatamente il controllo del veicolo.

Tuttavia, possiamo fare un'ulteriore considerazione sulla conclusione del caso esaminato. Se è pacifico che la conducente abbia violato il dovere di sorveglianza, analizzando le circostanze concrete è possibile dubitare che anche prestando attenzione la guidatrice fosse in grado di evitare l'incidente. La vittima, infatti, al momento dell'incidente, stava attraversando una strada a quattro corsie, non illuminata, se non dai fari dell'auto, e al di fuori delle strisce pedonali. Dovremmo pertanto chiederci se, qualora lo stesso incidente fosse stato causato da un'auto a guida integralmente manuale, la decisione sarebbe stata la medesima. La risposta a tale quesito necessita della preliminare analisi degli standard di diligenza richiesti al conducente, che inevitabilmente cambiano nei diversi ordinamenti.

Se prendiamo come riferimento l'ordinamento giuridico italiano, probabilmente l'esito del giudizio sarebbe stato il medesimo. L'art. 2054 co. 1 cc stabilisce una presunzione di colpa in capo al conducente per i danni

---

<sup>92</sup> La ricostruzione dell'incidente è riportata dalle più importanti testate <https://www.bbc.com/news/technology-54175359> (ultima consultazione 29/10/2024); <https://www.reuters.com/article/usa-uber-idFRKBN1XU0IC/> (ultima consultazione 29/10/2024); <https://eu.azcentral.com/story/money/business/tech/2018/03/22/what-went-wrong-uber-volvo-fatal-crash-tempe-technology-failure/446407002/> (ultima consultazione 29/10/2024).

<sup>93</sup> L'accordo di patteggiamento è disponibile sul sito internet del Maricopa County Attorney: <https://www.maricopacountyattorney.org/DocumentCenter/View/2780/Rafaela-Vasquez-Plea-Agreement> (ultima consultazione 29/10/2024).

<sup>94</sup> *Ibidem*.

arrecati a causa della circolazione del veicolo di cui era alla guida, se non prova di «aver fatto tutto il possibile per evitare il danno». In altri termini, il conducente deve fornire la prova di aver osservato, nei limiti della normale diligenza, un comportamento esente da colpa e conforme alle regole del Codice della Strada, ponendo in essere ogni attività ed accorgimento, comprese le manovre di emergenza, idonei a impedire il danno, da valutarsi con riferimento alle circostanze del caso concreto<sup>95</sup>. Tuttavia, analizzando la giurisprudenza in materia, si può osservare che lo standard di diligenza richiesto al conducente non è sempre il medesimo. In particolare, la dottrina parla di colpa relazionale, intendendo con tale sintagma la graduazione della diligenza richiesta in base alla categoria cui appartengono i soggetti coinvolti<sup>96</sup>. In caso di investimento di un pedone si richiede uno standard particolarmente elevato poiché questo è soggetto debole della circolazione. In questa ipotesi ci si avvicina molto a una responsabilità oggettiva poiché il conducente di fatto sarà sempre responsabile, pur se diligente<sup>97</sup>: questo persino a fronte di un comportamento colposo del danneggiato, a meno che non riesca a provare un comportamento del tutto imprevedibile del pedone, considerabile alla stregua del caso fortuito, interruttivo del nesso di causalità<sup>98</sup>. Applicando tale standard al caso *de quo* è possibile immaginare una condanna della conducente.

Valutazione almeno parzialmente diversa viene effettuata negli Stati Uniti dove il pedone che attraversa al di fuori delle strisce pedonali, e dunque tiene un comportamento imprudente, viene gravato delle conseguenze di tale comportamento, imponendogli di prestare massima attenzione nel compiere tale attraversamento rischioso e addossandogli la responsabilità in caso di danno. Nella maggior parte degli Stati<sup>99</sup>, infatti, i pedoni vengono sanzionati per l'attraversamento al di fuori delle zone consentite e di tale comportamento colposo si terrà conto in fase di determinazione della responsabilità. Una corretta analisi della questione non può tralasciare il fatto che le regole di responsabilità sono diverse nei 50 Stati americani: molti applicano la regola della *comparative fault*, in base alla quale il risarcimento dipende dal grado di responsabilità di ciascuna parte nella causazione dell'incidente, da determinarsi in base alla colpa di ciascuno; altri Stati, in particolare Alabama, Maryland, North Carolina, Virginia e Washington DC, applicano la più rigida regola della *contributory fault*, detta anche regola dell'1% poiché è sufficiente un

<sup>95</sup> Cfr. Cassazione penale, Sezioni Unite, sentenza n. 20137 del 13 luglio 2023; Cassazione civile, Sez. III, sentenza n. 100031 del 29 aprile 2006; Cassazione civile, Sez. II, sentenza n. 7575 del 9 luglio 1991; Cassazione penale, sentenza n. 3439 del 21 luglio 1989.

<sup>96</sup> F. CAFAGGI, *Profili di relazionalità della colpa. Contributo ad una teoria della responsabilità extracontrattuale*. Padova, 1996, p. 140.

<sup>97</sup> Cfr. C. CASTRONOVO, *Responsabilità civile*. Milano, 2018, p. 754; F. GALGANO, *I fatti illeciti*. Padova, 2008, p. 130; F. GALGANO, *Trattato di diritto civile*, II, Padova, 2009, p. 1007; M. FRANZONI, *Colpa presunta e responsabilità del debitore*. Padova, 1988, p. 152; M. FRANZONI, *L'illecito* (2. ed.), Milano, 2010, p. 613; S. RODOTÀ, *Il problema della responsabilità civile*. Milano, 1967, p. 161.

<sup>98</sup> Cfr. Cassazione civile, Sez. III, sentenza n. 20140 del 13 luglio 2023, Cassazione civile, Sez. III, sentenza n. 8663 del 4 aprile 2017; e Cassazione civile, Sez. VI, sentenza n. 4551 del 22 febbraio 2017.

<sup>99</sup> Fa eccezione, ad esempio, la California in cui il Freedom to Walk Act (Assembly Bill 2147), in vigore dal 1 gennaio 2023, afferma la libertà per il pedone di attraversare liberamente la strada a meno che non vi sia un pericolo immediato di collisione con un veicolo. Testo disponibile al link [https://leginfo.ca.gov/faces/billNavClient.xhtml?bill\\_id=202120220AB2147](https://leginfo.ca.gov/faces/billNavClient.xhtml?bill_id=202120220AB2147) (consultato il 25/7/2024). Per approfondire, <https://sepulvedalawgroup.com/whos-at-fault-when-a-pedestrian-is-hit-while-jaywalking/> (ultima consultazione 29/10/2024).

contributo colposo minimo, pari anche solo all'1% perché il risarcimento sia totalmente escluso<sup>100</sup>. Nell'incidente da noi esaminato, considerando la condotta fortemente negligente del pedone, potremmo dubitare che, nel caso in cui si fosse trattato di un veicolo a guida completamente umana, il verdetto sarebbe stato favorevole alla vittima. La considerazione che dunque ci sembra di poter fare riguarda le diverse aspettative che si nutrono in relazione alla guida umana, di cui si ammette la fallibilità, e alla guida autonoma, dalla quale si pretende la perfezione. Se i due modelli di guida venissero valutati con lo stesso criterio non ci sarebbe dubbio sulla maggiore sicurezza della guida autonoma: non sembra corretto il ragionamento perpetrato dai detrattori di tale tecnologia, che ne vorrebbero bandire l'utilizzo facendo leva sulla sua mancanza di affidabilità, per il semplice fatto che viene preso come riferimento uno standard diverso. Piuttosto, la particolare pericolosità delle auto semi-autonome si spiega in un fenomeno psicologico, il cd. *automation bias* o problema della *complacency*. Questo consiste nell'eccessivo affidamento che l'essere umano tende a riporre nel funzionamento della macchina, anche laddove sarebbe prescritta una verifica perché è accertato che la tecnologia possa sbagliarsi. Il meccanismo psicologico che porta a presumere il sistema funzionante anche in una situazione di pericolo, porta l'uomo ad assumere un livello di attenzione nettamente inferiore rispetto a quello necessario quando l'attività è integralmente demandata a lui. Ne consegue che i veicoli oggi in circolazione, di livello 2 o 3, sono in realtà i più pericolosi poiché il conducente è portato a delegare completamente la guida a un sistema non in grado di gestire qualsiasi situazione e che viceversa necessita di una costante sorveglianza<sup>101</sup>.

### 5.1 La responsabilità per i danni cagionati dalle auto completamente driverless

Se la responsabilità per i danni cagionati dalle auto dotate di sistemi ausiliari alla guida continua a ruotare attorno alla figura del conducente, altrettanto non si può dire per i danni causati dalle *driverless car* che, come dice il termine stesso, sono sprovviste del conducente. Si tratta dei veicoli di livello 4 e 5, entrambi veicoli autonomi, sprovvisti dei comandi di guida, che si differenziano perché quelli di livello inferiore sono capaci di muoversi in autonomia solo in determinate condizioni, si pensi ad esempio a un taxi locale, abilitato a svolgere solo il percorso aeroporto-centro città, mentre quelli di livello superiori sono autonomi in ogni circostanza.

Nel nostro ordinamento la soluzione potrebbe rinvenirsi nel terzo comma dell'art. 2054 cc, applicato analogicamente alle auto autonome. La norma afferma la responsabilità oggettiva del proprietario per i

---

<sup>100</sup> Cfr. J.D. CURTIS LEE, *What If a Pedestrian Is Hit While Jaywalking?*, disponibile al link <https://www.nolo.com/legal-encyclopedia/what-if-a-pedestrian-is-hit-while-jaywalking.html> (ultima consultazione 29/10/2024); J.D. HAWKINS, *What Happens if You Hit a Jaywalker?*, disponibile al link <https://www.hawkfirm.com/what-happens-if-you-hit-a-jaywalker/> (ultima consultazione 29/10/2024); SIMEONE&MILLER, *Is a Jaywalker liable for an accident?*, disponibile al link <https://www.simeonemiller.com/blog/is-a-jaywalker-liable-for-an-accident-simeone-miller-llp/> (ultima consultazione 29/10/2024); C. ARNOLD, *How could the new jaywalking laws in California affect pedestrian accident claims?*, disponibile al link <https://www.justice4you.com/blog/effects-of-legal-jaywalking-on-accident-claims.html> (ultima consultazione 10/08/2024).

<sup>101</sup> La letteratura sul tema è vasta. V.R. PARASURAMAN, D.H. MANZEY, *Complacency and Bias in Human Use of Automation: An Attentional Integration*, in *Human Factors*, 3, 2010, p. 381; K.M. WILSON, S. YANG, T. ROADY, J. KUO, M.G. LENNÉ, *Driver Trust & Mode Confusion in an On-Road Study of Level-2 Automated Vehicle Technology*, in *Safety Science*, 5, 2020; J. GASPAR, C. CARNEY, *The Effect of Partial Automation on Driver Attention: A Naturalistic Driving Study*, in *Human Factors*, 8, 2016, p. 1261; V.A. BANKS, A. ERIKSSON, J. O'DONOGHUE, N.A. STANTON, *Is Partially Automated Driving a Bad Idea? Observations From an On-Road Study*, in *Applied Ergonomics*, 68, 2018, p. 140.

danni causati dalla circolazione del suo veicolo, in quanto soggetto dotato dell'effettiva disponibilità giuridica e del potere di utilizzo esclusivo del veicolo, ovvero della possibilità di escludere terzi dall'uso dello stesso. È dunque possibile rinvenire l'*eadem ratio* che consente l'applicazione analogica: il proprietario del veicolo è chiamato a rispondere in virtù del nesso che lo lega al mezzo, pur essendo qualcun altro a condurre il veicolo. Una situazione non dissimile ci sembra si verifichi quando a condurre il veicolo non sia un altro conducente umano ma un pilota automatico, un'intelligenza artificiale.

## 5.2 La responsabilità del produttore

Infine, dev'essere valutata la possibilità che il danno sia causalmente riconducibile a un difetto del veicolo e dunque a essere responsabile sia (anche) il produttore. Anche in questo caso è interessante la comparazione tra la disciplina che si potrebbe applicare in Italia, e dunque a livello europeo, considerato che la disciplina della responsabilità del produttore è frutto del recepimento della direttiva comunitaria 85/374/CEE, e quella statunitense. Innanzitutto, bisogna precisare che il quarto comma dell'art. 2054 cc prevede la responsabilità oggettiva solidale di conducente e proprietario del veicolo anche quando il danno sia derivato da vizi di costruzione o da difetto di manutenzione. La previsione però non elimina la possibile responsabilità del produttore quando il veicolo messo in circolazione sia difettoso, ossia non offre la sicurezza che ci si può legittimamente attendere (art. 117 cod cons). Due sono gli aspetti che crediamo debbano essere analizzati particolare attenzione. *In primis* bisogna soffermarsi sulle circostanze che incidono sulla valutazione di difettosità e, in particolare, la sua presentazione nel mercato, nonché le istruzioni e avvertenze fornite al consumatore. La seconda questione, su cui inoltre si annida la differenza tra l'ordinamento italiano e quello statunitense, è il difficile rapporto tra sicurezza e difettosità.

Le modalità con cui il veicolo viene pubblicizzato, ovvero quello che nel testo dell'art. 117 cod cons viene definito «modo in cui il prodotto è stato messo in circolazione, la sua presentazione [...], le istruzioni e le avvertenze», rivestono una speciale rilevanza. È innegabile, infatti, che una commercializzazione aggressiva che esalta le capacità di guida autonoma del veicolo rischia di inasprire il cd *automation bias*, rendendo il guidatore che dovrebbe prestare attenzione ancora più disattento. Esempio di questo fenomeno può rinvenirsi nella pubblicità ingannevole portata avanti dal colosso statunitense Tesla, nei confronti del quale sono state intentate diverse *class action* per la pubblicità capziosa del programma *Autopilot*<sup>102</sup>, rispetto al quale il conducente non veniva sufficientemente messo in guardia sulla necessaria costante supervisione durante il suo funzionamento<sup>103</sup>. Rispetto a un veicolo presentato come integralmente autonomo, l'aspettativa del grande pubblico è di essere trasportati in sicurezza da una macchina in grado di guidarsi da sé, senza la necessità di supervisione, elemento sottaciuto ma pur sempre necessario. Dall'altro lato, il produttore per andare esente da responsabilità deve dimostrare la ricorrenza di una delle esimenti di cui

<sup>102</sup> In particolare, si veda il caso *Nachman v. Tesla, Inc.*, 22-CV-5976 (RPK) (ST), (E.D.N.Y. Sep. 30, 2023).

Inoltre, il Dipartimento della motorizzazione della California (California DMV) il 28 luglio 2022 ha presentato reclamo, accusando Tesla di pubblicità ingannevole. Disponibile al sito [https://regmedia.co.uk/2022/08/08/tesla\\_dealer\\_complaint.pdf](https://regmedia.co.uk/2022/08/08/tesla_dealer_complaint.pdf) (ultima consultazione 29/10/2024).

<sup>103</sup> <https://www.automobilismo.it/tesla-class-action-per-pubblicita-ingannevole-32132> (ultima consultazione 29/10/2024), e <https://www.motorionline.com/elon-musk-accusato-di-pubblicita-ingannevole-riguardo-allautonomia-delle-sue-vetture/> (ultima consultazione 29/10/2024).

all'art. 118 cod cons<sup>104</sup>. La lett. e) prevede che il produttore non sia responsabile «se lo stato delle conoscenze scientifiche e tecniche, al momento in cui il produttore ha messo in circolazione il prodotto, non permetteva ancora di considerare il prodotto come difettoso» e pertanto il cd. rischio da sviluppo grava sul proprietario del bene. Ci si chiede pertanto se il rispetto della normativa in materia di sicurezza dei prodotti sia sufficiente a integrare tale esimente. La Corte di Giustizia, conformemente a quanto emerge dal considerando 36 della direttiva 2001/95/CE<sup>105</sup> in materia di sicurezza dei prodotti, ha chiaramente affermato che l'area della *product safety* e quella della *product liability* sono nettamente separate<sup>106</sup>. Il rispetto dei requisiti di sicurezza è funzionale alla messa in commercio del prodotto ma non è sufficiente a provare che il danno non era prevedibile allo stato delle conoscenze scientifiche e tecniche al loro livello più avanzato. Pertanto, l'osservanza della normativa in materia di sicurezza è condizione necessaria ma non sufficiente per la non difettosità del prodotto: se gli standard non sono stati rispettati il prodotto sarà sicuramente difettoso, ma la loro osservanza non garantisce la non difettosità del prodotto. Inoltre, sicurezza e difettosità divergono anche sotto il profilo della valutazione che deve essere compiuta: la sicurezza viene valutata *ex ante* in relazione alla sussistenza di requisiti prefissati propri del prodotto astrattamente considerato, viceversa l'esistenza del difetto viene valutata *ex post* in relazione alla vita concreta di quel prodotto<sup>107</sup>. Tale netta separazione si impone anche ragionando *a contrario*. Sovrapporre i due ambiti comporterebbe l'esclusione della responsabilità del produttore per tutti i prodotti soggetti a omologazione, ovvero quei prodotti che per essere immessi nel mercato devono rispettare determinati standard di sicurezza, tra cui anche le automobili. Si tratta chiaramente di una conclusione inaccettabile.

---

<sup>104</sup> «La responsabilità è esclusa:

- a) se il produttore non ha messo il prodotto in circolazione;
- b) se il difetto che ha cagionato il danno non esisteva quando il produttore ha messo il prodotto in circolazione;
- c) se il produttore non ha fabbricato il prodotto per la vendita o per qualsiasi altra forma di distribuzione a titolo oneroso, né lo ha fabbricato o distribuito nell'esercizio della sua attività professionale;
- d) se il difetto è dovuto alla conformità del prodotto a una norma giuridica imperativa o a un provvedimento vincolante;
- e) se lo stato delle conoscenze scientifiche e tecniche, al momento in cui il produttore ha messo in circolazione il prodotto, non permetteva ancora di considerare il prodotto come difettoso;
- f) nel caso del produttore o fornitore di una parte componente o di una materia prima, se il difetto è interamente dovuto alla concezione del prodotto in cui è stata incorporata la parte o materia prima o alla conformità di questa alle istruzioni date dal produttore che la ha utilizzata».

<sup>105</sup> «La presente direttiva non dovrebbe interferire con i diritti delle vittime di cui alla direttiva 85/374/CEE del Consiglio, del 25 luglio 1985, relativa al ravvicinamento delle disposizioni legislative, regolamentari ed amministrative degli Stati membri in materia di responsabilità per danno da prodotti difettosi».

<sup>106</sup> Cfr. E. AL MUREDEN, *Il danno da prodotto conforme tra responsabilità per esercizio di attività pericolosa ed armonizzazione del diritto dell'Unione Europea*, in *Il corriere giuridico*, 5, 2020, p. 686.

<sup>107</sup> Cfr. A. AMIDEI, *Intelligenza artificiale e product liability: sviluppi del diritto dell'Unione Europea*, in *Giur. It.*, 7, 2019, p. 1722; G. F. SIMONINI, *L'approccio olistico nel danno da prodotto*, in *Danno e responsabilità*, 2, 2018, p. 139; C. AMATO, *Product liability and product security: Present and future*, in S. LOHSSE, R. SCHULZE, D. STAUDENMAYER (a cura di), *Liability for artificial intelligence and the internet of things*, Baden-Baden, 2019, p. 79; R. MONTINARO, *Difetto del prodotto, norme tecniche ed informazione sui rischi di danno alla salute*, in *Responsabilità civile e previdenza*, 1, 2019, p. 278; E. BELLISARIO, *Il danno da prodotto conforme tra regole preventive e regole risarcitorie*, in *Europa e diritto privato*, 3, 2016, p. 841.

Negli Stati Uniti l'approccio è opposto<sup>108</sup>. La definizione degli standard federali di sicurezza<sup>109</sup> è di centrale importanza poiché, oltre a consentire una diffusione meno rischiosa dei veicoli autonomi, fissa il livello di responsabilità dei produttori. Secondo la cd. *preemption doctrine*, infatti, il rispetto degli standard federali di sicurezza delineati dalla NHTSA impedisce il configurarsi di una responsabilità in capo al produttore<sup>110</sup>. Questo principio che la giurisprudenza trae dall'interpretazione del *National Traffic and Motor Vehicle Safety Act* del 1966 in relazione ai veicoli tradizionali<sup>111</sup> dovrebbe estendersi anche ai veicoli automatizzati. La diretta conseguenza di tale interpretazione, molto favorevole al produttore è responsabilizzare il proprietario-utilizzatore su cui da ultimo viene a gravare il costo risarcitorio dei danni. Parte della dottrina ritiene invece che i costi dei nuovi incidenti causati dall'automazione della circolazione stradale debbano essere socializzati<sup>112</sup>, evitando di rendere la disciplina eccessivamente gravosa per il proprietario-utilizzatore. Una delle proposte consiste nel classificare la guida autonoma come *ultrahazardous activity*. In relazione a queste attività il legislatore predispone discipline particolarmente rigorose volte ad assicurare uno svolgimento ragionevolmente sicuro<sup>113</sup> e che al contempo limitano l'operatività della *preemption doctrine*, facendo persistere la responsabilità del produttore nonostante il rispetto di tutte le regole preventive<sup>114</sup>. Questa interpretazione non è così distante da quella parte di dottrina italiana che ha ipotizzato una classificazione delle attività che coinvolgono l'utilizzo dell'intelligenza artificiale come attività pericolose, facendole rientrare, per quanto concerne la responsabilità, nell'ambito applicativo dell'art. 2050 c.c.<sup>115</sup>. L'operazione consentirebbe di estendere di molto la tutela assicurata al danneggiato, considerato che la responsabilità per l'esercizio di attività pericolose è una responsabilità oggettiva la cui sussistenza dipende solo dall'accertamento del danno e del nesso di causalità, prescindendo da ogni valutazione in termini soggettivi e dalla difettosità del prodotto che caratterizza la responsabilità del produttore. Inoltre, in ambito europeo, tale strumento si aggiunge e non sostituisce la tutela offerta dalla disciplina del

<sup>108</sup> Per un'ampia disamina comparativa v. U. RUFFOLO, E. AL MUREDEN, *Autonomous Vehicles e responsabilità nel nostro sistema ed in quello statunitense*, in *Giurisprudenza italiana*, 7, 2019, p. 1707.

<sup>109</sup> Il Self Drive Act (H.R. 3388 del 7 settembre 2017 disponibile al link <https://www.congress.gov/bill/115th-congress/house-bill/3388/titles>) assegna alla legge federale il ruolo di aggiornare i *vehicles safety standards*, eliminando i riferimenti allo human driver, e di delineare i caratteri delle *safety assessment certifications* degli *Highly Automated Vehicles*, promuovendo la loro sperimentazione e diffusione.

<sup>110</sup> Cfr. M. GEISTFELD, *The Regulatory Sweet Spot for Autonomous Vehicles*, in *Wake Forest L. Rev*, 53, 2018, p. 120.

<sup>111</sup> Si veda il noto leading case *Geier v. American Honda Company*, 529 U.S.861 (2000).

<sup>112</sup> K.S. ABRAHAM, R.L. RABIN, *Automated Vehicles and Manufacturer Responsibility for Accidents: A New Legal Regime for a New Era*, in *Virginia Law Review*, 1, 2019, p. 127 e ss.

<sup>113</sup> Cfr. A. QUERCI, *Sicurezza e danno da prodotti medicali*, Torino, 2011, p. 202.

<sup>114</sup> Cfr. A. QUERCI, *Responsabilità da prodotto negli Usa e in Europa. Le ragioni di un revirement "globale"*, in *Nuova Giur. Civ. Comm.*, 2, 2011, p. 118.

<sup>115</sup> Cfr. C. SCOGNAMIGLIO, *Responsabilità civile ed intelligenza artificiale: quali soluzioni per quali problemi?*, in *Responsabilità civile e previdenza*, 4, 2023, p. 1073; L. COPPINI, *Robotica e intelligenza artificiale: questioni di responsabilità civile*, in *Politica del diritto*, 4, 2018, p. 735; M. SCIALDONE, *Il diritto dei robot: la regolamentazione giuridica dei comportamenti non umani*, in E. PIETRAFESA, F. MARZANO, T. MEDICI (a cura di), *La rete e il fattore C: Cultura, Complessità, Collaborazione*, Volume II, Roma, 2016, p. 76; A. SANTOSUOSSO, M. TOMASI, *Diritto, scienza, nuove tecnologie*, Padova, 2021, p. 329; C. LAENZA, *Intelligenza artificiale e diritto: ipotesi di responsabilità civile nel terzo millennio*, in *Resp.civ. prev.*, 3, 2021, p. 1018. In particolare, in relazione alle auto a guida autonoma v. U. RUFFOLO, *Intelligenza artificiale ed automotive: la responsabilità dei veicoli self-driving e driverless*, in U. RUFFOLO (a cura di), *Intelligenza artificiale: il diritto, i diritti, l'etica*, Milano, 2020, p. 159.

produttore poiché responsabilità basate su presupposti diversi<sup>116</sup>. Per quanto interessante, questa tesi ci pare criticabile alla luce della constatazione che l'utilizzo in sé dell'intelligenza artificiale, non può essere considerato tout court un'attività pericolosa, come invece è stato sostenuto, poiché essa rappresenta la massima espressione della tecnica cui oggi siamo giunti, mezzo integrativo o addirittura correttivo delle imprecisioni umane<sup>117</sup>. Questa considerazione è *a fortiori* valida per l'ambito della guida autonoma: se ad oggi la circolazione dei veicoli non è considerata un'attività pericolosa, ancor meno potrà essere considerata tale la circolazione automatizzata che si stima sia in grado di ridurre gli incidenti, ovvero la pericolosità per i soggetti coinvolti nella stessa.

## 6. Conclusioni

La circolazione dei veicoli autonomi pone un duplice problema etico, da un lato le regole di condotta che devono essere impartite alla macchina, determinando il suo comportamento di fronte a dilemmi etici, dall'altro l'equità delle regole di responsabilità. La prima questione impone di ragionare a ritroso indagando quali siano le regole etiche di comportamento che l'essere umano segue, per scoprire che non esiste una morale condivisa a livello globale, presupposto necessario per dare una regola univoca alla macchina. In questo contesto, una possibile soluzione, seppur parziale e difficile da implementare, potrebbe consistere nell'installazione dell'Ethical Knob, meccanismo che tuttavia non risolverebbe totalmente il problema, riservando all'utilizzatore del veicolo l'ultima parola.

La questione più generica della responsabilità dei sistemi di intelligenza artificiale rimane aperta dal punto di vista etico, come anche giuridico: escludendo una responsabilità del sistema di IA, questa rimane agli umani legati in una catena causale complessa che può rendere difficile ragionare nei termini tradizionali della responsabilità individuale. L'intervento di molteplici soggetti anche digitali, aprendo la strada a forme di responsabilità diffusa e relazionale. Una soluzione bilanciata deve bilanciare l'interesse all'innovazione tecnologica con un'adeguata tutela delle vittime di incidenti. Le teorizzazioni dell'etica devono accompagnarsi a quelle del diritto. Di fronte al fenomeno della digitalizzazione, che sta penetrando sempre di più nella società, vi è bisogno di un dialogo tra le discipline che guidi i regolatori alla formulazione della norma "più giusta", e di nuovi (o meno nuovi) capisaldi. Difatti, da un punto di vista giuridico, la questione della responsabilità non può rimanere irrisolta, ma necessita di punti fermi, pena il collasso del sistema che non dà certezze di sicurezza in punto di risarcimento. Abbiamo voluto fornire quella che ci sembra la proposta più condivisibile, e che rinviene tali certezze nella responsabilità oggettiva del proprietario del veicolo autonomo fonte del danno, nonché in quella del produttore in caso di *driverless car* difettosa. Non si può sottacere, infatti, che il diritto, pur dovendo guardare con favore alle suggestioni provenienti

---

<sup>116</sup> Sulla cumulabilità delle tutele cfr. A. FUSARO, *La responsabilità secondo altre leggi*, in G. ALPA, M. BIN., P. CENDON (a cura di), *La responsabilità del produttore*, in F. GALGANO (dir. da), *Trattato di dir. comm. e dir. pubbl. dell'econ.*, Vol XIII. Padova, 1989, p. 265; M. BIANCA, *Diritto civile, V, La responsabilità*. Milano, 1994, p. 739; G. ALPA, M. BESSONE, Z. ZENCOVICH, *I fatti illeciti*, in P. RESCIGNO (diretto da), *Trattato di diritto privato*, VI, Torino, 1995, p. 390.

La stessa CGUE, nel caso C-402/03 Skov Aeg v. Bilka, aveva affermato che la direttiva 374/85 lasciava impregiudicati i modelli di responsabilità contrattuale ed extracontrattuale preesistenti nei singoli stati membri purché basati su elementi diversi, quali ad esempio la garanzia per i vizi o la colpa.

<sup>117</sup> Cfr. M. COSTANZA, *Intelligenza Artificiale e diritto - L'Intelligenza Artificiale e gli stilemi della responsabilità civile*, in *Giurisprudenza Italiana*, 7, 2019, p. 1657; G. D'ALFONSO, *Intelligenza artificiale e responsabilità civile. Prospettive di regolamentazione europea*, in *Revista de Estudios Jurídicos y Criminológicos*, 6, 2022, p. 174.

dall'etica, ha un fine diverso, ossia mantenere l'ordine sociale, evitando che si verifichi lo scenario moralmente più inaccettabile: che nell'incertezza complessiva a rimanere inciso dal danno sia solo il danneggiato. Per questo motivo le regole di responsabilità oggettiva, per quanto criticabili da un punto di vista etico, perché non connesse a una condotta colposa dell'agente, sono necessarie per garantire la cd. socializzazione del danno.