

# ELEMENTI DI ROBOETICA

## Analisi di alcune metaetiche nella progettazione e programmazione di veicoli autonomi

Fiorella OPERTO

(Scuola di Robotica, Genova)

Robotics is rapidly becoming one of the leading fields of science and technology, so that very soon humanity is going to coexist with a totally new class of technological artifacts: the robots. Far from indulging in Sci-Fi's scenarios, we could nonetheless describe the coming robot revolution as the first invasion of an alien race on Earth. Robots will assist humans in several circumstances, solving many problems, from daily life to productive activities to healthcare, from amusing kids to safeguarding the well-being of our planet. Not all of these applications are going to raise ethical, social or psychological problems. However, as in every other field of science and technology, sensitive areas open up with a new array of problems that constitute the content of Roboethics today. This essay addresses some key general notions of Roboethics that are at the upstream of the regulatory rules for advanced robots' applications. We then analyse the two levels of Roboethics applications: the applied ethics that attempts to provide philosophical (ethical) answers to new issues arising from advanced robotics; and a second level (Robot Ethics or Machine Ethics), derived from Roboethics, regarding the code of conduct that designers implement in the Artificial Intelligence of robots. Next, we shortly highlight that, alongside the techno-scientific idea of Robotics, there is also a Robotics ideology shaped over centuries as legends and myths by human's ingenuity. This cluster of literary visions, wishes and hopes attributes to robots some magic powers embodied in their Artificial Intelligence, so that it is imagined that robots could "decide" autonomously their own behaviour, selected on the basis of a deep analysis of the context. How this ideology affects our idea of Robotics still today? As Case Study application, we compare some real ELS issues arising from self-driving vehicles use—as outlined by experts in the field—with some of Robotics Ideology idola as expressed by the most common interpretation of the so-called Trolley Paradox.

Keywords: roboethics, autonomous vehicles, machine ethics.

### 1. Introduzione

La robotica sta rapidamente sviluppandosi in uno dei capi di tecnoscienza più innovativi; tra pochi anni non solo gli esperti, anche il pubblico generale sarà in contatto, o opererà, con questi nuovi artefatti tecnologici, i robot. Non vogliamo

indulgere in scenari fantascientifici, ma possiamo dire che gli umani saranno a contatto con la prima intelligenza aliena della storia.

I robot stanno già assistendo gli umani in diversi compiti pericolosi, difficili o ripetitivi, e soprattutto i robot di servizio, professionali o personali, stanno entrando nelle nostre vite: negli ospedali, nelle scuole, nelle manifatture, nel tempo libero<sup>1</sup>.

Non tutte queste applicazioni robotiche sono destinate a sollevare seri problemi etici, legali e sociali. Ma, come in ogni altro campo di scienza e tecnologia, l'impatto che i robot avranno sulla nostra vita dovrà essere valutato rispetto alle raccomandazioni di Roboetica sviluppate da vari anni<sup>2</sup>.

Da alcuni anni, i professionisti che lavorano nel campo della robotica e il grande pubblico si stanno ponendo, a diversi livelli e con varie impostazioni, domande riguardanti gli aspetti etici, sociali e legali della ricerca robotica e delle sue applicazioni. La robotica è una nuova scienza ancora allo stato nascente, nata dalla fusione di molte discipline appartenenti al campo delle scienze umane e di quelle naturali. Chiunque, anche da amatore, si avvicini a essa, ne riconoscerà lo sviluppo caotico che caratterizza una rivoluzione scientifica, che invade i campi tradizionali del sapere e della tecnologia e che impone problemi nuovi e complessi di natura etica, filosofica, sociale, legale<sup>3</sup>. Sono soprattutto le ricerche e applicazioni nel settore della robotica di servizio applicata al medicale, alla biorobotica, alla robotica per assistenza e alla robotica militare a sollevare inquietudini e perplessità.

La Roboetica è nata nel 2003 dal robotico Gianmarco Veruggio. Infatti, tra il 2003 e il 2004 si era intensificato, parallelamente all'affacciarsi dei robot sulle scene dell'*edutainment* (termine che unisce *education* ed *entertainment*), il dibattito sui rischi costituiti dall'ingresso dei robot nelle nostre società. Nel febbraio del 2005 il Network europeo della robotica della Commissione Europea finanziò il progetto per un Atelier sulla Roboetica. Organizzato da Scuola di Robotica nel 2006 a Genova<sup>4</sup>, l'Atelier

---

<sup>1</sup> Bruno SICILIANO e Oussama KHATIB (a cura di), *Springer Handbook of Robotics*, II Edition, Springer, Berlin-Heidelberg 2016.

<sup>2</sup> Gianmarco VERUGGIO, Fiorella OPERTO, e George BEKEY, *Roboethics: Social and Ethical Implications*, in B. SICILIANO e O. KHATIB (a cura di), *The Springer Handbook of Robotics*, cit., pp. 2135-2160.

<sup>3</sup> Fiorella OPERTO, *Ethics in Advanced Robotics*, in "IEEE Robotics & Automation Magazine – Special Issues on Robot Ethics, Robotics and Automation Society", n. 18, v. 1, anno 2011, pp. 72-78.

<sup>4</sup> <http://www.euron.org/projects/roboethics.html>.

produsse la *Roadmap* della Roboetica<sup>5</sup>. Il processo di *editing* della *Roadmap* sulla Roboetica ha coinvolto molti studiosi e robotici. Nel 2006 la *Roadmap* è stata pubblicata sul sito di *Euron*<sup>6</sup>, e nel 2007 è stata presentata ufficialmente durante *l'International Conference on Robotics and Automation* (ICRA 2007). Da allora, la Roboetica è entrata ufficialmente tra le Etiche Applicate, oggetto di una ormai vasta letteratura.

Questo articolo presenta alcuni concetti generali di Roboetica che sono stati acquisiti da diversi enti internazionali (Commissione Europea, ONU, UNESCO) e enti regolatori (ISO). Analizziamo successivamente i due livelli della Roboetica: l'etica applica che ha come obiettivo di studiare i problemi etici, legali e sociali della robotica avanzata (etica); e il secondo livello (*Robot Ethics*, o *Machine Ethics*), derivato dalla prima, che riguarda le norme di funzionamento (“comportamento”) che i progettisti, programmatori e utenti di robot dovrebbero seguire.

Delineeremo poi brevemente che, parallelamente al recente sviluppo della Robotica, continua a vivere una ideologia robotica che ha incorporato elementi da mitologie antiche (la ribellione dei robot, il Golem, Frankenstein). È un insieme di strutture letterarie, visioni, antropomorfizzazioni di macchine e meccanismi, miti e leggende che sono evidentemente di natura letteraria e nulla hanno a che vedere con la realtà della robotica e del suo funzionamento (ma molto sul nostro rapporto con i nostri artefatti). Sebbene queste espressioni letterarie siano interessanti e importanti da studiare, oggi l'ideologia robotica si sovrappone, nella visione del grande pubblico, alla realtà del settore, un po' come la chimica fu confusa con, o suffusa di, alchimia ancora nel Settecento europeo. Concezioni e rappresentazioni su un robot che “decida”, “si emozioni” o “si sviluppi come un umano” fanno parte ancora di questa sovrapposizione.

Infine, analizzeremo un caso studio su problemi reali di Roboetica che possono derivare dall'impiego di veicoli autonomi in percorsi pubblici e cercheremo di applicare la Roboetica ad alcuni di questi problemi.

---

<sup>5</sup> Gianmarco VERUGGIO, *The EURON Roboethics Roadmap, European Robotics Research Network, Atelier on Roboethics, 2005-2007*. <http://www.roboethics.org>.

<sup>6</sup> [www.euron.org](http://www.euron.org).

## 2. Roboetica: un'etica applicata relativa ai robot, per gli umani

Similmente alla prima, seconda e terza Rivoluzione Industriale, ma in modo più complesso ed originale, la quarta rivoluzione industriale, basata sull'introduzione del computer, dell'Intelligenza Artificiale e della Robotica nei processi produttivi e nella società, ha introdotto cambiamenti sociali ed economici profondi che trovano società, decisori e noi tutti, impreparati.

È la prima volta nella storia umana che ci troviamo a collaborare su diversi campi con un'intelligenza artificiale. La convergenza di robotica, AI, reti e big data, senza contare gli sviluppi nelle neuroscienze, rappresentano oggi un *melting pot* straordinario, dove la sinergia tra queste tecnologie sta imprimendo un ritmo esponenziale al progresso tecnologico e alle possibili applicazioni<sup>7</sup>. E con moltissime opportunità e benefici, tutto questo sta sollevando problemi etici nuovi. Alcuni campi di applicazione della robotica hanno già posto seri problemi nel campo dei diritti umani: la robotica militare, ovvero sistemi robotici applicati a sistemi d'arma (robot mobili e droni armati). E il settore della bionica, con la progettazione e impiego di impianti ibridi e impianti *Brain-Computer*.<sup>8</sup>

Inoltre, il fatto che i robot possano collegarsi alla rete Internet, o ad altre reti, permette lo sviluppo di sistemi robotici distribuiti, aggregati di robot i cui processori connessi formano un processore parallelo la cui potenza permetterebbe di affrontare problemi complessi, elaborare i dati provenienti dall'intera rete di robot e monitorare così, per esempio, dalla singola cellula a organismi complessi<sup>9</sup>. Fra poco tempo, molto prima di quanto forse pensiamo, impiegare un robot sarà usuale come usare uno *smartphone* oggi. Tra una decina di anni vedremo persone acquistare robot come oggi gadget digitali: il mercato offrirà robot standard e l'utente li personalizzerà con *app*, sensori e altre tecnologie digitali come oggi scarichiamo *app* per il nostro *smartphone*, aggiornando il "formato standard" per le funzioni desiderate.

I robot così conformati, dotati di *Machine Learning*, di capacità di apprendimento, coabiteranno con noi, e apprenderanno e memorizzeranno i pattern dei nostri

---

<sup>7</sup> Gianmarco VERUGGIO, TedTalk, 2018, <https://www.youtube.com/watch?v=FO68leHdOFo>.

<sup>8</sup> Guglielmo TAMBURRINI, *Robot ethics: a view from the philosophy of science*, in R. CAPURRO e M. NAGENBORG (a cura di), *Ethics and Robotics*, IOS Press, Heidelberg 2009, pp. 11-22.

<sup>9</sup> Dezheng SONG, Ken GOLDBERG, Nak Young CHONG, *Networked Telerobots*, in B. SICILIANO e O. KHATIB (a cura di), *Springer Handbook of Robotics*, cit., pp. 759-771.

ambienti, il nostro profilo (stato di salute, abitudini, preferenze, routine) e ogni altro pattern abilitato a renderli più “intelligenti”.

Oltre che con gli umani, questi robot avanzati scambieranno dati e software tra di loro, e tra banche di dati, senza diretto intervento umano, con l’obiettivo di immagazzinare, per esempio, dati e soluzioni già trovate a problemi complessi. Vedremo, nel caso studio, come veicoli dotati di autonomia e di IA possano acquisire e scambiare dati sullo stato delle strade, sulla situazione meteo, il traffico, e, ovviamente, il nostro profilo come “passeggero”.

In questi casi, rispetto a robot dotati di capacità di apprendimento, i problemi chiamati ELS (acronimo di *Ethical, Legal, and Social Issues*) saranno maggiori, coinvolgendo questioni di privacy, identità personale, dignità, discriminazione e accessibilità<sup>10</sup>.

Nei robot, l’apprendimento è basato sulla maggior quantità e qualità possibili di modelli (pattern) di situazioni/risposte che vengono inseriti nel robot, modelli riferenti a situazioni che il robot può incontrare nelle sue operazioni e a possibili cambiamenti dell’ambiente in cui la macchina dovrà operare. Su questi, il robot integra le situazioni già sperimentate nell’ambiente in cui opera. Ovviamente, non potendo immaginare tutte le possibili situazioni, i modelli sono su base statistica, e maggiore sarà la quantità di esempi e variabili, migliore sarà la risposta del robot e il suo livello di apprendimento. Tuttavia, la predizione di correttezza sarà sempre approssimativa e, in ultima analisi, non vi è modo di predire con certezza quale sarà la performance di una macchina che apprende in un ambiente poco o nulla strutturato, di fronte a nuove missioni. Rifacendosi alla teoria dell’apprendimento computazionale—un campo di studi tra l’intelligenza artificiale, la statistica, l’informatica teorica, e i metodi di *machine learning*—i redattori della *Roadmap* e i partner del progetto europeo *Ethicbots* sono giunti alla conclusione che né i progettisti, né i produttori di robot né gli utenti finali siano in grado, oggi, di predire con precisione le azioni di una *learning machine*, di un

---

<sup>10</sup> Fiorella OPERTO, *Ethics in Advanced Robotics. ELS Issues in Advanced Robotics*, in “IEEE Robotics & Automation Magazine – Special Issues on Robot Ethics, Robotics and Automation Society”, n. 18, v. 1, anno 2011, pp. 72-78.

robot dotato di capacità di apprendimento, in un ambiente umano altamente modificabile<sup>11</sup>.

L'ideologia robotica citata, ovvero la sovrapposizione tra la rappresentazione della realtà dei robot e la mitologia antica, determina anche un *bias* sugli aspetti ELS della stessa.

Per Roboetica (scritto spesso come Robot Etica, o *Robot Ethics*) molti hanno inteso la capacità dei robot dotati di apprendimento di “decidere” autonomamente, tra diverse scelte di comportamento, quella più “etica” e anche di far emergere una nuova “etica”, più “etica” di quella degli umani, perché non affetta da passioni ed emozioni<sup>12</sup>.

In realtà, l'elemento chiave della Roboetica non è un'etica artificiale, ma la discussione e lo studio sui principi etici degli umani, di chi queste macchine progetta, produce e impiega.

Noi robotici dobbiamo assicurare il nostro impegno per aumentare la consapevolezza del pubblico circa le problematiche della robotica, affinché la società possa prendere parte attiva nel processo di creazione di una coscienza collettiva, in grado di individuare e prevenire un uso errato della tecnologia. La speranza è che si possa giungere a un'etica condivisa da tutte le culture, tutte le nazioni e le religioni, così che la costruzione e l'impiego di robot contro gli esseri umani sia considerato un crimine contro l'umanità.<sup>13</sup>

Durante il primo Simposio Internazionale sulla Roboetica (2004), l'antropologa Daniela Cerqui presentò tre principali posizioni, di scienziati e ricercatori, rispetto all'etica nella scienza e tecnologia:

- Scienziati e ricercatori che non sono interessati alle conseguenze etiche delle loro attività. Giudicano il proprio intervento professionale come limitato all'ambito tecno scientifico e non ritengono di avere alcuna responsabilità sociale o morale rispetto alla loro attività professionale.

---

<sup>11</sup> Edoardo DATTERI, Hykel HOSNI, e Guglielmo TAMBURRINI, *An inductionless and default-based analysis of machine learning procedures*, in L. MAGNANI (ed.), *Model Based Reasoning in Science and Engineering*, College Publications, London 2006, pp. 379-399.

<sup>12</sup> Ronald Craig ARKIN, Patrick ULAM, e Alan R. WAGNER, *Moral Decision-making in Autonomous Systems: Enforcement, Moral Emotions, Dignity, Trust and Deception*, in “Proceedings of the IEEE”, n. 100, v. 3, March 2012, pp. 571-589.

<sup>13</sup> Gianmarco VERUGGIO e Keith ABNEY, *Roboethics: the applied ethics for a new science*, in P. LIN, G. BEKEY, e K. ABNEY (a cura di), *Robot Ethics*, MIT Press, Cambridge 2011, p. 355. Traduzione a cura dell'autrice.

- Scienziati e ricercatori che sono interessati ai problemi etici della loro professione sul breve termine. Da questo punto di vista, le questioni sono espresse in termini di “buono” o “cattivo”, e riferite abbastanza rigidamente a qualche valore culturale. Per esempio, questi scienziati capiscono che i robot devono aderire alle convenzioni sociali. Questo includerà “rispettare” e aiutare gli uomini in diversi contesti.
- Scienziati e ricercatori che guardano alle questioni etiche da un punto di vista di lungo periodo e che si pongono la domanda che precede ogni altra, nella Roboetica: i robot sono necessari, sono utili, sono indispensabili nel mondo d’oggi, con tutti i gravi problemi che ci affliggono? La risposta è sempre relativa all’impiego delle tecnologie: se i robot saranno impiegati a beneficio degli umani, sì, saranno molto utili<sup>14</sup>.

Rispetto al primo atteggiamento, sappiamo che questo contiene risposte che fanno parte di un modo “positivistico”, ottocentesco, di considerare la tecnoscienza. Forse oggi nessun ricercatore e produttore di tecnologia si esprimerebbe più in questi termini, almeno non pubblicamente. Di fatto, tuttavia, molti professionisti che lavorano in questi settori si adeguano nella pratica a questo atteggiamento di responsabilità limitata.

Il secondo profilo è senza dubbio più complesso: per semplificare, dobbiamo ricordare che i valori usati per definire “cattivo” e “buono” sono relativi. Nel nostro caso, sono i valori di riferimento dei paesi industrializzati. Senza una Roboetica, il singolo ricercatore è in ogni caso lasciato solo a decidere se continuare a lavorare su aspetti sensibili della propria attività, o smettere. Una decisione spesso drammatica. Ma la Roboetica è veramente solo un problema individuale del singolo ricercatore, di chi impiega i robot, del singolo che deve fare i conti con la propria coscienza? Oppure è un problema sociale che deve essere affrontato a un livello istituzionale?

Il terzo profilo corrisponde a quei ricercatori, produttori e utilizzatori su e di robot che considerano la loro attività all’interno del generale stato del mondo. Vorrebbero che la robotica andasse a beneficio di tutti. La domanda successiva è: chi decide che cosa sia il “beneficio” degli umani?

---

<sup>14</sup> Daniela CERQUI, Jutta WEBER, e Karsten WEBER (a cura di), *Ethics in Robotics*, in “International Review of Information Ethics”, n. 6, anno 2006.

### 3. Quale etica per la Roboetica?

La prima domanda di etica generale che si posero i redattori della *Roadmap* è stata: quale etica per la Roboetica? Infatti, la nostra etica individuale, quella trasmessaci dalla famiglia, dagli insegnanti o appresa con letture e con l'esperienza, anche se condivisa da molti appartenenti alla nostra comunità, non è praticata in egual modo da tutto il nostro gruppo di pari, tantomeno da tutti gli esseri umani sul pianeta. Non solo esistono molte etiche individuali, nel mondo, ma anche diverse teorie etiche (utilitarismo, generale e normativo; l'etica deontologica o kantiana; etiche religiose e laiche; la *virtue ethics*, e così via)<sup>15</sup>. Da queste derivano etiche descrittive e normative, ed etiche applicate ai diversi settori della scienza e tecnologia. In questo ultimo caso, si richiede il contributo comune sia di moralisti e filosofi sia di esperti dei settori interessati. Per quanto riguarda la bioetica, per esempio, sappiamo quali approfondite e accese discussioni vi siano relativamente ad aspetti della vita e della morte. Inoltre, rispetto alla Roboetica, l'internazionalizzazione della ricerca e la globalizzazione economica fanno sì che i robotici e i produttori di robot appartengano a Nazioni e culture assai lontane le une dalle altre. Si può quindi facilmente immaginare quanto sia difficile universalizzare concetti sensibili come la dignità dell'individuo o il rispetto della privacy<sup>16</sup>.

I redattori della *Roadmap* decisero così di seguire i principi universalmente accettati – almeno sulla carta – espressi nelle Carte Fondamentali dei Diritti Umani: la Dichiarazione Universale dei Diritti Umani firmata nel 1948 e promossa dalle Nazioni Unite, e la Carta dei diritti fondamentali dell'Unione europea (Nizza, 2000). Non tutte le Nazioni del mondo, come sappiamo, hanno adottato tutte le norme (prescrizioni e divieti) ivi contenute: per esempio, in molte Nazioni del settore detto avanzato vige la pena di morte. E ancora, in molte Nazioni del mondo la posizione della donna e del bambino rispetto ai diritti fondamentali non è tutelata, o, peggio, non è nemmeno presa in considerazione. Inoltre, vi sono Nazioni che proteggono i diritti degli animali e molte

---

<sup>15</sup> Adriano FABRIS, Sergio BARTOLOMMEI, e Edoardo DATTERI, *Quale etica per la robotica?*, in "Teoria", n. XXVII, vol.2, anno 2007, pp. 7-17.

<sup>16</sup> Fiorella OPERTO, *Ethics in Advanced Robotics. ELS Issues in Advanced Robotics*, in "IEEE Robotics & Automation Magazine – Special Issues on Robot Ethics, Robotics and Automation Society", n. 18, v. 1, anno 2011, pp. 72-78.



no; e più in generale, esiste il problema dei diritti complessivi del nostro pianeta, che sostiene le vite di tutti noi.

Molti dei co-redattori della *Roadmap* sulla Roboetica si sono detti coscienti che i principi contenuti nella Dichiarazione Universale dei Diritti Umani e nella Carta dei diritti fondamentali dell'Unione europea non solo non sono accettati da molte Nazioni, ma non sono veramente applicati da quelle che firmarono le due Carte. Inoltre, quei principi, sebbene consistenti e fondamentali, non esauriscono né soddisfano gli aspetti originali che i prodotti della robotica potranno sviluppare rispetto ai nuovi diritti e ai nuovi divieti. Abbiamo visto come la diffusione di Internet e delle reti abbia creato bisogni e prescrizioni che sono diventati anche aspirazione a diritti e nuovi problemi sociali.

Nel secolo scorso, infatti, a seguito dei drammatici eventi della Prima e Seconda Guerra Mondiale, e di alcuni Olocausti, molti studiosi hanno avanzato l'idea che la nozione di responsabilità legata a un ruolo dovesse essere sostituita da un globale senso di responsabilità umana. È stato dopo questi eventi che il principio di "diritto umano" indipendentemente dalla cittadinanza, ceto sociale, genere è entrato in vigore come un principio attivo.

#### **4. Una ideologia robotica**

Nel 18mo secolo, uno degli impegni principali dei ricercatori e studiosi di elettricità e magnetismo, oltre alla ricerca, fu di liberare il settore da credenze magiche e leggende. Ancor prima, i chimici dovettero fare i conti con l'alchimia e le superstizioni collegate, per liberare la chimica dalla leggenda.

La robotica è un settore tecnoscientifico – una scienza, sostengono alcuni – relativamente giovane e il suo mondo culturale, le immagini e la narrazione collegati risentono ancora della fantascienza e della letteratura, influenzate dalle colorate descrizioni, dai personaggi e dalle storie che per il grande pubblico sono assai più affascinanti di queste macchine ancora poco note. Il comportamento dei robot è spesso ancora antropomorfizzato, e lo è, sorprendentemente, anche da ricercatori ed esperti che, per lo più provenienti da studi di ingegneria, prendono a prestito nella loro narrazione ai non esperti, temi dalla letteratura e dai miti, non essendovi ancora una letteratura robotica corretta.

Purtroppo, questa ideologia robotica sta frenando la comprensione della robotica da parte del pubblico e potrebbe impedire o rallentare scelte sociali generali corrette.

In realtà, i robot narrati *qua robot*, secondo la loro natura di macchine autonome e dotate di apprendimento, possono essere assai più affascinanti di molte descrizioni che riusano concetti letterari un po' logori, o che traslano sui robot comportamenti umani.

Un altro aspetto della ideologia robotica è lo stereotipo per cui queste macchine sarebbero entità intelligenti, quasi perfette, razionali. Spesso il robot è rappresentato con capacità decisionali adatte a situazioni imprevedute, essendo dotato di rapidità e potenza di calcolo, guidato da algoritmi precisi ed esente dagli squilibri umani dovuti alle emozioni.

È interessante notare che questa rappresentazione contraddice la nostra quotidiana esperienza di interazione con macchine digitali che, lontano dall'essere perfette, esibiscono comportamenti incomprensibili, sono afflitte da *bug* e si rompono quando più ne avremmo bisogno.

E soprattutto, i robot sono straordinari soprattutto con i compiti routinari; meno, almeno ora, in situazioni imprevedute.

L'effetto di una ideologia è precisamente di mascherare la realtà; non solo, di costruirne un'altra che è finzione, di offrire un paradigma, delle spiegazioni, adatte al nostro ambiente culturale e sociale. Un esempio. Se paragoniamo il numero di incidenti aerei con il numero di incidenti mortali su strada, e anche con quelli su altri mezzi (treno, bus, ecc) siamo confermati nell'opinione che il viaggio in aereo è il più sicuro, per kilometraggio. Ma quando accade un incidente aereo, la reazione del pubblico è amplificata precisamente dal fatto che non ci aspetteremmo un tale evento. Siamo ormai preparati ad accettare i morti dovuti a incidenti automobilistici perché abbiamo condiviso, almeno una volta, l'esperienza di poter incorrere in un incidente automobilistico grave. Conosciamo le emozioni coinvolte che ci appaiono così "normali", prossime. Ma pochi di noi condividono con i piloti di aereo l'esperienza di un incidente. Quindi, non ci aspettiamo un incidente aereo, che diventa eccezionale.

Ecco perché ha ragione un esperto quando sostiene che il vero "Paradosso" delle macchine autonome in genere (e, in questo caso, dei veicoli autonomi) è l'"Over-Trusting Autonomy Paradox", come vedremo.

Nel paragrafo seguente cerchiamo di individuare elementi di Robotica in un caso studio: questioni ELS nei veicoli autonomi sul mercato e in corso di test. Cerchiamo di

individuare le metaetiche che sottendono le decisioni dei progettisti e dei produttori di questi veicoli.

## 5. Un caso studio: alcune metaetiche nella progettazione di veicoli autonomi

Da alcuni anni, le società produttrici di autoveicoli stanno progettando veicoli senza pilota, o autonomi. Sul mercato delle autovetture si trovano già diverse marche, con dati livelli di autonomia. Ad oggi, nessuna marca di autoveicoli ha messo sul mercato modelli con autonomia 5, che è il livello dove il veicolo ha la capacità di guidare in completa autonomia indipendentemente dalle condizioni ambientali<sup>17</sup>. In questo caso, i veicoli sono robot autonomi dotati di capacità di apprendimento.

La pubblicità dei produttori di veicoli autonomi promette oggi autonomia 3 o 4, ma vedremo che siamo ancora lontani da questi livelli, soprattutto in tutte le situazioni ambientali.

A seguito di alcuni incidenti, soprattutto due mortali che hanno coinvolto una Tesla e un veicolo Uber – nel marzo del 2018 quest’ultimo ha ucciso una donna in Arizona, ed è stato il primo incidente mortale che abbia coinvolto un pedone – alcuni ministeri dei Trasporti e Autorità analoghe alla *US National Highway Traffic Safety Administration* (NHTSA) – in Germania e Norvegia – hanno emanato direttive di sicurezza relative. In particolare, la NHTSA ha chiesto ai produttori su quali basi fossero stati progettati i veicoli, tenendo conto degli effetti ELS: quali fossero le decisioni tecnologiche sottostanti i comportamenti dei veicoli autonomi.

Patrick Lin, direttore dell’*Ethics and Emerging Sciences Group* al *California Polytechnic State University* di San Luis Obispo scrive che il Cal Poly, la Stanford CARS e la Daimler hanno iniziato discussioni su questi aspetti, cui hanno partecipato anche Google/Waymo e Ford<sup>18</sup> (Lin, 2018).

Oggi, rispetto ai problemi ELS in veicoli autonomi, l’argomentazione più diffusa e discussa è espressa dal cosiddetto *Trolley Paradox*, il paradosso del tram, o del carrello ferroviario. Il dilemma descrive un tram che corre sul proprio binario, e che per un guasto o simile non possa frenare, possa solo deviare su uno dei due binari. Su un

---

<sup>17</sup> Per i livelli di autonomia, si veda lo *US National Highway Traffic Safety Administration*: <https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety>.

<sup>18</sup> Patrick LIN, *Who’s at Fault in Uber’s Fatal Collision?*, in “IEEE Spectrum”, March 2018. <https://spectrum.ieee.org/cars-that-think/transportation/safety/reflecting-on-ubers-fatal-crash>.

binario c'è una persona anziana, sull'altro due bambini. Il manovratore può solo deviare sull'uno o sull'altro binario. Che cosa sceglierà?

Questo classico esperimento di pensiero ha molte versioni, e quella aggiornata sui veicoli autonomi, attribuisce a questi la decisione ultima sul proprio comportamento, a seconda di come sarebbe stato programmato in tali situazioni.

Questo dilemma, se interessante per studenti di filosofia, non facilita in realtà l'analisi degli aspetti reali degli ELS in veicoli autonomi. Al contrario, nell'opinione di chi scrive (e di altri studiosi<sup>19</sup>) può ingenerare confusione.

Le case automobilistiche hanno presente e tengono conto dei problemi ESL quando progettano e producono veicoli autonomi? In un commento duro ma realistico, Brian Casey della Stanford Law School, afferma:

Le case automobilistiche, premute dalla necessità di massimizzare i profitti, saranno certamente poco attente a esoterici problemi di giusto e sbagliato rispetto a questioni per loro più concrete come l'attribuzione di responsabilità legale. Robotici, fatevi da parte. Avvocati, a voi il campo!<sup>20</sup>

Vediamo perché, a opinione di chi scrive, il Paradosso del Trolley possa qui sviare la discussione, citando un esperto.

L'anno scorso, il direttore del Toyota Research Center (TRI), Gill Pratt, diede una interessante intervista al IEEE Magazine, *Spectrum*, una delle più diffuse e autorevoli riviste di ingegneria e sistemi. Toyota sta investendo lungo i prossimi quattro anni più di un miliardo di dollari in Intelligenza Artificiale, con focus principale sulla guida autonoma. In questa intervista, datata ma tutt'ora illuminante, Gyll Pratt presenta i problemi ELS che stanno incontrando – non solo la Toyota – ma tutte le case impegnate in veicoli autonomi<sup>21</sup>.

---

<sup>19</sup> Patrick LIN, *Why Ethics Matters for Autonomous Cars*, in M. MAURER, J.C. GERDES, B. LENZ, H. WINNER (a cura di), *Autonomous Driving. Technical, Legal and Social Aspects*, Springer, Berlin Heidelberg 2016, pp. 69-85; Bryan CASEY, *Amoral Machines, or: How Roboticians Can Learn to Stop Worrying and Love the Law*, in "Northwestern University Law Review", n. 111, v. 5, anno 2017, pp. 231-250.

<sup>20</sup> Aarian MARSHALL, *Lawyers, Not Ethicists, Will Solve the Robocar 'Trolley Problem'*, in "Wired", May 28, 2017. <https://www.wired.com/2017/05/autonomous-vehicles-trolley-problem/>. Traduzione a cura dell'autrice.

<sup>21</sup> Evan ACKERMAN, *Toyota's Gill Pratt on Self-Driving Cars and the Reality of Full Autonomy*, in *Spectrum*, January 2017. <https://spectrum.ieee.org/cars-that-think/transportation/self-driving/toyota-gill-pratt-on-the-reality-of-full-autonomy>.

Sulla base delle sue osservazioni, e di alcuni altri studi, tra cui quello dell' *US Energy Department* del 2018<sup>22</sup>, cerchiamo di analizzare quelli che pensiamo siano gli stereotipi nella rappresentazione generale del funzionamento di veicoli autonomi, e la loro realtà.

Circa il reale livello di autonomia, purtroppo la pubblicistica non è chiara. Secondo l'esperto, «nessuno si avvicina al livello 5». Anche se questa affermazione si riferisce a un anno fa, ci sembra ancora attuale, se per livello 5 si intenda, come è, la capacità di un veicolo di muoversi nel traffico, in modo completamente autonomo, e in ogni condizione meteo, di traffico e altro.

La rappresentazione generale di questi veicoli tende a oscurare alcuni punti. Il primo, come abbiamo detto, è descrivere le capacità attuali dei robot come straordinarie in situazioni complesse. È vero il contrario: «Le macchine (in questo caso, i veicoli autonomi, n.d.r.) sono straordinarie nel mantenersi vigili, ma gli umani lo sono a gestire situazioni difficili nel traffico».<sup>23</sup>

Nel 1980 Hans Moravec, Rodney Brooks, Marvin Minsky e altri svilupparono il cosiddetto “Paradosso Moravec” secondo cui ragionamenti di alto livello richiederebbero relativamente piccola potenza computazionale, mentre capacità sensoriali (o di sensomotori) di basso livello richiedono risorse computazionali notevoli. Sebbene datata, la seguente riflessione di Moravec è ancora attuale. «È relativamente facile realizzare un computer che esibisca livello di performance maturi su test di intelligenza, o in una partita scacchi, è difficile e quasi impossibile fornire loro le capacità percettivo-motorie di un bambino di un anno».<sup>24</sup>

L'altro punto spesso oscurato dalla pubblicistica delle case automobilistiche è l'implicita definizione stessa di “autonomia”: in realtà, i veicoli autonomi sono collegati a diversi altri sistemi, da cui dipendono per la raccolta dati, per le informazioni dai sensori, e così via, tanto che sarebbe più chiaro chiamarle *Connected Autonomous Vehicles, Veicoli Autonomi Connessi*, dato che il loro funzionamento dipende da molti altri operatori e reti.

---

<sup>22</sup> U.S. ENERGY INFORMATION ADMINISTRATION, *Autonomous Vehicles: Uncertainties and Energy Implications*, 2018. <https://www.eia.gov/outlooks/aeo/pdf/AV.pdf>.

<sup>23</sup> Evan ACKERMAN, *Toyota's Gill Pratt on Self-Driving Cars and the Reality of Full Autonomy*, in “IEEE Spectrum”, January 2017. <https://spectrum.ieee.org/cars-that-think/transportation/self-driving/toyota-gill-pratt-on-the-reality-of-full-autonomy>. Traduzione a cura dell'autrice.

<sup>24</sup> Hans MORAVEC, *Mind Children. The Future of Robot and Human Intelligence*, Harvard University Press, Harvard 1990, p. 15. Traduzione a cura dell'autrice.

Una tendenza opposta ma complementare è quella di attribuire alla tecnologia possibili effetti nocivi etici e sociali, che in realtà dipendono dalla progettazione della tecnologia, dalla realizzazione e dal loro uso – da parte di umani.

Abbe Mowshowitz, nel suo *Technology as excuse for questionable ethics*, afferma:

I nostri problemi sociali, che sembrano immutabili, sono veri e presenti, ma cercare di individuarne la causa nella “tecnica” o nella “autonomia tecnologica” è un errore pericoloso. Non dovremmo prendercela con la tecnologia per i nostri guai (..) Soltanto azioni di umani possono essere alienate e disumane. La reificazione della tecnologia permette l’illegittimo trasferimento di responsabilità da umani a costrutti sociali immaginari, il che ostacola la nostra capacità di comprendere e affrontare le sfide etiche molto reali poste dall’impiego della tecnologia.<sup>25</sup>

Non è corretto attribuire alle auto le morti sulle strade, che ammontavano a quasi un milione nel 2013<sup>26</sup>, ma senza dubbio la spinta delle case automobilistiche ad aumentare la potenza dei veicoli, in assenza di altri interventi, è una delle precondizioni di queste stragi. Poiché esiste una correlazione diretta tra velocità del veicolo e numero di incidenti mortali<sup>27</sup>, dovette essere chiaro che un aumento della potenza degli stessi, in assenza di altri interventi, avrebbe prodotto necessariamente questo drammatico risultato. L’applicazione tecnologica non è “neutra”, e in questo senso è da considerarsi “coinvolta” negli effetti, ma siamo noi a progettarela, ad applicarla. Chi decise, ed accettò, che la velocità fosse più importante di così tante vite umane? E come accadde che monetizzammo il rischio e le morti tramite le assicurazioni? Queste sono state di fatto delle scelte sociali.

Le metaetiche sottese a queste decisioni influenzano le applicazioni tecnologiche, e lo fanno “di nascosto”.

Nel prossimo paragrafo analizziamo alcune metaetiche<sup>28</sup> che sottostanno a scelte tecnologiche sui veicoli autonomi. Che lo sia dichiarato o meno, ogni scelta tecnologica

---

<sup>25</sup> Abbe MOWSHOWITZ, *Technology as excuse for questionable ethics*, in “AI & Society”, n. 22, v. 3, anno 2008, p. 275. Traduzione a cura dell’autrice.

<sup>26</sup> Fonte: WHO: [https://www.who.int/gho/road\\_safety/mortality/en/](https://www.who.int/gho/road_safety/mortality/en/).

<sup>27</sup> «For car occupants in a crash with an impact speed of 80 km/h, the likelihood of death is 20 times what it would have been at an impact speed of 30 km/h». Fonte, WHO, *idem*.

<sup>28</sup> Per una definizione di cosa si intende per metaetica si veda la relativa voce della Stanford Encyclopedia of Philosophy (2007): «Metaethics is the attempt to understand the metaphysical, epistemological, semantic, and psychological, presuppositions and commitments of moral thought, talk, and practice. As such, it counts within its domain a broad range of questions and puzzles, including: Is morality more a matter of taste than truth? Are moral standards culturally relative? Are

deriva da una data concezione della vita umana; che sia espresso o meno, ognuna di queste scelte implica l'adesione a una teoria etica. Lo *US National Highway Traffic Safety Administration* (NHTSA) ha aggiornato questa comprensione quando ha emesso una direttiva in 15 punti per le norme 2016 sui veicoli autonomi, dove uno dei 15 punti riguarda la trasparenza circa le decisioni etiche adottate.

## 6. Trasparenza nel processo decisionale sulle tecnologie adottate

Vi sono molte decisioni quotidiane che coinvolgono decisioni etiche nascoste. Per esempio, quando applichiamo il principio del minimo rischio stiamo utilizzando un concetto della teoria morale utilitaristica (non-maleficenza).

Come abbiamo citato, in Roboetica e per la *Roboethics Roadmap* sono stati adottati i principi etici incorporati nella Dichiarazione Generale dei Diritti Umani.

Nel caso di veicoli autonomi, i principi e i diritti umani coinvolti sono:

- Rispetto e protezione della dignità e della privacy umane;
- Diritto all'integrità fisica e mentale della persona;
- Diritto alla libertà e sicurezza;
- Diritto alla protezione dei dati personali.

Aree maggiormente implicate sono:

- Dignità umana e privacy;
- Questioni di attribuzione di responsabilità;
- Effetti psicologici sul passeggero umano.

L'autonomia del veicolo ha effetto su tutti questi aspetti, e principi, perché questa coinvolge questioni di sicurezza, protezione della privacy, attribuzione di responsabilità in caso di danni.

---

there moral facts? If there are moral facts, what is their origin? How is it that they set an appropriate standard for our behavior? How might moral facts be related to other facts (about psychology, happiness, human conventions...)? And how do we learn about the moral facts, if there are any? These questions lead naturally to puzzles about the meaning of moral claims as well as about moral truth and the justification of our moral commitments. Metaethics explores as well the connection between values, reasons for action, and human motivation, asking how it is that moral standards might provide us with reasons to do or refrain from doing as it demands, and it addresses many of the issues commonly bound up with the nature of freedom and its significance (or not) for moral responsibility».

Rispetto a quest'ultimo aspetto, poiché livelli crescenti di autonomia dipendono dalle capacità di apprendimento dei veicoli autonomi, occorre essere consapevoli che, come sostiene Gyll Pratt, i robot che apprendono manifestano ancor oggi un problema: non è possibile sapere come sia stato elaborato il comportamento esibito. Il “veicolo che parla” è ancora l'obiettivo di molte ricerche in tal senso.

Questa incertezza significa che, in caso di incidente, gli ingegneri non possono essere certi di quello che sia accaduto. Da un punto di vista morale, questa finestra di attribuzione di responsabilità dovrebbe richiedere un livello di autonomia che assicuri la sicurezza del o dei passeggeri.

Quando il rationale che governa le applicazioni tecnologiche non sia chiaramente espresso, gli utenti non possono responsabilmente decidere quale veicolo acquistare, come usarlo, e conoscerne i rischi. Il consenso informato è basato sulla conoscenza ed è importante soprattutto quando sono coinvolte tecnologie innovative, la cui performance, per esempio, non sia completamente testata in tutte le situazioni possibili.

Questo significa che decidere il livello di autonomia di un veicolo, sceglierne la campagna pubblicitaria (per esempio, non promettere neanche mediante immagini e altre espressioni mediatiche, performance non sicure), rendere trasparenti le metaetiche, e far conoscere i rischi coinvolti, tutto ciò corrisponde a un comportamento etico da parte dei produttori.

Circa il rispetto del diritto alla privacy, ricordiamo che i veicoli autonomi sono, come detto, collegati a diversi *spot*, diversi punti di connessione: reti interne, collegamenti con la casa produttrice, con infrastrutture pubbliche e private, con altri veicoli, con lo *smartphone* o altri *device* del passeggero. Il comportamento del passeggero sarà monitorato per tutto il viaggio e i suoi dati condivisi con altre reti, che useranno il profilo dell'umano così ottenuto per aggiornare dati e pattern. Saranno, questi dati del passeggero, usati e condivisi o venduti? Saranno utilizzati per aumentare il premio assicurativo, per identificare comportamenti specifici e a loro volta ancora condivisi? In tutto questo, come sarà rispettato il diritto alla privacy?

Un altro problema ELS che sottende una meta-decisione riguarda il livello di sicurezza richiesto e offerto dal veicolo rispetto al livello di autonomia progettato. La domanda “Quale livello di sicurezza è considerato accettabile?” richiede una risposta basata su analisi complesse.

Come società, siamo pronti ad accettare che veicolo autonomi siano più affidabili degli umani, di quanto? 10 per cento? (..) Non saprei, per essere onesto, se questa decisione



spetti a noi, se sia il nostro ruolo, come tecnologi, di fornire la risposta alla società. Questa è la funzione del governo, di altri, che ci dicano ‘Forse questo è sufficiente a salvare una vita’. Oppure, saremmo noi a dover dire: ‘non useremo questo fino a che non sia 10 volte più affidabile (di un pilota umano, n.d.r.). Non sono sicuro, ma penso che fino a che non avremo una risposta a questa domanda, dovremo essere molto prudenti e non introdurre tecnologie che non siano in linea con le aspettative della società.’<sup>29</sup>

Un ultimo aspetto che citiamo dal punto di vista ELS è la struttura gerarchica umano-veicolo. Il problema, in termini roboetici, è quello del livello decisionale in team ibridi, umani-robot. Che deciderà in ultima analisi in situazioni complesse? Come abbiamo visto, i robot sono ottimi in quanto a capacità di esecuzione di compiti routinari, mentre gli umani sono – ancora – più affidabili nel gestire situazioni inattese. Quali saranno le scelte dei produttori? Chi controllerà chi, in questi casi?

## 7. Conclusioni

Come detto nell’Introduzione, i robot saranno la prima entità artificiale intelligente e autonoma con cui avremo a che fare. Vi sono molti problemi i cui contorni ci sono sconosciuti, soprattutto a chi non è nato in un’epoca digitale. Tendiamo, noi non nativi digitali, a impiegare paradigmi analogici. Inoltre, la ancor presente ideologia robotica rende confusi i problemi.

Analizzando le decisioni tecnologiche possibili, e attuate, sui veicoli autonomi, possiamo studiare un caso attuale di roboetica applicata, allontanandoci così da *bias* in cui stereotipi e poca conoscenza possano indurci. E, in ultima analisi, come in tutti i problemi che coinvolgono tecnologie innovative, un alto livello di trasparenza decisionale è necessaria.

---

<sup>29</sup> Evan ACKERMAN, *Toyota's Gill Pratt on Self-Driving Cars and the Reality of Full Autonomy*, in “IEEE Spectrum”, January 2017. <https://spectrum.ieee.org/cars-that-think/transportation/self-driving/toyota-gill-pratt-on-the-reality-of-full-autonomy>. Traduzione a cura dell’autrice.

## Nota bibliografica

- Evan ACKERMAN, *Toyota's Gill Pratt on Self-Driving Cars and the Reality of Full Autonomy*, in “IEEE Spectrum”, January 2017. <https://spectrum.ieee.org/cars-that-think/transportation/self-driving/toyota-gill-pratt-on-the-reality-of-full-autonomy>.
- U.S. ENERGY INFORMATION ADMINISTRATION, *Autonomous Vehicles: Uncertainties and Energy Implications*, 2018. <https://www.eia.gov/outlooks/aeo/pdf/AV.pdf>.
- Ronald Craig ARKIN, Patrick ULAM, e Alan R. WAGNER, *Moral Decision-making in Autonomous Systems: Enforcement, Moral Emotions, Dignity, Trust and Deception*, in “Proceedings of the IEEE”, n. 100, v. 3, March 2012, pp. 571-589.
- Bryan CASEY, *Amoral Machines, or: How Roboticists Can Learn to Stop Worrying and Love the Law*, in “Northwestern University Law Review”, n. 111, v. 5, anno 2017, pp. 231-250.
- Daniela CERQUI, Jutta WEBER, e Karsten WEBER (a cura di), *Ethics in Robotics*, in “International Review of Information Ethics”, n. 6, anno 2006.
- Edoardo DATTERI, Hykel HOSNI, e Guglielmo TAMBURRINI, *An inductionless and default-based analysis of machine learning procedures*, in L. MAGNANI (ed.), *Model Based Reasoning in Science and Engineering*, College Publications, London 2006, pp. 379-399.
- Adriano FABRIS, Sergio BARTOLOMMEI, e Edoardo DATTERI, *Quale etica per la robotica?*, in “Teoria”, n. XXVII, vol.2, anno 2007, pp. 7-17.
- David HAUSSLER, *On Decision Theoretic Generalizations of the PAC Model for Neural Net and Other Learning Applications*, in “Information and Computation”, n. 100, v. 1, September 1992, pp. 78-150.
- Patrick LIN, *Why Ethics Matters for Autonomous Cars*, in M. MAURER, J.C. GERDES, B. LENZ, H. WINNER (a cura di), *Autonomous Driving. Technical, Legal and Social Aspects*, Springer, Berlin Heidelberg 2016, pp. 69-85.

Patrick LIN, *Who's at Fault in Uber's Fatal Collision?*, in "IEEE Spectrum", March 2018. <https://spectrum.ieee.org/cars-that-think/transportation/safety/reflecting-on-ubers-fatal-crash>.

Aarian MARSHALL, *Lawyers, Not Ethicists, Will Solve the Robocar 'Trolley Problem'*, in "Wired", May 28, 2017. <https://www.wired.com/2017/05/autonomous-vehicles-trolley-problem/>.

Hans MORAVEC, *Mind Children. The Future of Robot and Human Intelligence*, Harvard University Press, Harvard 1990.

Abbe MOWSHOWITZ, *Technology as excuse for questionable ethics*, in "AI & Society", n. 22, v. 3, anno 2008, pp. 271-282.

Fiorella OPERTO, *Ethics in Advanced Robotics. ELS Issues in Advanced Robotics*, in "IEEE Robotics & Automation Magazine – Special Issues on Robot Ethics, Robotics and Automation Society", n. 18, v. 1, anno 2011, pp. 72-78.

Bruno SICILIANO, Oussama KHATIB (a cura di), *Springer Handbook of Robotics*, II Edition, Springer, Berlin Heidelberg 2016.

Dezhen SONG, Ken GOLDBERG, Nak Young CHONG, *Networked Telerobots*, in B. SICILIANO e O. KHATIB (a cura di), *Springer Handbook of Robotics*, cit., pp. 759-771.

Guglielmo TAMBURRINI, *Robot ethics: a view from the philosophy of science*, in R. CAPURRO e M. NAGENBORG (a cura di), *Ethics and Robotics*, IOS Press, Heidelberg 2009, pp. 11-22.

Gianmarco VERUGGIO, *The EURON Roboethics Roadmap, European Robotics Research Network, Atelier on Roboethics, 2005-2007*. <http://www.roboethics.org>.

Gianmarco VERUGGIO, TedTalk, 2018, <https://www.youtube.com/watch?v=F068leHdOFo>.

Gianmarco VERUGGIO e Keith ABNEY, *Roboethics: the applied ethics for a new science*, in P. LIN, G. BEKEY, e K. ABNEY (a cura di), *Robot Ethics*, MIT Press, Cambridge 2011, pp. 347-363.

Gianmarco VERUGGIO, Fiorella OPERTO, e George BEKEY, *Roboethics: Social and Ethical Implications*, in B. SICILIANO e O. KHATIB (a cura di), *The Springer Handbook of Robotics*, cit., pp. 2135-2160.