

Francesca Cerea

Responsabilità civile e sistemi “intelligenti”



Giappichelli

CAPITOLO I

INTELLIGENZA ARTIFICIALE: INQUADRAMENTO CONCETTUALE E PROSPETTIVE TEORICHE IN TEMA DI SOGGETTIVITÀ

SOMMARIO: 1. Un problema di definizioni – 2. Sul funzionamento dei sistemi “intelligenti”: dagli algoritmi ai modelli di apprendimento. – 3. Intelligenza Artificiale e robotica: scienze distinte (con frequenti occasioni di intersezione). – 4. Sullo *status* dei prodotti “intelligenti”: *res* o soggetti giuridici digitali? – 5. Ricorso all’applicazione analogica degli artt. 2047-2049 c.c.: un evitabile snaturamento delle fattispecie. – 6. Le responsabilità da Intelligenza Artificiale: tra adattamenti e nuovi paradigmi.

1. *Un problema di definizioni*

Sarebbe considerevolmente più semplice discutere di Intelligenza Artificiale (IA) potendo quantomeno prendere le mosse da una definizione condivisa di questo tipo di tecnologia, la cui nascita risale convenzionalmente alla metà degli anni '50 del secolo scorso ¹.

¹Una completa visione dell’IA viene espressa per la prima volta da Alan Turing il cui lavoro segna il punto di svolta decisivo sul tema. V. A. TURING, *Computing Machinery and Intelligence*, in *Mind*, 1950, 433 ss. Il termine “Intelligenza Artificiale” viene però coniato qualche anno più tardi, nel 1956, da John McCarthy. V. J. MCCARTHY *et al.*, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, 1955. La *proposal*, elaborata da McCarthy, Minsky, Shannon e Rochester per un workshop di due mesi al Dartmouth College, divenuto poi il luogo di nascita ufficiale della disciplina, è reperibile per intero in R. CHRISLEY-S. BEGEER, *Artificial Intelligence. Critical Concepts in Cognitive Science*, Routledge, London, 2000 o consultabile all’indirizzo: <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>. Per una ricostruzione storica completa dello sviluppo di questa tecnologia v. N.J. NILLSON, *The Quest for Artificial Intelligence. A History of Ideas and Achievements*, Cambridge University Press, Cambridge, 2010; P. MCCORDUCK, *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*, CRC Press, New York, 2004; P. STONE *et al.*, *Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel*, Stanford University, Stanford, 2016, consultabile all’indirizzo: <http://ai100.stanford.edu/2016-report>; G. SARTOR, *L’informatica giuridica e le tecnologie dell’informazione*, Giappichelli, Torino, 2012, 277 ss. Per un resoconto degli avvenimenti e dei protagonisti del Dartmouth Summer Re-

Tuttavia, a fronte dell'imponente applicazione di sistemi artificiali negli ambiti più disparati², si è assistito nel corso del tempo ad un proliferare sempre maggiore di tentativi definitori. È infatti possibile individuare all'interno della grande mole di descrizioni di Intelligenza Artificiale fornite in letteratura due distinte modalità di avvicinamento al fenomeno: un primo approccio, che si potrebbe definire “pensiero vs azione”, vede contrapporsi l'idea che l'intelligenza consista esclusivamente in processi di pensiero e ragionamento alla diversa idea che essa includa altresì l'interazione con l'ambiente; un secondo approccio, definibile come “umano vs razionale”, pone in alternativa definizioni che misurano il successo di un sistema intelligente in base alla capacità di riprodurre l'intelligenza umana, per sua natura limitata, e definizioni che utilizzano come parametro uno standard ideale, quale il concetto di razionalità, prescindendo dai limiti della razionalità umana³.

Su altro piano si pone poi la nota distinzione filosofica tra Intelligenza Artificiale debole (*weak Artificial Intelligence*) e Intelligenza Artificiale forte (*strong Artificial Intelligence*).

Con l'espressione IA debole ci si riferisce all'idea che le macchine possano agire in modo intelligente o, per meglio dire, come se fossero intelligenti; per IA forte si intendono invece quegli artefatti che sarebbero effettivamente in grado di pensare, non limitandosi ad una mera simulazione del pensiero umano⁴. L'Intelligenza Artificiale debole è quella di cui effettivamente oggi

search Project alla luce dei successivi cinquanta anni di ricerca si veda J. MOOR, *The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years*, in *AI Magazine*, 2006, 87 ss.

²Dalla medicina all'industria, dai trasporti alla finanza, dall'informazione alla difesa, dalla giustizia all'educazione, sono questi solo alcuni dei settori fortemente intaccati dall'impiego di tecnologie intelligenti.

³La distinzione è contenuta in S.J. RUSSEL-P. NORVIG, *Artificial Intelligence. A modern approach*, Englewood Cliffs, Prentice Hall, 2010, 4. Si rinvia al capitolo I del loro lavoro per l'approfondimento degli approcci citati che sono stati tutti oggetto di sviluppo nella storia dell'IA.

⁴Tale caratterizzazione si deve a J.R. SEARLE, *Minds, Brains and Programs*, in *The Behavioural and Brain Science*, 1980, 417 ss. il quale sostiene peraltro, mediante il celebre esperimento mentale della cosiddetta “stanza cinese”, l'impossibilità teorica di realizzare sistemi informatici capaci di pensiero in senso proprio negando in linea di principio l'esistenza di un'IA forte, in netta contrapposizione con McCarthy. La dicotomia tra *strong AI* e *weak AI* appare in realtà ormai superata sul piano pratico, ben pochi ricercatori che si occupano di Intelligenza Artificiale ritengono infatti che l'esito di tale dibattito possa sortire una qualche ricaduta significativa e preferiscono pertanto dedicarsi allo studio dei meccanismi dell'intelligenza in quanto tale. Non a caso per riferirsi all'IA in senso forte si utilizza anche l'espressione GOFAI (*Good Old-Fashioned Artificial Intelligence*), coniata da J. HAUGELAND, *Artificial Intelligence. The Very Idea*, MIT Press, Cambridge, 1985 che sta ad indicare che si tratta di un sistema superato essendo stato pressoché abbandonato sin dagli anni '90 a causa di evidenti limiti concettuali. V. L. FLORIDI, *Philosophy and Computing. An Introduction*, Routledge, London and New York, 1999, 132 ss. Principale sosteni-

si dispone, essendo l'Intelligenza Artificiale forte stata abbandonata ormai diverso tempo addietro, per essere, in realtà, più una impostazione filosofica che un vero e proprio approccio alla creazione⁵.

Ancora, sul piano pratico è possibile rinvenire, da un lato, sistemi intelligenti puramente *software-based*, in grado di agire nel mondo virtuale, quali assistenti vocali, *software* per l'analisi delle immagini, strumenti di elaborazione del linguaggio naturale, motori di ricerca, sistemi di riconoscimento vocale e facciale e dall'altro sistemi informatici di AI incorporati in dispositivi *hardware*, si pensi, solo per fare qualche esempio, a robot avanzati, auto a guida autonoma o droni⁶. Peraltro, la distinzione tra i sistemi esperti appena richiamati non è sempre netta, potendo gli uni essere integrati negli altri; è il caso, solo per fare un esempio, dei robot sociali o dei droni che incorporano il modello linguistico attualmente più noto: Chat-GPT⁷.

Nonostante lo scenario multiforme sin qui accennato un comune sforzo definitorio a livello europeo è stato compiuto prima dallo stesso Parlamento⁸, poi dalla Commissione europea nell'ambito della predisposizione

tore della corrente critica in tema di *strong AI* fu il filosofo Hubert Dreyfus. V. H. DREYFUS, *What Computers Can't Do: The Limits of Artificial Intelligence*, Harper e Row, New York, 1979; ID, *What Computers Still Can't Do: A Critique of Artificial Reason*, MIT Press, Cambridge, 1992; H. DREYFUS-S. DREYFUS, *Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*, Basil Blackwell, Oxford, 1986.

⁵ M. BROUSSARD, *La non intelligenza artificiale. Come i computer non capiscono il mondo*, Franco Angeli, Milano, 2019. Così l'autrice: «L'intelligenza artificiale forte (o generale) è quella di Hollywood. È l'intelligenza di robot senzienti (che potrebbero voler conquistare il mondo), di computer coscienti, di macchine che "pensano" come gli umani. L'intelligenza artificiale debole (o limitata) è diversa: è un metodo matematico per la previsione».

⁶ Cfr. sul punto il lavoro dell'*High-Level Expert Group on Artificial Intelligence* nominata dalla Commissione europea dal titolo "*A definition of AI: Main Capabilities and Disciplines*", Bruxelles, 8.4.2019 che riprende la definizione di Intelligenza Artificiale contenuta nella Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale europeo e al Comitato delle Regioni dal titolo "*L'intelligenza artificiale per l'Europa*", Bruxelles, 25.4.2018.

⁷ ChatGPT è un modello linguistico (prodotto da OpenAI) addestrato su un enorme mole di testi e interazioni umane, che gli consente di generare risposte coerenti e grammaticalmente corrette a un'ampia gamma di suggerimenti e domande poste in linguaggio naturale. Attualmente i test svolti da Microsoft intendono verificare se il sistema sia in grado di andare oltre la mera interpretazione testuale e approcciarsi alle peculiarità del mondo fisico per essere efficacemente applicato in robotica. Lo scopo è quello di consentire una interazione uomo-macchina più semplice, che non costringa l'utilizzatore ad apprendere complessi linguaggi di programmazione o dettagli tecnici circa il sistema robotico. Si rinvia all'articolo di Microsoft: *ChatGPT for Robotics: Design Principles and Model Abilities*, 2023.

⁸ Si veda a riguardo la Risoluzione del Parlamento europeo recante "*Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*", 16.2.2017 e in particolare l'allegato alla Risoluzione.

di un piano coordinato sull'intelligenza artificiale⁹ e da ultimo nel c.d. AI Act¹⁰ – come emendato nel giugno 2023 – il quale considera come «sistema di intelligenza artificiale» qualsiasi «sistema automatizzato progettato per operare con livelli di autonomia variabili e che, per obiettivi espliciti o impliciti, può generare output quali previsioni, raccomandazioni o decisioni che influenzano gli ambienti fisici o virtuali» (art. 3, punto 1).

Si tratta della *Proposta di Regolamento del Parlamento Europeo e del Consiglio in tema di Intelligenza Artificiale* che, tra gli altri profili, si occupa anche di fornire una definizione comune di sistemi di AI che sia chiara, al fine di garantire la certezza del diritto e al tempo stesso flessibile, per agevolare i futuri sviluppi tecnologici (considerando 6)¹¹.

Al netto di future eventuali modifiche, la nozione di Intelligenza Artificiale di cui all'AI Act appare sufficientemente ampia da ricomprendere l'eterogenea gamma di applicazioni che sono espressione di questo genere di tecnologia e nel contempo esaustiva per il giurista che voglia indagare il tema.

⁹La Commissione considera sistemi di AI tutti quei «sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere specifici obiettivi». La definizione si rinviene nella Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale europeo e al Comitato delle Regioni dal titolo “*L'intelligenza artificiale per l'Europa*”, Bruxelles, 25.4.2018. L'*High-Level Expert Group on Artificial Intelligence*, nel già citato lavoro, partendo proprio da questa nozione di IA, ha poi formulato una definizione più articolata, che è stata, peraltro, considerata dalla Commissione europea quale punto di partenza per lo sviluppo di una definizione operativa di AI. Cfr. il *Technical report* del *Joint Research Center (JRC)* della Commissione europea dal titolo *Defining Artificial Intelligence. Towards and Operational Definition and Taxonomy of AI*, Luxembourg, 27.2.2020.

¹⁰*Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione*, 21.4.2021, COM/2021/206 final.

¹¹La Proposta di Regolamento risale al 21.4.2021 e non ha ad oggi ancora terminato l'iter legislativo. Tuttavia, il 6 dicembre 2023 il Parlamento e il Consiglio europeo hanno raggiunto un accordo politico provvisorio che rappresenta un significativo passo in avanti verso la finalizzazione di un testo definitivo. Nei mesi successivi i lavori proseguiranno per la definizione degli aspetti tecnici del Regolamento, il quale dovrà poi essere portato all'attenzione del Coreper e nuovamente approvato da entrambe le istituzioni. V. il comunicato stampa reperibile sul sito del Consiglio: *Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world*. Per una breve analisi della Proposta v. M. TAMPIERI, *L'intelligenza artificiale e le sue evoluzioni. Prospettive civilistiche*, Wolters Kluwer Cedam, Milano, 2022, 48 ss.

2. Sul funzionamento dei sistemi “intelligenti”: dagli algoritmi ai modelli di apprendimento

Per il proprio funzionamento qualunque sistema di Intelligenza Artificiale si avvale di algoritmi, vale a dire di sequenze di istruzioni elementari che consentono di risolvere un problema trasformando i dati iniziali del problema stesso nel risultato, cioè nell'insieme di dati finali che rappresentano la soluzione¹². Per dirla in modo ancor più elementare, nell'asciutto linguaggio della matematica, l'algoritmo è: «una successione finita di passi (intesi come istruzioni), ognuno dei quali definito ed eseguibile, che opera su dati producendo risultati»¹³. Affinché tali istruzioni siano in grado di governare il funzionamento di una Intelligenza Artificiale è però necessario compiere un ulteriore passaggio, vale a dire trasformare l'algoritmo in un programma, traducendo le istruzioni algoritmiche in un linguaggio di programmazione comprensibile alla macchina. Invero, trovato e definito l'algoritmo, il programma non ne è che il pedissequo svolgimento¹⁴.

L'Intelligenza Artificiale è dunque definibile come una branca dell'informatica che imita le funzioni cognitive degli esseri umani mediante il ricorso ad algoritmi, che utilizzano una imponente mole di dati e una costellazione di tecniche e approcci diversi (metodi simbolici, metodi che si basano sulla concatenazione di regole (se X allora Y), metodi statistici, approcci di apprendimento automatico, solo per citarne alcuni) per riuscire a esprimere in formule la complessità del comportamento umano.

Nel corso del tempo l'IA ha visto affermarsi in modo particolare due paradigmi principali: quello simbolico (approccio *top-down*)¹⁵, basato sulla ma-

¹²T.H. CORMEN *et al.*, *Introduction to Algorithms*, MIT Press, Cambridge, 2013 volendo fornire una definizione informale di algoritmo si esprimono in questi termini: «Informally, an algorithm is any well-defined computational procedure that takes some value, or set of values, as input and produces some value, or set of values, as output. An algorithm is thus a sequence of computational steps that transform the input into the output».

¹³A. BRUSAMOLIN MANTOVANI, *Elementi di informatica*, Cedam, Padova, 69.

¹⁴«Come l'architetto disegna il progetto di un edificio (e in esso può consistere l'originalità) e poi lo affida all'ingegnere e al geometra (a seconda della sua complessità) per l'esecuzione, così in perfetta analogia, l'analista disegna il progetto di un programma, riportando in esso le fasi dell'algoritmo pensato e poi lo affida al programmatore per la sua realizzazione (anche se in questo caso «realizzare» non significa «costruire» o «fabbricare», ma creare un nuovo scritto, destinato, però, questa volta, ad essere «letto» ed applicato non più dall'uomo, ma dal computer)». A tale efficace metafora fa ricorso Cfr. R. BORRUSO, *Computer e diritto. Analisi giuridica del computer*, Giuffrè, Milano, 1988, 192.

¹⁵L'elaborazione di questo metodo, su cui si basa molta di quella che potrebbe essere chiamata IA “classica”, si deve a Newell e Simon, autorevoli pionieri dell'indirizzo da loro denominato “Ipotesi del sistema simbolico fisico” (*physical system hypothesis*). A. NEWELL-

nipolazione di simboli come metodo di approssimazione dell'intelligenza umana, e quello sub-simbolico (approccio *bottom-up*), costruito invece sull'applicazione di procedure statistiche o numeriche. I modelli simbolici facenti ricorso alla logica, pur essendo sistemi che giungono a risultati in modi facilmente comprensibili per un osservatore umano, presentano limiti sia in termini di *performance* sia in termini di capacità di apprendimento. Le tecniche sub-simboliche, invece, godono oggi di un rinnovato successo proprio grazie alla loro maggiore efficacia, ma scontano il fatto di giungere a risultati difficilmente spiegabili e interpretabili *ex post*. Proprio a fronte degli svantaggi e vantaggi propri di ciascun approccio, diverse sono le ricerche che mirano ad unificare e sfruttare in maniera sinergica il paradigma simbolico e sub-simbolico per ottenere “tecniche ibride”.

Questo è ciò che attualmente un'IA è realmente in grado di fare, nonostante spesso l'(ab)uso di parole antropomorfe quali *Machine Learning* (apprendimento automatico), *Neural Network* (rete neurale) e *Deep Learning* (apprendimento profondo) induca a pensare che le macchine siano davvero in grado di pensare, di avere consapevolezza o addirittura empatia¹⁶.

In effetti, il ricorso alla stessa idea di “intelligenza” risulta fuorviante perché sembra lasciar intendere che esista un'entità, benché artificiale, a cui è consentito attribuire tale qualità, tipicamente umana (e, in altre forme, animale). Forse, come opportunamente messo in luce in dottrina, sarebbe più saggio allontanare il *focus* dalla definizione di intelligenza antropica, prestando, invece, attenzione al risultato del processo del sistema. In questo modo, «se il processo è qualificato intelligente quando è svolto da un essere umano, allora lo si può qualificare intelligente se è svolto da una macchina. Quindi l'Intelligenza Artificiale può essere definita la scienza di far fare ai computer cose che richiedono intelligenza quando vengono fatte dagli esseri umani»¹⁷.

H.A. SIMON, *Computer Science as Empirical Inquiry: Symbols and Search*, in *Communications of the ACM*, 1976, 113 ss.

¹⁶ A. VESPIGNANI, *L'algoritmo e l'oracolo. Come la scienza predice il futuro e ci aiuta a cambiarlo*, Il Saggiatore, Milano, 2019. Sul punto anche A. SANTOSUOSSO, *Intelligenza artificiale e diritto. Perché le tecnologie di IA sono una grande opportunità per il diritto*, Mondadori, Milano, 2020, osserva: «[v]i è un limite conoscitivo che porta gli umani a riprodurre quella che considerano la migliore intelligenza disponibile, cioè quella umana. Questo limite porta a usare un linguaggio antropomorfo e, quindi, ad antropomorfizzare le macchine che produciamo».

¹⁷ Così G. FINOCCHIARO, *Intelligenza artificiale e responsabilità*, in *Contr. e impr.*, 2020, 2, 713 che aggiunge: «Appare dunque un metodo meno fuorviante e meno condizionante nell'approccio metodologico, rispetto ad altri metodi spesso adottati senza piena consapevolezza metodologica, e che non induce ad assunzioni implicite di soggettività, quello di confrontare i risultati di due processi: se il primo è qualificato intelligente, la medesima qualificazione si attribuirà anche al secondo. È utile e anzi necessario sfrondare il di-

Non pare allora superfluo, alla luce degli accennati fraintendimenti derivanti dall'antropomorfizzazione del discorso attorno alle tecnologie in parola, cercare di mettere un po' d'ordine nella terminologia appena citata partendo proprio dall'espressione *Machine Learning* (ML), con cui si indica uno specifico settore del vasto campo dell'Intelligenza Artificiale e segnatamente «quella branca dell'informatica che permette ad una macchina di imparare ad eseguire un compito senza essere stata esplicitamente programmata per farlo»¹⁸. Il sistema, esattamente come accade per applicazioni più semplici, osserva e raccoglie informazioni sull'ambiente che lo circonda ma in più, ed è questo l'elemento caratterizzante, ipotizza l'esistenza di una legge statistica sottostante che cerca di scoprire e di utilizzare per il compimento di attività future¹⁹.

L'idea che sta alla base di questo approccio di tipo sub-simbolico è che l'algoritmo impara da solo a identificare relazioni precise presenti nei dati osservati, senza avere regole e modelli espliciti pre-programmati. Così facendo, invece di scrivere algoritmi con milioni di istruzioni specifiche per svolgere un compito particolare, l'algoritmo di *machine learning* apprende dai dati migliorando le proprie prestazioni in modo adattivo a mano a mano che impara dalle informazioni che sta elaborando. Quanto più grande è il set di dati, tanto più accurata sarà l'individuazione delle relazioni anche impercettibili tra gli stessi²⁰.

scorso dalla retorica che inesorabilmente lo popola. La fattispecie si delinea così in modo neutro e senza "pre-giudizi". Un conto è affermare che un robot intelligente ha operato in un certo modo, altro è affermare che un *software* basato su un'applicazione di intelligenza artificiale ha prodotto un determinato risultato. Le parole usate e la narrazione scelta possono portare con sé un condizionamento».

¹⁸La definizione si deve a Arthur Lee Samuel il cui lavoro è probabilmente il primo esempio di ricerca nel campo dell'apprendimento automatico coronato dal successo. V. A.L. SAMUEL, *Some studies in machine learning using the game of checkers*, in *IBM Journal of Research and Development*, 1959, 210 ss. Più formale la definizione di *machine learning* fornita da Tom M. Mitchell nel 1997 che si esprime in questi termini: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E". V. T.M. MITCHELL, *Machine Learning*, McGraw-Hill, New York, 1997.

¹⁹A prima vista potrebbe sembrare lo stesso percorso verso la conoscenza del mondo che effettua l'uomo, ma c'è una significativa differenza: mentre l'uomo, attraverso un meccanismo di astrazione, cerca una spiegazione generale dei fenomeni particolari che osserva, l'agente, in questa sua ricerca, si concentra unicamente su quel determinato fenomeno e non va alla ricerca di una legge universale. Ciò che interessa all'agente è trovare la particolare funzione che lega certi *input* a certi *output*, un modello assai più semplice, dunque, dell'astrazione umana. Così G. D'ACQUISTO, *Intelligenza artificiale. Elementi*, Giappichelli, Torino, 2021, 127 ss.

²⁰Mette in luce l'importanza di grandi volumi di dati al fine del funzionamento ottimale dei sistemi di IA anche la Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni,

Gli innumerevoli algoritmi di ML vengono tradizionalmente suddivisi in tre categorie.

In base al modo in cui sono strutturati i dati analizzati e alle metodologie di apprendimento usate si distingue, invero, tra apprendimento automatico supervisionato (*supervised*), non supervisionato (*unsupervised*) e rinforzato (*reinforced*).

Pur a rischio di una semplificazione eccessiva è possibile descrivere le differenze tra i tre modelli nel modo che segue.

Nel caso di *supervised learning* (apprendimento supervisionato), il più diffuso, al sistema vengono forniti sia dei dati in input sia le informazioni relative ai risultati desiderati, affinché esso identifichi in autonomia una regola generale che colleghi i dati in ingresso con quelli in uscita, in modo poi da poter utilizzare tale regola per l'esecuzione di compiti simili. I dati forniti al sistema (ad es. immagini, file di testo, video) sono, dunque, etichettati, cioè annotati con l'identificazione di determinate proprietà o caratteristiche, o classificati in base a ciò che rappresentano, con l'obiettivo di fornire maggiore contesto ai *dataset*. Nonostante tale attività (c.d. *data labeling*)²¹ rappresenti un significativo dispendio di tempo e di risorse economiche per le aziende del settore, garantisce dati generalmente più accurati capaci di migliorare l'affidabilità delle previsioni del modello.

Si parla invece di *unsupervised learning* (apprendimento non supervisionato) quando al sistema vengono forniti solamente set di dati in entrata privi di qualsivoglia indicazione circa il risultato sperato. Spetta quindi al sistema risalire agli schemi e ai modelli che si celano negli input individuando la struttura logica senza che questi siano preventivamente etichettati. Tra le applicazioni più comuni di tale approccio rientrano alcuni dispositivi di diagnosi medica per il rilevamento, la classificazione e la segmentazione

dal titolo “*L'intelligenza artificiale per l'Europa*” COM/2018/237 final, Bruxelles, 25.4.2018. La Commissione, infatti, osserva: «Sono necessari ingenti volumi di dati per sviluppare l'IA. L'apprendimento automatico, un tipo di IA, opera mediante l'individuazione di modelli a partire dai dati disponibili e la successiva applicazione di questa conoscenza ai dati nuovi. Quanto più è grande il set di dati, tanto più accurata sarà l'individuazione delle relazioni anche impercettibili tra i dati. Quando si tratta di utilizzare l'IA, gli ambienti ad alto contenuto di dati offrono anche le maggiori opportunità, perché i dati sono il mezzo attraverso il quale l'algoritmo apprende e interagisce con il suo ambiente».

²¹ L'attività di etichettatura dei dati può avvenire, in realtà, secondo metodi diversi ciascuno dei quali presenta vantaggi e svantaggi. La scelta può dipendere dalla complessità dell'attività, nonché dalla dimensione, dall'ambito e dalla durata del progetto. Si rinvia alla panoramica dei principali approcci all'annotazione dei dati riportata sul sito di IBM alla voce *Etichettatura dei dati*. Tutti gli approcci sono, comunque, soggetti ad errore umano (ad esempio, errori di codifica, errori di immissione manuale), il che può compromettere la qualità dei dati e determinare imprecisioni nella loro elaborazione e nella creazione di modelli. Per tale ragione i controlli per garantire la qualità del processo di annotazione sono fondamentali per preservare la qualità dei dati.

delle immagini, impiegati in radiologia per effettuare diagnosi rapide e accurate sui pazienti²².

La terza categoria definita *reinforced learning* (apprendimento per rinforzo) costituisce invece una sorta di sottoinsieme dell'apprendimento non supervisionato poiché anche in questo caso i dati forniti non sono etichettati. Tuttavia, la tecnica di apprendimento è differente laddove in tale ipotesi il sistema deve interagire con un ambiente dinamico in cui è immerso e da cui riceve dati di input e raggiungere un obiettivo, al raggiungimento del quale riceve una ricompensa, imparando dagli errori identificati mediante "penalità".

Esistono poi altre numerose sottocategorie di *Machine Learning* che identificano in realtà altrettanti approcci pratici di applicazione degli algoritmi di autoapprendimento²³, tra cui rientra una delle forme più popolari ed efficaci di sistemi sub-simbolici di apprendimento nota come rete neurale artificiale (*artificial neural network*).

Questa classe di algoritmi, ispirata all'analogia con la struttura e il funzionamento del cervello biologico, pur avendo una storia lunga quanto quella dell'Intelligenza Artificiale²⁴, rispetto alla quale si è mossa in parallelo, ha visto la propria ascesa con il nuovo millennio.

²²Tale modalità di apprendimento è impiegata anche nel settore *marketing*, per la definizione della personalità dei clienti al fine di comprendere i tratti comuni e le abitudini di acquisto degli stessi. L'apprendimento senza supervisione (e in particolare il *clustering*, vale a dire una tecnica di estrazione dei dati che raggruppa dati non etichettati in base alle loro somiglianze o differenze) consente alle aziende di creare profili della personalità degli acquirenti migliori e, di conseguenza, di allineare i messaggi promozionali ai prodotti in modo più appropriato.

²³Per una spiegazione dei diversi modelli di algoritmi cui si fa ricorso v. C. O'NEIL, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown Publishing Group, New York, 2016 (trad. it. *Armi di distruzione matematica*, Bompiani, Milano, 2017); G. D'ACQUISTO, *op. cit.*, 161 ss.

²⁴Lo studio delle reti neurali ha inizio sin dagli anni '40 con il fondamentale lavoro di McCulloch e Pitts, in cui viene proposto un primo modello formale di neurone, con il duplice obiettivo di capire come funziona il cervello e di progettare macchine in grado di simularne il comportamento. Si veda W.S. MCCULLOCH-W. PITTS, *A logical calculus of the ideas immanent in nervous activity*, in *Bulletin of Mathematical Biophysics*, 5, 115 ss., 1943. Si deve però a Rosenblatt (F. ROSENBLATT, *The Perceptron – A Perceiving and Recognizing Automaton*, in *Cornell Aeronautical Laboratory*, 1957; ID., *The perceptron: A probabilistic model for information storage and organization in the brain*, in *Cornell Aeronautical Laboratory*, 1958, 65, 6, 386) la formulazione del primo modello di rete di neuroni in grado di apprendere a riconoscere elementi, modificando la struttura delle connessioni: il c.d. "Perceptrone" (*Perceptron*). Dopo il lavoro del 1969 di Minsky e Papert (M. MINSKY-S. PAPERT, *Perceptrons*, MIT Press, Cambridge, 1969) volto a mettere in luce i significativi limiti del Perceptrone, la ricerca sulle reti neurali viene fortemente ridimensionata (si parla di "primo inverno dell'AI") fino all'inizio degli anni '80 quando l'algoritmo *Backpropagation* ridona nuova linfa alle reti neurali. V. D.E. RUMELHART-G.E. HINTON-R.J. WILLIAMS, *Learning Internal Representations by Error Propagation*, in D.E. RUMELHART-J.L.

È infatti solo recentemente che, grazie alla centuplicata capacità computazionale dei computer, allo sviluppo dell'*Internet of Things* (IoT), ormai declinabile come *Internet of Everything* (IoE) e all'enorme quantità di dati (*Big Data*) di cui oggi gli algoritmi di ML possono disporre, hanno iniziato ad essere sviluppate reti neurali di dimensioni maggiori articolate su più livelli di apprendimento che hanno condotto all'evoluzione *self-learning* del fenomeno AI, facendo compiere a questa tecnologia un salto qualitativo determinante²⁵.

Le reti neurali sono formate da neuroni artificiali interconnessi, il cui meccanismo è ispirato ad una schematizzazione del funzionamento dei neuroni del sistema nervoso centrale umano. L'idea alla base del sistema è che l'unità elementare di elaborazione, il neurone artificiale appunto, sia in grado di acquisire dati in input nel livello di neuroni in ingresso (*input layer*) e fornire dati di output attraverso il livello di neuroni in uscita (*output layer*). Fra questi due livelli possono essere presenti uno o più strati per così dire “nascosti” (*hidden layers*) di neuroni artificiali che si occupano della vera e propria elaborazione dei dati. Il numero di livelli intermedi presente tra quello in ingresso e quello in uscita consente di distinguere tra *simple neural networks*, ove il livello “nascosto” è uno soltanto e *deep neural networks*, vale a dire reti neurali profonde in quanto formate da un numero di livelli intermedi maggiore di uno²⁶. Nella gran parte dei casi le reti neurali profonde possiedono molti strati, ciascuno con un gran numero di neuroni, per un

MCCLELLAND, *Parallel Distributed Processing – Explorations in the Microstructure of Cognition*, vol. I, MIT Press, Cambridge, 1986, 318 ss. Dopo un altro “inverno dell'AI”, nel 1995 Vapnik e Cortes introducono nuovi algoritmi di apprendimento automatico detti *Support Vector Machines* (SVM) molto efficaci per attività come la categorizzazione di testi e l'elaborazione del linguaggio naturale (C. CORTES-V. VAPNIK, *Support vector networks*, in *Machine Learning*, 20, 273 ss., 1995). A partire dal 2012, con l'avvento di sempre maggiori quantità di dati, cresce l'attenzione nei confronti delle “Reti Neurali Convolutionali” (*Convolutional Neural Networks* (CNN)) introdotte dal pionieristico lavoro degli anni '90 di Yann LeCun e le reti neurali profonde diventano una componente fondamentale dell'informatica moderna (Y. LECUN *et al.*, *Gradient-based learning applied to document recognition*, in *Proceedings of the IEEE*, 86, 11, 1998, 2278 ss.). Nel 2018 lo stesso LeCun, con i colleghi Hinton e Benjo, vengono insigniti del Turing Award per il loro apporto concettuale e ingegneristico al campo dell'IA in tema di reti neurali artificiali.

²⁵ Cfr. T. TAULLI, *Artificial Intelligence Basics. A Non-Technical Introduction*, Apress, New York, 2019; R. ROVATTI, *Intelligenza Artificiale e “Big Data”: un percorso da dati a informazioni (con molteplici occasioni di errore)*, in C. BOTTARI (a cura di), *La salute del futuro. Prospettive e nuove sfide del diritto sanitario*, Bononia University Press, Bologna, 2020, 175 ss. Taulli sottolinea come il rapido progresso dei sistemi esperti sia, soprattutto, dovuto alla crescita esponenziale dei volumi di dati, alla creazione di nuove e più performanti infrastrutture informatiche e all'impiego di una nuova tecnologia di chip (le c.d. GPU, unità di elaborazione grafica) dotata di una velocità molte volte superiore a quella delle CPU convenzionali e, dunque, in grado di diminuire significativamente i tempi dei processi di elaborazione.

²⁶ T. TAULLI, *op. cit.*

totale di molte migliaia, a volte milioni, di unità tra loro collegate tramite connessioni “pesate”²⁷. Ogni connessione, invero, non è altro che un valore numerico (il “peso” appunto) che ne stabilisce l’importanza relativa (fornisce una misura di quanto “conta” tale *input* nel neurone) e viene moltiplicato per il valore del neurone successivo collegato. Oltre che dai “pesi” il comportamento di una rete neurale dipende dai c.d. “bias”, vale a dire da valori numerici costanti da sommarsi alla moltiplicazione che coinvolge il “peso” menzionato²⁸. È attraverso il continuo aggiustamento di “pesi” e “bias” (operazione che avviene negli strati “nascosti”, durante la fase di addestramento) che una rete neurale è in grado di generalizzare e modellizzare un problema nel mondo reale, modulando l’output e l’input di ogni singolo neurone fino al raggiungimento di una soluzione accettabile. Proprio a causa della immensa mole di numeri coinvolti nessuna rete neurale profonda possiede parametri definiti manualmente da un programmatore, essendo per converso sottoposta ad un processo di *Machine Learning*.

Sarebbe quindi sufficiente parlare di reti neurali per includere la parte legata al *Deep Learning* (DL), trattandosi di un metodo di addestramento e utilizzo delle stesse; tuttavia, tale espressione ha ormai acquisito una propria autonomia di impiego, a fronte del fatto che la gran parte degli esempi ipotizzabili nel mondo *Machine Learning* si riferiscono al DL (si pensi, solo per fare qualche esempio, ai *software* per il riconoscimento di immagini e ai veicoli a guida autonoma).

3. *Intelligenza Artificiale e robotica: scienze distinte (con frequenti occasioni di intersezione)*

Spesso oggetto di confusione e di sovrapposizione con il campo dell’Intelligenza Artificiale è la robotica, vale a dire quel settore dell’ingegneria che si occupa della progettazione e costruzione di robot.

Esattamente come per la gran parte delle tecnologie complesse non esiste una definizione univoca e ampiamente condivisa di robot, essendo piuttosto rinvenibile una grande mole di descrizioni offerte nel corso del tempo, di pari passo con l’evoluzione della scienza e della tecnologia in grado di creare macchine sempre più sofisticate²⁹.

²⁷ A mero titolo di esempio si tenga conto che AlexNet, sistema di IA per il riconoscimento di immagini, è composta da più di 800 mila neuroni ed è definita da più di 60 milioni di parametri. Nel 2012 AlexNet ha dimostrato di saper riconoscere un’immagine con un’accuratezza maggiore rispetto a quella dell’essere umano medio. V. R. ROVATTI, *op. cit.*, 180 ss.

²⁸ L’output di un neurone è, quindi, espresso dalla formula $output = input \times pesi + bias$.

²⁹ Emblematiche le parole di J. ENGELBERGER, *Robotics in service*, MIT Press, Cam-

Il termine “robot” è stato impiegato, per la prima volta, in RUR (Rossum’s Universal Robots), una commedia del 1921 del drammaturgo Karel Čapek, che, su suggerimento del fratello Josef, utilizzò la parola ricavata dal ceco “robota”, ossia “lavoro forzato”, per riferirsi ad operai creati dall’uomo e destinati al suo servizio³⁰.

Complice una ricostruzione fantascientifica diffusasi nell’immaginario collettivo attraverso i *mass media*, la letteratura e la filmografia sul tema, i robot vengono sovente intesi come macchine che pensano e si comportano come un essere umano. Le cose non stanno esattamente così ed è, pertanto, necessaria qualche precisazione.

Va anzitutto operato un primo distinguo tra macchine dotate di un corpo (*embodied*) e macchine che non ce l’hanno (*non-embodied*). Tali categorie sono a loro volta suscettibili di ulteriore declinazione essendo possibile distinguere tra macchine, sia *embodied* che *non-embodied*, dotate di qualche forma di Intelligenza Artificiale e macchine che ne sono prive³¹.

Le macchine *embodied* prive di IA sono in grado di attuare movimenti e di produrre lavoro fisico. Si tratta di artefatti programmati per operare in modo automatico, svolgendo lavori gravosi o ripetitivi per aumentare la produttività e le prestazioni degli operatori che li usano. Essendo guidati dall’uomo (o programmati da lui) non prendono decisioni autonome: il loro

bridge, 1989 secondo cui: «*I can’t define a robot, but I know one when I see one*» che danno il senso della impossibilità di individuare una sola definizione condivisa del termine “robot”. Il problema viene messo in luce, tra gli altri, anche da A. SANTOSUOSSO-C. BOSCARATO-F. CAROLEO, *Robot e diritto: una prima ricognizione*, in *Nuova giur. civ. comm.*, 7-8, 2012, 494 ss. Gli Autori segnalano la mancanza di un generale consenso su cosa sia un robot e quali macchine possano essere qualificate come tali, elencando una serie di possibili criteri di classificazione. L’assenza di una definizione scientifica idonea ad abbracciare l’intera, variegata tipologia delle macchine “intelligenti” viene evidenziata anche da C. SALAZAR, *Umano, troppo umano... o no? Robot, androidi e cyborg nel “mondo del diritto” (prime notazioni)*, in *BioLaw Journal – Rivista di BioDiritto*, 1/2014, 258-259; N. NEVEJANS, *Les robots: tentative de définition*, in A. BENSAMOUN, *Les robots*, Mare & Martin, Le Kremlin-Bicêtre, 2015, 79 ss.

³⁰ K. ČAPEK, *R.U.R. Rossum’s Universal Robots*, Prague, 1920. Il termine “robot” acquista ancor più popolarità grazie all’opera di Asimov, cui si deve la formulazione delle c.d. tre leggi sulla robotica contenute per la prima volta nel racconto *Runaround* (in italiano “circolo vizioso”) e riprese dall’Autore in altre opere. Le tre leggi, nel loro testo originale, sono: “(1) *A robot may not injure a human being or, through inaction, allow a human being to come to harm.* (2) *A robot must obey any orders given to it by human beings, except where such orders would conflict with the First Law.* (3) *A robot must protect its own existence as long as such protection does not conflict with the First or Second Law*». I. ASIMOV, *Runaround in Astounding Science Fiction*, Street & Smith, New York, 1942; ID., *I, Robot*, Collins, London, 1968.

³¹ La classificazione si deve a R. CINGOLANI-D. ANDRESCIANI, *Robots, macchine intelligenti e sistemi autonomi: analisi della situazione e delle prospettive*, in G. ALPA (a cura di), *Diritto e intelligenza artificiale*, Pacini, Pisa, 2020, 23 ss.

agire dipende dal programma o dal teleoperatore, cioè dall'essere umano. Si pensi, solo per fare qualche esempio, ai robot cui si fa ricorso in ambito spaziale o in situazioni di pericolo (incendi e terremoti) o all'automazione industriale.

In tempi recenti alcune macchine *embodied* sono state dotate di Intelligenza Artificiale acquisendo crescenti capacità cognitive e decisionali che consentono loro di prendere scelte autonome, senza il controllo di un operatore³². Possono essere macchine non antropomorfe (per esempio le auto a guida autonoma), oppure veri e propri umanoidi sviluppati per interagire con gli esseri umani.

I prodotti tecnologici *non-embodied* non sono in grado di compiere lavoro o di fare movimenti e fanno parte di quelle tecnologie comunemente chiamate digitali. Anche per questo tipo di prodotti è possibile distinguere tra macchine senza corpo prive di IA (come, ad esempio, la tv e la radio) e macchine "*non-embodied*" dotate di IA (si pensi a sistemi di analisi delle immagini radiografiche per la diagnosi di patologie o a *software* di riconoscimento del linguaggio naturale come il già citato Chat-GPT)³³.

L'impiego di sistemi dotati di AI solleva questioni giuridiche piuttosto complesse, tra cui la possibile attribuzione di soggettività giuridica a simili artefatti. L'ipotesi nasce, anzitutto, con riferimento ai robot *embodied* intelligenti³⁴ che rappresentano il connubio tra due scienze, la robotica e l'Intelligenza Artificiale, nate da realtà e tecnologie diverse (la meccatronica, da un lato e la *computer science*, dall'altro). Il contaminarsi di questi due mondi tecnologici porta, infatti, alla creazione di quelle che vengono chiamate macchine intelligenti e autonome (*AIS Autonomous Intelligent Systems*)³⁵.

È pensando a questo genere di macchine che viene messa a punto la descrizione, ad oggi più diffusa, di robot come «*an autonomous system which exists in the physical world, can sense its environment, and can act on it to*

³² Si tratta di prodotti che scontano quello che prende il nome di Paradosso di Moravec. Si deve, infatti, allo studioso austriaco Hans Moravec la formulazione della tesi secondo la quale il ragionamento di alto livello richiede pochissimo calcolo, mentre le capacità motorie, anche relativamente semplici, necessitano di capacità computazionali elevatissime. È per tale ragione che per sistemi robotici dotati di AI risulta più semplice avere prestazioni paragonabili a quelle umane sul piano logico-razionale, mentre abilità sensoriali e motorie banali risultano spesso inarrivabili. V. H. MORAVEC, *Mind Children. The Future of Robot and Human Intelligence*, Harvard University Press, Cambridge, 1988, 15.

³³ Il settore delle tecnologie digitali ha visto uno sviluppo sensibilmente più rapido del settore robotico proprio per le difficoltà connesse al "*Paradosso di Moravec*" richiamato nella nota precedente.

³⁴ Per poi essere formulata, come si dirà, anche rispetto ai sistemi intelligenti privi di supporto fisico (*software*).

³⁵ R. CINGOLANI-D. ANDRESCIANI, *op. cit.*, 29.

*achieve some goals»*³⁶. Esiste infatti un certo consenso attorno all'idea che «*robots are mechanical objects that take the world in, process what they sense, and in turn act upon the world»*³⁷.

Si tratta di artefatti il cui avvento costituisce un passo fondamentale nella storia della scienza e della tecnologia e possiede notevoli implicazioni anche sul piano antropologico, legale ed etico. Ciò è dimostrato dall'attenzione dedicata al tema dal Parlamento europeo che, al fine di stabilire una definizione comune europea di robot autonomo intelligente, ne ha individuato le seguenti caratteristiche: a) capacità di acquisire autonomia grazie a sensori e/o mediante lo scambio di dati con il proprio ambiente (interconnettività) e l'analisi di tali dati; b) capacità di apprendimento attraverso l'esperienza e l'interazione; c) forma del supporto fisico del robot; d) capacità di adeguare il suo comportamento e le sue azioni all'ambiente³⁸.

Come accennato, tali connotazioni del prodotto intelligente inducono alcuni giuristi a chiedersi quale sia lo *status* privatistico da riconoscere all'artefatto robotico e, in particolare, se lo stesso vada ricompreso nella categoria di prodotto o se sia opportuno qualificarlo come un soggetto agente, con ricadute significativamente diverse sul piano della disciplina in tema di responsabilità civile.

Questo, a ben vedere, appare il senso pratico di tentare un riconoscimento di personalità al sistema di Intelligenza Artificiale: individuare colui che, sul piano civile, possa essere tenuto a rispondere dell'eventuale danno cagionato; ma una tale osservazione non può non evidenziare come la questione divenga ancor più complessa, perché a nulla serve l'attribuzione di una qualche "soggettività" ad un ente, se allo stesso non si riconosca al contempo anche una capacità patrimoniale in grado di far fronte alle pretese creditorie e una capacità di costituirsi e stare in giudizio per difendersi dalle stesse.

Ma su tali questioni si tornerà nel prosieguo.

³⁶ La definizione si deve a M.J. MATARIĆ, *The Robotics Primer*, MIT Press, Cambridge, 2007.

³⁷ V. CALO, *Robotics and the lessons of cyberlaw*, in *Cal. Law. Rev.*, 2015, 529. Il "sense-think-act cycle" rappresenta una modellizzazione dell'intelligenza umana particolarmente influente nel settore robotica e Intelligenza Artificiale. V. R. PFEIFER-C. SCHEIER, *Understanding intelligence*, MIT Press, Cambridge, 1999, 37; R.A. BROOKS, *Intelligence Without Reason*, MIT Press, Cambridge, 1991, 569 s.

³⁸ V. Risoluzione del Parlamento europeo recante *Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*, 16.2.2017 e in particolare l'allegato alla Risoluzione nella parte dedicata a *Definizione e classificazione dei "robot intelligenti"*.

4. *Sullo status dei prodotti “intelligenti”: res o soggetti giuridici digitali?*

L'autonomia e la capacità di apprendere e acquisire esperienza mediante processi decisionali di natura algoritmica, caratteristiche essenziali del robot intelligente, hanno spinto ad interrogarsi sulla natura di soggetto giuridico dell'artefatto robotico e dunque sulla sua possibile inclusione nelle categorie giuridiche esistenti o in una nuova categoria *ad hoc*³⁹.

A ciò non si è sottratto il Parlamento europeo che, nella già citata Risoluzione, propone «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»⁴⁰.

³⁹ Nel corso degli anni si è assistito a diverse proposte per estendere una sorta di personalità giuridica alle tecnologie digitali emergenti, alcune risalenti addirittura al secolo scorso. Si vedano ad esempio L. SOLUM, *Legal Personhood for Artificial Intelligences*, in *North Carolina Law Review*, 70, 1992, 1231; S. LEHMAN-WILZIG, *Frankenstein Unbound: Towards a Legal Definition of Artificial Intelligence*, in *Futures*, 13(6), 1981, 442; G. TADDEI ELMI, *I diritti dell'intelligenza artificiale tra soggettività e valore: fanta-diritto o ius condendum?*, in L. LOMBARDI VALLAURI (a cura di), *Il meritevole di tutela*, Giuffrè, Milano, 1990, 685 ss.; ID., *Soggettività artificiali e diritto*, relazione tenuta il 3 giugno 2004 in occasione della chiusura del Master sul Diritto delle nuove tecnologie informatiche organizzato dal CSIG sezione di Firenze; H. PUTNAM, *Robots: Machines of Artificially created Life?*, in *Journal of Philosophy*, 1964, 668 ss.; P. ZIFF, *The feelings of Robots*, in *Analysis*, 1959, 64 ss.; G. TEUBNER, *Rights of Non-humans? Electronic Agents and Animals as New Actors in Politics and Law*, in *Journal of Law and Society*, 2006, 497 ss. Più recentemente cfr. G. TEUBNER, *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi*, Edizioni Scientifiche Italiane, Napoli, 2019, *passim*; R. VAN DER HOVEN VAN GENDEREN, *Do we need to legal personhood in the age of robot and IA?*, in M. CORRALES-M. FENWICK-N. FORGÓ (eds.), *Robotics, AI and the Future of Law*, Springer, Singapore, 2018, 15 ss.; H. EIDENMÜLLER, *The rise of robots and the law of humans*, in *Oxford Legal Studies Research Paper No. 27/2017*, 2017, 765 ss.

⁴⁰ Tale posizione si discosta peraltro dallo studio della Direzione generale delle Politiche interne del Parlamento europeo nel quale si legge: «We believe it would be inappropriate and out-of-place not only to recognise the existence of an electronic person but to even create any such legal personality. Doing so risks not only assigning rights and obligations to what is just a tool, but also tearing down the boundaries between man and machine, blurring the lines between the living and the inert, the human and the inhuman. Moreover, creating a new type of person – an electronic person – sends a strong signal, which could not only reignite the fear of artificial beings but also call into question Europe's humanist foundations. Assigning person status to a non-living, non-conscious entity would therefore be an error since, in the end, humankind would likely be demoted to the rank of a machine. Robots should serve humanity and should have no other role, except in the realms of science-fiction». V. N. NEVEJANS, *European Civil Law Rules in Ro-*

L'impostazione appena descritta sembra dare per assodata l'inadeguatezza dell'attuale sistema di responsabilità civile, contrattuale ed extracontrattuale, a regolare il settore delle tecnologie intelligenti, inadeguatezza cui si fornisce risposta ipotizzando di attribuire una qualche soggettività al robot dotato di AI.

Il problema si è posto anche in dottrina ed è stato affrontato mediante la teorizzazione da parte di alcuni di una «soggettività giuridica piena» e di altri di una «soggettività giuridica parziale» degli "agenti software"⁴¹.

La prospettiva di ricorrere ad una non meglio specificata «personalità giuridica elettronica», piena o parziale, seppur teoricamente e tecnicamente praticabile dal punto di vista giuridico rischia⁴², però, di risolversi nell'introduzione di una mera *fictio* foriera più di equivoci interpretativi ed esiti abnormi che di soluzioni realmente efficaci⁴³.

A riguardo va anzitutto svolta qualche precisazione circa il concetto di

botics, Study for the JURI Committee, ottobre 2016. Sul punto v. F. CAROCCIA, *Soggettività giuridica dei robot?*, in G. ALPA (a cura di), *Diritto e intelligenza artificiale*, cit., 222 ss.; F. SARZANA DI S. IPPOLITO, *La legislazione internazionale in tema di intelligenza artificiale: dai diritti civili alla soggettività della persona elettronica*, in A. D'ALOIA (a cura di), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, Franco Angeli, Milano, 2020, 217 ss.; R. TREZZA, *Diritto e Intelligenza artificiale. Etica-Privacy-Responsabilità-Decisione*, Pacini, Pisa, 2020, 56 s. Quest'ultimo osserva come l'ipotesi del Parlamento europeo «potrebbe incentivare una totale "deresponsabilizzazione" del produttore e dell'ideatore, un pericoloso *laissez faire*, considerato che i produttori e gli ideatori agirebbero su un mercato globale, con regole non uniformi [...]».

⁴¹ L'espressione è utilizzata da G. TEUBNER, *Ibridi ed attanti. Attori collettivi ed enti non umani nella società e nel diritto*, Mimesis, Milano, 2015. L'Autore teorizza una capacità giuridica degli agenti elettronici delimitata all'ambito della rappresentanza e della responsabilità degli ausiliari, idea poi ripresa da Teubner anche in *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi*, cit., 63-90. La questione ha suscitato interesse anche nella dottrina spagnola: come dimostrano, tra gli altri, J. ERCILLA GARCIA, *Normas de derecho civil y robótica. Robots inteligentes, personalidad jurídica, responsabilidad civil y regulación*, Aranzadi, Navarra, 2018, *passim*; C. ROGEL VIDE (ed.), *Los robots y el derecho*, Reus, Madrid, 2018, *passim*.

⁴² Come osserva F. CAROCCIA, *op. cit.*, 232 ss. non sembrano esservi ostacoli teorici che impediscano al legislatore di considerare le macchine come soggetti, potendo il diritto introdurre le soluzioni tecniche che ritiene adeguate a rispondere ad esigenze pratiche.

⁴³ Di tale avviso A. AMIDEI, *Robotica Intelligente e responsabilità: profili e prospettive evolutive del quadro normativo europeo*, in U. RUFFOLO (a cura di), *Intelligenza artificiale e responsabilità*, Giuffrè, Milano, 2017, 63 ss.; G. PASSAGNOLI, *Ragionamento giuridico e tutele nell'intelligenza artificiale*, in *Persona e Mercato*, 3, 2019, 79 ss.; G. COMANDÈ, *Responsabilità e accountability nell'era dell'Intelligenza Artificiale*, in F. DI CIOMMO-O. TROIANO (a cura di), *Giurisprudenza e Autorità indipendenti nell'epoca del diritto liquido*, Studi in onore di Roberto Pardolesi, La Tribuna, Piacenza, 2018, 1010; E. PALMERINI, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, in *Resp. civ. e prev.*, 6, 2016, 1837 ss.; M. TAMPIERI, *L'intelligenza artificiale e le sue evoluzioni. Prospettive civilistiche*, cit., 103 ss.

autonomia su cui spesso si fa leva per il riconoscimento di uno *status* giuridico specifico in capo ai sistemi di IA. Secondo l'impostazione del Parlamento europeo l'autonomia, definita come «*capacità di prendere decisioni e metterle in atto nel mondo esterno*»⁴⁴, solleva diversi interrogativi sia sul piano della riconducibilità di tali sistemi alla categoria delle *res*, sia circa la capacità delle ordinarie regole in tema di responsabilità civile di disciplinare il tema, tenuto conto del fatto che «*più i robot sono autonomi, meno possono essere considerati come meri strumenti nelle mani di altri attori*»⁴⁵.

Va però chiarito che quando ci si riferisce ad una capacità della macchina di apprendere (*learning*) o auto-apprendere (*self-learning*) autonomamente (*rectius*: automaticamente) si intende dire che la macchina è diventata più accurata nell'eseguire un singolo, specifico compito in base a precisi parametri definiti da una persona⁴⁶.

Questo tipo di apprendimento, più automatico che autonomo, non implica intelligenza o, per meglio dire, implica un'intelligenza tarata sulla funzione da svolgere, «*un'intelligenza algoritmica molto precisa e riproducibile, specializzata e mirata alla soluzione di specifiche classi di problemi ma poco flessibile, poco adattativa e poco fantasiosa*»⁴⁷.

Per questo motivo non deve impressionare il fatto che un'IA possa battere il campione del mondo di scacchi, né da ciò vi è da desumere che il robot sia realmente intelligente o dotato di autocoscienza⁴⁸, tenuto conto che la

⁴⁴ Risoluzione del Parlamento europeo recante «*Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*», 16.2.2017 punto AA.

⁴⁵ *Ivi*, punto AB.

⁴⁶ Come chiarito dalla stessa Commissione europea nella Relazione al Parlamento europeo, al Consiglio e al CESE dal titolo «*Relazione sulle implicazioni dell'intelligenza artificiale, dell'Internet delle cose e della robotica in materia di sicurezza e di responsabilità*», 19.2.2020, 8, nota 38: «*Finora l'espressione "capace di autoapprendimento" è stata utilizzata nell'ambito dell'intelligenza artificiale per lo più per indicare che le macchine sono in grado di apprendere durante l'addestramento; non è ancora un requisito che le macchine dotate di intelligenza artificiale continuino ad apprendere anche dopo la loro messa in funzione; al contrario, in particolare nel settore della salute, le macchine dotate di intelligenza artificiale cessano normalmente di apprendere dopo che il loro addestramento si è concluso con successo. Pertanto, per il momento il comportamento autonomo dei sistemi di intelligenza artificiale non implica che il prodotto svolga compiti non previsti dagli sviluppatori*».

⁴⁷ R. CINGOLANI, *L'altra specie, Otto domande su noi e loro*, Il Mulino, Bologna, 2019, 108 ss. In tema anche N. CRISTIANINI, *La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano*, Il Mulino, Bologna, 2023, 9 ss. Sul punto v. anche M. CHIRIATTI, *Incoscienza artificiale. Come fanno le macchine a prevedere per noi*, LUISS University Press, Roma, 2021, 71 «il calcolo senza significato può al massimo esprimere l'ossimoro dell'"intelligenza incosciente". Per comprendere il senso dei calcoli occorre un'esperienza cosciente».

⁴⁸ La gran parte dei ricercatori esclude che sarà possibile in futuro ottenere un'Intelligenza Artificiale cosciente, tuttavia ci si interroga sulle possibili sfide etiche, sociali e re-

macchina è in grado di eseguire soltanto quell'attività e, per riuscirci, consuma una quantità di energia infinitamente più grande di quella che consuma l'uomo⁴⁹.

Ne consegue che la macchina pur potendo migliorare nello svolgimento di compiti programmati è incapace di comprendere i rapporti causa-effetto tra i dati raccolti⁵⁰ e di compiere un giudizio autenticamente valoriale⁵¹, li-

gatorie che un sistema algoritmico *apparentemente* consapevole potrebbe comunque portare con sé. Sul tema v. A. SETH, *Why Conscious AI Is a Bad, Bad Idea*, in *Nautilus*, 2023. Dello stesso autore *We Are Beast Machines*, *ivi*, 2021: «Consciousness has more to do with being alive than with being intelligent. We are conscious selves precisely because we are beast machines. Experiences of being you, or of being me, emerge from the way the brain predicts and controls the internal state of the body. The essence of selfhood is neither a rational mind nor an immaterial soul. It is a deeply embodied biological process, a process that underpins the simple feeling of being alive that is the basis for all our experiences of selfhood». Dello stesso avviso M. CHIRIATTI, *op. cit.*, 31 ss.

⁴⁹ Il riferimento è alla nota partita tra il computer Deep Blue prodotto da IBM e Garry Kasparov avvenuta nel maggio 1997 e spesso utilizzata come simbolo dell'ascesa dell'Intelligenza Artificiale. V. F. HSU, *Behind Deep Blue: Building the Computer that Defeated the World Chess Champion*, Princeton University Press, Princeton, 2002. Come scrive il programmatore e consulente Neville-Neil: «Abbiamo avuto quasi mezzo secolo di competizioni tra gli esseri umani e il computer al gioco degli scacchi, ma questo significa forse che uno qualunque di quei computer sia intelligente? No, niente affatto. Per due ragioni. La prima è che gli scacchi non sono una prova di intelligenza; sono una prova di una particolare abilità – l'abilità nel giocare a scacchi. Se fossi in grado di battere un campione mondiale di scacchi; eppure, fossi incapace di passare il sale a tavola a chi me lo chiede, sarei intelligente? La seconda ragione è che pensare in termini scacchistici era una prova di intelligenza basata sulla falsa premessa culturale che bravi giocatori di scacchi fossero menti brillanti, persone intellettivamente più dotate delle persone attorno a loro. Sì, molte persone intelligenti possono eccellere negli scacchi, ma gli scacchi, o qualunque altra singola abilità, non denotano intelligenza». Così G.V. NEVILLE-NEIL, *The Chess Player Who Couldn't Pass the Salt*, in *Communications of the ACM*, 60, 4, March 24, 2017, richiamato da M. BROUSSARD, *op. cit.*, 121 s.

⁵⁰ I sistemi di Intelligenza Artificiale si dimostrano efficienti nell'individuare, se esistono, correlazioni tra i dati, le quali però possono anche essere del tutto casuali e fuorvianti. Si pensi al noto caso di "Compas", sistema progettato per supportare i giudici americani nelle loro decisioni, al fine di calcolare la probabilità di recidiva degli imputati. L'algoritmo individuava correlazioni errate e discriminatorie fra tratti somatici e propensione a delinquere, in particolare tendeva a sottostimare il rischio di recidiva quando il convenuto era caucasico e a sovrastimarne quando era un afroamericano. Risultati simili dipendono dal fatto che per quanto l'IA sia efficiente nel far emergere correlazioni tra i dati, non lo è altrettanto nella valutazione del rapporto causa-effetto tra gli stessi. M. CHIRIATTI, *op. cit.*, 69 ss.

⁵¹ Prerequisito di qualsiasi giudizio di tipo valoriale è l'averne una coscienza. La coscienza persegue il senso di giustizia attraverso l'intelligenza – che la definisce – e la volontà, che la attua. Sul punto si richiama l'esempio tratto da una vicenda statunitense proposto da E. CATERINI, *Artificial Intelligence, persona e soggetto*, in *Tecnologie e Diritto*, 2022, 223 «Il software rileva anche il lieve eccesso di velocità al passare dell'auto. Il giudice esonera da responsabilità l'autista ultranovantenne, ammalato di tumore, che alla guida dell'auto accompagna in ospedale il figlio ultrasessantenne per una importante tera-

mitandosi a risolvere in modo efficiente problemi, anche difficili, all'interno di un dominio specifico, attraverso criteri quantitativi. Il concetto di "autonomia" della macchina ad oggi non è dunque assoluto, ma trova il suo limite e la sua reale dimensione se collocato entro una cornice funzionale, nella quale il robot altro non è che una entità specializzata capace di eseguire con ottimi risultati i compiti per cui è stato programmato, ma incapace di eseguire tutti gli altri⁵². Ogni previsione offerta da un sistema algoritmico è una valutazione quantitativa, mai qualitativa, diversamente da ciò che avviene nelle decisioni umane che solo raramente possono dirsi frutto di semplice calcolo, essendo basate su valori incommensurabili e quindi incomputabili⁵³.

Certamente il linguaggio ormai diffuso quando si discute dell'operatività dei sistemi di AI, descritta abitualmente in termini di "azione", "decisione", "scelta", si presta a facili fraintendimenti e può solo apparentemente dirsi neutrale, laddove finisce per attribuire a procedure automatizzate (di fatto fondate su analisi di dati, modelli statistici e formulazioni previsionali) quei presupposti che caratterizzano la razionalità e la capacità deliberativa dell'essere umano⁵⁴.

pia. Quale intelligenza artificiale avrebbe saputo giudicare il valore dell'azione di quell'uomo? Soltanto il senso concreto e il valore etico della giustizia ha saputo raggiungere tale conclusione». Sul punto anche P. PERLINGIERI, *Sul trattamento algoritmico dei dati*, ivi, 2020, 184.

⁵² Risulta quindi piuttosto artificioso il frequente confronto tra intelligenza naturale e Intelligenza Artificiale, laddove nel primo caso si tratta di un'intelligenza globale e libera di svilupparsi secondo direzioni imprevedibili, mentre la seconda è un'intelligenza commisurata allo scopo da raggiungere e priva di caratteristiche quali intuito, capacità di improvvisare ed emotività che rendono gli umani così diversi tra loro. V. R. CALO, *Robotics and the Lessons of Cyberlaw*, in *California Law Review*, 2015, 513 ss. il quale sostiene che «*There are analytic and technical reasons to believe robots will never think like people*» e R. CINGOLANI, *L'altra specie, Otto domande su noi e loro*, cit., 110 che, attraverso una efficace metafora, definisce il robot come «*un centometrista dell'intelletto, mentre l'umano è un decatleta della mente*».

⁵³ Di questo avviso M. CHIRIATTI, *op. cit.*, 108. Sul punto anche E. CATERINI, *Artificial Intelligence, persona e soggetto*, cit., 225 s. osserva: «La saggezza, intesa come qualità umana capace di ponderare e penetrare le cose e gli uomini, non è programmabile; la giustizia, intesa come qualità dell'uomo capace di contemperare l'interesse proprio con quello degli altri, tanto da esprimersi in un'esperienza equa e corretta, non è programmabile».

⁵⁴ Mette opportunamente in luce questo aspetto M. ZANICHELLI, *Ecosistemi, opacità, autonomia: le sfide dell'intelligenza artificiale in alcune proposte recenti della Commissione europea*, in A. D'ALOIA (a cura di), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, cit., 76 ss. L'Autrice, opportunamente, osserva come: «Si estende oggi alle macchine artificiali forse in modo troppo acritico e scontato, almeno sul piano del lessico, l'attribuzione di quel processo decisionale che nell'agire pratico è premessa necessaria di ogni scelta. Proprio l'atto del decidere dovrebbe invece identificare una delle principali