Intelligenza artificiale e *imaging* diagnostico
Implicazioni per il TSRM

Giornata mondiale della radiologia
8 novembre 2020

La Sezione Aspetti giuridici e medico-legali della FNO TSRM e PTSRP elabora pareri inerenti il campo di esercizio dei professionisti, sulla base della normativa vigente e della più autorevole letteratura, avvalendosi di esperti dello specifico settore, medici legali, giuristi e qualsiasi altro professionista la cui competenza è utile a dirimere i quesiti oggetto di studio e approfondimento.

I medesimi sono, altresì, vagliati dall'Ufficio legale della FNO TSRM e PSTRP e necessitano di approvazione da parte del Comitato centrale.

Tali pareri, sottoposti a periodica revisione, possono, inoltre, costituire un'occasione di confronto e di crescita interprofessionale. Per tale motivo, eventuali osservazioni e/o suggerimenti possono essere inviate al seguente indirizzo di poste elettronica: federazione@tsrm.org.

Naturalmente, questo parere costituisce espressione di una attività meramente consultiva e non già di amministrazione attiva, che non può avere natura immediatamente applicativa; in particolare si deve segnalare che ogni questione va affrontata tenendo conto degli aspetti specifici e del contesto particolare che l'ha generata.

Dunque, l'espressione generale di questo parere non può sostituirsi agli opportuni e specifici pareri relativi al caso personale e concreto.

Autori

Marta D'Agostino	Laurea in Giurisprudenza
Massimiliano Paganini	Laurea in Scienze cognitive e processi decisionali

Coordinamento

Roberto Di Bella	Laurea magistrale in Scienze delle professioni sanitarie tecniche diagnostiche

Presentazione

L’intelligenza artificiale è uno strumento; molto sofisticato, ma pur sempre soltanto uno strumento.

Il modo in cui l’intelligenza artificiale impatterà sulla società, sulla sanità, sulle persone assistite, sulle professioni sanitarie e sul TSRM dipenderà dal modo in cui lo strumento sarà gestito (o non gestito).

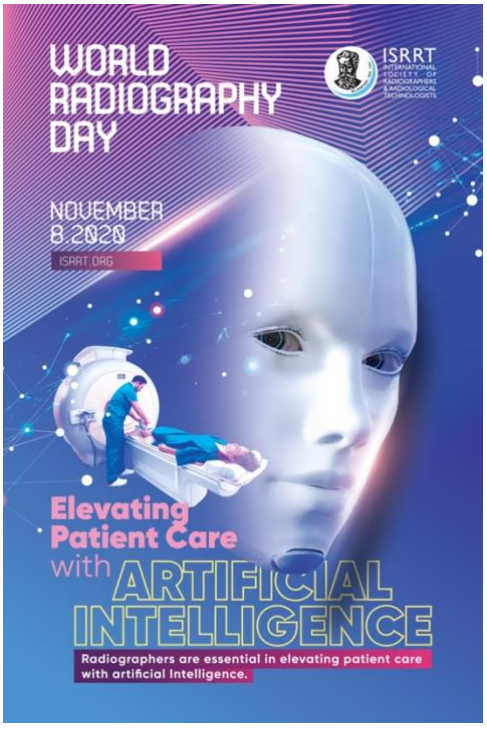
Un uso responsabile, ad impatto sicuro e positivo, dell’intelligenza artificiale è possibile solo se la si conosce in tutti i suoi aspetti: teoretico, logico, informatico, operativo, giuridico, medico-legale, etico, etc...

Questa monografia del gruppo Aspetti giuridici e medico-legali della Federazione nazionale, curata dal collega Massimiliano Paganini e dall’Avv. Marta D’Agostino, affronta uno degli aspetti sopra elencati, valutandone le valenze per il TSRM.

La sua lettura consentirà ai TSRM di avere una prima qualificata rappresentazione del tema, che dovrà necessariamente far parte della formazione universitaria e continua dei Tecnici sanitari di radiologia medica, affinché l’intelligenza artificiale sia uno degli strumenti a loro disposizione. E non il contrario.

Presidente
Commissione di albo nazionale TSRM
Carmela Galdieri

Presidente
FNO TSRM e PSTRP
Alessandro Beux



Intelligenza artificiale e imaging diagnostico.
Implicazioni per il Tecnico sanitario di radiologia medica

SOMMARIO: 1. Introduzione; 2. L'Intelligenza Artificiale nel mondo; 2.1. In particolare: in Italia; 3. Qualificazione giuridica dell'Intelligenza Artificiale; 3.1. Alcune considerazioni di carattere preliminare alla qualificazione giuridica e alla ricostruzione del nesso di causalità; 3.2. L'Intelligenza Artificiale intesa come se fosse un bambino, un dipendente una attività pericolosa, una cosa o un animale; 3.2.1 Le Intelligenze Artificiali utilizzate nell'ambito sanitario e la Legge 8 marzo 2017, n. 24; 3.3. L'Intelligenza Artificiale intesa come un prodotto; 3.3.1 In particolare: l'Intelligenza Artificiale intesa come dispositivo medico; 3.4. L'Intelligenza Artificiale più evoluta considerata come una persona; 3.4.1 Le intelligenze Artificiali più evolute e il c.d. *dilemma situation*; 3.4.2 *EPersons*: rilievi critici e prospettive; 4. Le difficoltà nella ricostruzione del nesso causale: *in dubio pro machina;* 5. I dati trattati ed utilizzati nell'addestramento delle Intelligenze Artificiali; 6. Conclusioni, La "Competenza Artificiale": monitoraggio, formazione e informazione dei TSRM quali strumenti principali per fronteggiare il fenomeno.

1. Introduzione. – Nel 1956 presso l'Università di Dartmouth (Hannover) si svolse il *Dartmouth Summer Research Project on Artificial Intelligence*. Tale evento organizzato da John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon viene considerato l'inizio della ricerca nell'ambito dell'Intelligenza Artificiale (termine coniato per la prima volta da McCarthy).

In realtà, la nascita dell'Intelligenza Artificiale può essere fatta risalire ad Alan Mathison Turing che nella prima parte del Novecento aveva teorizzato una macchina ideale in grado di eseguire algoritmi e dotata di un nastro potenzialmente infinito su cui leggere e/o scrivere simboli.

L'idea di Intelligenza Artificiale di Turing è contenuta in un articolo del 1950; il matematico (ma anche logico, crittografo, filosofo) per spiegare la propria idea si avvalse di un gioco di società popolare nell'Inghilterra vittoriana, cioè il gioco dell'imitazione, che consisteva nel chiudere un uomo e una donna in stanze separate obbligandoli a rispondere alle domande di una terza persona. L'uomo doveva rispondere dicendo sempre la verità, la donna mentendo sempre. L'intervistatore doveva indovinare in quale stanza si nascondeva l'uomo e in quale la donna. Nella variante proposta da Turing, nelle stanze vengono chiuse una macchina e una persona e il gioco consiste nel capire qual è l'essere umano. Se la macchina riesce a ingannare l'intervistatore, allora ha superato il *test*, e secondo Turing dimostra di essere intelligente[1].

In altri termini, quella di Turing è una definizione empirica di "intelligenza": è intelligente chi si comporta in modo intelligente; è importante solo ciò che avviene fuori della stanza.

D'altra parte, empirica è anche la definizione di Intelligenza Artificiale fornita da McCarthy e Minsky: l'Intelligenza Artificiale consiste «*nel far sì che una macchina si comporti in modi che sarebbero definiti intelligenti se fosse un essere umano a comportarsi così*»[2].

A contestare tale concetto di intelligenza è John Searle, filosofo della mente, attraverso il noto esperimento della "stanza cinese". Searle ipotizza di chiudere in una stanza una persona alla quale vengono rivolte delle domande in cinese. Il soggetto in questione restituisce effettivamente all'esterno delle risposte in cinese, tuttavia, per farlo, ricorre alla consultazione di un manuale che non comprende. In sintesi, il soggetto restituisce risposte in cinese, ma non comprende il cinese. Se per

[1] A. M. TURING, *Computing Machinery and Intelligence*, in *Mind*, vol. 49, 1950, p. 433 ss..

[2] J. MCCARTHY *et al.*, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, August 31, 1955, in *AI magazine*, vol. 27, n. 4, 2006, p. 12.

Turing l'importante è solo ciò che si osserva al di fuori della stanza, per Searle è fondamentale, invece, anche ciò che avviene all'interno della medesima[3].

È proprio il termine "intelligenza" abbinato all'aggettivo "artificiale" ad essere al centro di una parte del dibattito, in quanto se si accettasse una definizione di "intelligenza" come quella fornita, ad esempio, dallo Zingarelli, cioè quale «*complesso delle facoltà mentali e pratiche che consentono all'uomo di ragionare, comprendere la realtà, fronteggiare situazioni nuove*» si potrebbe giungere alla conclusione di Margaret Boden, secondo la quale l'intelligenza implica la coscienza[4].

Come scrive Jerry Kaplan:

«*Se McCarthy avesse scelto una espressione più prosaica, che non suggeriva una sfida al dominio e alla cognizione umani, come "elaborazione simbolica" o "informatica analitica" ... il progresso nel campo sarebbe apparso per ciò che è: l'incessante avanzamento dell'automazione*»[5].

Per fare chiarezza, dunque, è utile partire dalla proposta, ampiamente condivisa, di Searle di suddividere il campo dell'IA in due filoni: "ipotesi forte" e "ipotesi debole" dell'Intelligenza Artificiale[6].

Per "ipotesi forte" dell'IA si intende la possibilità per l'uomo di costruire dei calcolatori che pensano e sanno di pensare, cioè delle macchine coscienti. Se di ciò si stesse parlando (e quindi anche questo scritto dovrebbe farlo) si dovrebbero affrontare argomenti quali: cos'è la coscienza, da dove origina, come funziona, quali sono le teorie della mente (dualiste, monistiche, fisicaliste, emergentiste, ecc.), mappe neurali di primo e secondo ordine, coscienza nucleare, coscienza estesa ed altro ancora.

Nel nostro caso, invece, quando si parla di "Intelligenza Artificiale" si sta facendo riferimento alla c.d. "ipotesi debole" della stessa, ossia alla possibilità di costruire macchine che agiscono "come se" pensassero senza esserne, tuttavia, consapevoli.

Si tratta, in altri termini, di macchine che, forse, è più corretto definire "competenti" invece che "intelligenti", nell'accezione della Boden.

La competenza di un sistema non è necessariamente abbinata alla comprensione da parte del medesimo. Come chiarisce Daniel Dennett, può esistere competenza senza comprensione, così come avviene nel caso della Natura; è possibile, cioè, creare livelli molto attendibili di grande competenza senza comprensione nel caso di compiti piuttosto circoscritti ed ancora, la competenza è un effetto emergente di sistemi di competenza senza comprensione. Cosicché, ad esempio, la Natura, che si muove con apparente intenzionalità, è in realtà priva di mente, e la selezione naturale è il risultato competenziale di un procedere per tentativi ed errori. In altri termini, esistevano ragioni molto tempo prima che esistessero esseri che ragionano: si può dire, pertanto, che prima arriva la competenza poi, eventualmente, la comprensione[7].

A questo punto, risulta necessario approfondire il concetto di "competenza", cominciando dalla definizione più condivisa, vale a dire quella proposta da Spencer e Spencer: «*caratteristica intrinseca di un individuo causalmente collegata ad una performance eccellente in una mansione*»[8].

---

[3] J. R. SEARLE, *Mind, Brains and Programs. A Debate on Artificial Intelligence*, in *The Behavioral and Brain Science*, vol. 3, 1980, p. 128-135.

[4] R. MARMO, *Algoritmi per intelligenza artificiale*, Milano, 2020.

[5] J. KAPLAN, *Intelligenza artificiale. Guida al futuro prossimo*, Roma, 2017, cap. 2.

[6] J. R. SEARLE, *op. cit.*.

[7] D. C. DENNETT, *Dai batteri a Bach. Come evolve la mente*, Milano, 2018.

[8] L.M. SPENCER, M. S. SPENCER, *Competenza nel lavoro*, Milano, 1995.

Nonostante la definizione di competenza di Spencer e Spencer abbia determinato una convergenza generale, la definizione medesima presenta delle ambiguità che lasciano spazio a diverse interpretazioni che gli autori stessi hanno cercato di colmare.

In modo particolare si è cercato di chiarire come l'elemento intrinsecamente legato all'individuo fosse la capacità intesa come "dotazione personale" che permette di eseguire con successo una determinata prestazione, quindi la possibilità di riuscita nell'esecuzione di un compito o, in termini più estesi, di una prestazione lavorativa. Questa possibilità è condizionata dall'attitudine, che rappresenta il substrato costituzionale di una capacità. La capacità è, dunque, espressione di un'attitudine che ha trovato condizioni esterne e interne favorevoli al suo manifestarsi in comportamenti e prestazioni. Rifacendosi ad Aristotele, si potrebbe asserire che l'attitudine è capacità in potenza che da sola, tuttavia, non costituisce la competenza la quale deve presentare necessariamente altri due requisiti, quali la conoscenza e l'esperienza.

Se Madre Natura, intrinsecamente capace di condurre la selezione naturale e l'implementazione di sistemi complessi, procede per tentativi ed errori, dietro all'IA vi è, invece, un progettista intelligente che trasferisce, almeno inizialmente, le informazioni necessarie affinché le macchine possano agire in modo competente.

Scopo del citato convegno di Dartmouth era quello di studiare ogni aspetto dell'apprendimento o qualsiasi caratteristica dell'intelligenza per essere in grado di simularli. Di seguito le esatte parole di McCarthy e Minsky:

«*Lo studio procederà sulla base della congettura che tutti gli aspetti dell'apprendimento o qualsiasi altra caratteristica dell'intelligenza possa essere di principio descritta in modo così specifico che una macchina la possa simulare. Verrà fatto un tentativo per scoprire come si possa fare in modo che le macchine usino il linguaggio, formulino astrazioni e concetti, risolvano tipi di problemi ora riservati agli esseri umani*»[9].

Per raggiungere tale obiettivo furono intrapresi due filoni differenti: il primo di stampo simbolista, percorso da Herbert Simon, Allen Newell, John McCarthy; il secondo di stampo connessionista, intrapreso da Frank Rosenblatt.

La logica simbolica è una branca della matematica che si occupa di rappresentare concetti e affermazioni come simboli, e poi definisce varie trasformazioni per manipolare quei simboli allo scopo di ragionare deduttivamente, dalle ipotesi alle conclusioni o induttivamente, dalle osservazioni alle ipotesi[10].

Il filone simbolista è caratterizzato da un approccio *top-down* all'intelligenza, che viene considerata una proprietà indipendente dal supporto, biologico o di altro tipo, e che può essere descritta in termini simbolici o astratti. L'ipotesi fondante del sistema simbolico è che i simboli sono alla base dell'azione intelligente la quale viene definita come «*un comportamento appropriato ai fini del sistema e capace di adattarsi a quanto l'ambiente richiede*»[11].

Già nel 1956 Simon e Newell, due tra i massimi esponenti dell'approccio simbolista, con il supporto del programmatore Cliff Shaw, erano riusciti ad implementare un sistema di Intelligenza Artificiale (*Logic Theorist*) in grado di procedere alla dimostrazione di teoremi matematici; il sistema fu in grado di dimostrare trentotto dei cinquantadue enunciati dei *Principia Mathematica* (1910-13) di Bertrand Russell e Alfred North Whitehead.

---

[9] MCCARTHY *et al.*, *op. cit.*.

[10] KAPLAN, *op. cit.*.

[11] NEWELL, SIMON, in KAPLAN, *op. cit.*, cap. 2.

Successivamente a *Logic Theorist*, l'approccio simbolista ha condotto, nel tempo, all'implementazione dei c.d. "sistemi esperti", vale a dire sistemi che emulano il processo decisionale in un ambito specifico e che funzionano grazie ad una conoscenza di base costituita da dati e regole e dall'azione di un motore di inferenza in grado di generare nuovi assiomi a partire da quelli originari: in grado, cioè, di produrre nuova conoscenza.

In funzione degli scopi per i quali sono stati progettati, i sistemi esperti utilizzano diversi tipi di linguaggio quali, ad esempio, la logica *booleana*, con i noti operatori *NOT*, *AND* e *OR*; la logica *Fuzzy*, per la peculiarità di meglio modellare il ragionamento umano; i teoremi *bayesiani*, per le decisioni in ambito di incertezza; gli alberi decisionali.

I sistemi esperti, che hanno avuto una prima fase di implementazione dagli anni '60 agli anni '80, hanno recentemente trovato nuova linfa nei sistemi di *planning*, in cui, data una determinata condizione iniziale nota, si giunge a una o più condizioni finali desiderate grazie ad uno specifico *set* di operazioni disponibili basate su deduzioni simboliche e ragionamento euristico. Quest'ultimo viene introdotto nel sistema per evitare e/o ridurre una delle maggiori limitazioni dei sistemi esperti, vale a dire quella che in gergo viene definita "l'esplosione combinatoria", che si instaura nel momento in cui il sistema deve esplorare un alto numero di opzioni di scelta. In altri termini, il ragionamento euristico esclude a priori la valutazione di scelte non favorevoli, riducendo lo spazio di ricerca.

I sistemi esperti, dunque, sono in grado di riprodurre le prestazioni di una persona esperta, o di un *team* di specialisti, in un determinato settore di attività; in campo radiologico potrebbero essere utilizzati per definire l'appropriatezza (giustificazione) di una prestazione in funzione della prestazione stessa, del quesito clinico, del sesso, dell'età, del distretto da irradiare, della dose.

In generale, come sopra spiegato, i sistemi esperti (proposizionali, probabilistici, *bayesiani*) sono utili per la pianificazione di una sequenza di azioni finalizzata a raggiungere un obiettivo.

È il caso dei sistemi basati sulla logica *Fuzzy* (logica sfumata – non proposizionale – introdotta da Lofti Zadeh nel 1956) in grado di gestire *input* numerici trasformandoli in un linguaggio comune (c.d. "processo di fuzzyficazione"), elaborare un ragionamento (grazie ad una tabella delle regole contenente il pensiero di esperti) e restituire delle decisioni (*output*) in forma numerica (c.d. "processo di defuzzyficazione"). Sempre rimanendo nel campo della diagnostica, utilizzando un sistema *Fuzzy* sarebbe possibile definire la quantità idonea di mezzo di contrasto da somministrare ad un paziente e la successiva idratazione partendo dal peso del paziente medesimo, dal valore della creatinina (meglio GFR) e dalla presenza di ulteriori fattori di rischio quali il diabete e la cardiopatia.

Diversamente da quello simbolista, il filone connessionista costituisce un approccio *bottom-up* basato sulla premessa che la struttura specifica del cervello è fondamentale per far emergere la cognizione; di conseguenza l'*hardware* su cui poggia l'Intelligenza Artificiale deve imitare la struttura cerebrale dell'essere umano. A tale conclusione era giunto anche John von Neumann il quale riteneva che l'intelligenza fosse una proprietà emergente dei neuroni che formano il cervello; di conseguenza, riproducendo con formule matematiche la rete di neuroni del cervello si può creare una forma di Intelligenza Artificiale.

Nel 1957 Rosenblatt riusciva a realizzare il primo percettrone, cioè la prima copia artificiale di un neurone: il percettrone poteva ricevere diversi *input* in ingresso (simulazione dei dendriti), ma poteva dare un solo vettore in uscita (simulazione dell'assone). La propagazione dell'informazione costruita in base ad una logica binaria (stringhe 0-1) poteva avvenire solo quando il percettrone raggiungeva un certo valore soglia (simulazione del potenziale d'azione).

Il percettrone di Rosenblatt era costituito da un solo strato di neuroni ed aveva capacità di calcolo limitate: era in grado di effettuare unicamente le classificazioni in cui le classi fossero separabili in modo lineare. Tale limite veniva messo in risalto negli anni '70 da un articolo di McCarthy che dava il suo contributo, insieme alla bassa capacità di calcolo delle macchine dell'epoca,

ad un raffreddamento degli entusiasmi nei confronti dello sviluppo connessionista dell'IA che determinava una diminuzione dei fondi nel campo della ricerca.

Intorno alla fine degli anni '80, tuttavia, con la crescita esponenziale della potenza tecnologica, riprendeva vigore anche l'approccio connessionista all'IA. Secondo la legge di Moore, ogni due anni la massima potenza di calcolo nel mondo cresce in maniera esponenziale; intorno all'anno 2000 un *computer* di prezzo medio era in grado di effettuare circa 108 operazioni al secondo, corrispondenti a quelle svolte dal cervello di un insetto, nel 2025 tale potenza salirà a 1016 operazioni al secondo, pari a quelle svolte dal cervello umano, mentre nel 2045 verrà raggiunta una potenza di calcolo pari a quella di tutta l'umanità. Nel 2007 fu lo stesso Moore a porre un limite alla propria legge legandolo alle dimensioni dei *transistor* che attualmente hanno una dimensione di circa 14 nm (obiettivo 5-7 nm), tuttavia la possibilità di implementare un *computer* quantistico prefigura una potenza di calcolo colossale in grado di oltrepassare i limiti posti dalla legge di Moore.

A partire dal percettrone di Rosenblatt, la ricerca in ambito connessionistico ha consentito la realizzazione di reti neurali multistrato che, superando i limiti del monostrato, hanno reso possibile le attuale applicazioni nel campo del *machine learning* e del *deep learning.*

Per *machine learning* intendiamo l'implementazione di algoritmi che permettono l'apprendimento di informazione a partire dai dati disponibili oltre alla capacità di predire nuove informazioni alla luce di quelle apprese. I calcolatori basati sul *machine learning* sono, quindi, delle macchine in grado di apprendere. Secondo la definizione di Tom M. Mitchell:

«*si dice che un programma per computer impara dall'esperienza E nei confronti di una certa classe di compiti C e insieme di misure di performance P se le sue performance sui compiti di C, secondo le misure definite in P, migliorano grazie all'esperienza E*»[12].

I sistemi di *machine learning* sono ottimali per svolgere compiti di classificazione; dati degli *input* in ingresso il sistema restituirà degli *output* classificati. Per poter svolgere tali compiti, tuttavia, le macchine vanno addestrate; tale addestramento può avvenire in modo "supervisionato", così da insegnare al calcolatore a fronte degli *input* utilizzati quali siano gli *output* desiderati; oppure in modo "non supervisionato", fornendo alla macchina unicamente gli *input* e lasciando che sia la medesima a scoprire modelli nascosti e ad estrapolare caratteristiche salienti a partire dai dati di ingresso.

La capacità di costruire reti neurali multistrato ha consentito di implementare sistemi di *deep learning*, che costituiscono una evoluzione del *machine learning*, in grado di modellare astrazioni di alto livello sui dati, e oggi impiegati, ad esempio, nella *computer vision*. Lo sviluppo dei "sistemi di elaborazione parallela", delle "reti convoluzionali" e le "tecniche di correzione degli errori" (*backpropagation*) hanno consentito di ottenere importanti risultati circa il riconoscimento delle immagini da parte delle macchine, che allo stato attuale hanno raggiunto una accuratezza superiore a quella dei soggetti umani ed un margine di errore inferiore al 5%[13].

Ai fini di quanto verrà affrontato nella seconda parte del presente scritto è importante sottolineare, tuttavia, come i sistemi di *machine learning* e *deep learning*, a differenza dei sistemi basati sull'approccio simbolista, costituiscono delle scatole nere da cui ne consegue una incapacità di rendere conto dell'elaborazione effettuata; in altri termini, non vi è possibilità di capire perché tali macchine abbiano dato un risultato specifico, in quanto non è possibile descrivere la localizzazione della conoscenza che è distribuita sull'intera rete neurale.

---

[12] T. M. MITCHELL, *Machine Learning*, Milano, 1997.

[13] S. PARRA, M. TORRENS, *Intelligenza Artificiale. La strada verso la superintelligenza*, Milano, 2017.

Al termine di questa introduzione circa l'Intelligenza Artificiale occorre domandarsi quale sia effettivamente l'efficienza e l'efficacia degli algoritmi soprattutto in relazione alle abilità dei professionisti/esperti.

Nell'ambito della psicologia delle decisioni, in particolare, e delle scienze cognitive, in generale, il tema è stato ed è ampiamente studiato. Nel 1954 Paul Meehl pubblicava "*Clinical versus Statistical Prediction: A Theoretical Analysis and Review of the Evidence*", un piccolo volume contenente una revisione di circa venti studi che mettevano a confronto il giudizio di esperti di specifici settori verso previsioni statistiche elaborate combinando alcuni punteggi secondo una regola[14]. Come sottolineato da Daniel Kahneman, i venti studi presi in considerazione da Meehl erano destinati a diventare duecento, ma il punteggio nella gara tra algoritmi ed essere umani non cambiò: «*Circa il 60 per cento degli studi ha dimostrato che gli algoritmi sono stati assai più esatti. Dagli altri confronti risulta un pareggio per quanto riguarda l'accuratezza, ma un pareggio equivale a una vittoria per le regole statistiche. Non è stata documentata in maniera convincente nessuna eccezione*»[15].

Nel commentare tali risultati, a distanza di tempo dalla pubblicazione di "*Clinical versus Statistical Prediction: A Theoretical Analysis and Review of the Evidence*", lo stesso Meehl scrive: «*Nel campo delle scienze sociali non c'è nessuna controversia che, come questa, conti un così ricco* corpus *di studi qualitativamente vari che indichino in maniera tanto uniforme in un'unica direzione*»[16].

Le principali cause del perché gli esperti sono, in alcuni compiti specifici, meno affidabili degli algoritmi sono illustrate, oltre che da Meehl, da Simon, Kahneman, Amos Tsversky e Richard Thaler (solo per fare alcuni autorevoli nomi nel campo delle scienze cognitive).

Per Simon, ad esempio, la complessità crescente costituisce un eccessivo carico cognitivo per l'essere umano, motivo per cui, evoluzionisticamente, si è sviluppato il ragionamento euristico: la complessità riduce la validità. Gli esperti mostrano incoerenza nella formulazione di giudizi sommari su informazioni complesse[17].

Tale incoerenza è anche determinata dal condizionamento che il sistema euristico subisce dal contesto in cui l'individuo si trova a decidere e da alcuni effetti, riconducibili a *biases* ed euristiche, dei quali l'individuo è spesso inconsciamente soggetto.

Altro fattore estremamente importante, che incide sui giudizi degli esperti, è costituito da contesti incerti e poco stabili (tali contesti sono definiti "a bassa validità"): «*Quando la predittività è scarsa… l'incoerenza distrugge qualsiasi validità predittiva*»[18].

Per Kahneman, studi sperimentali alla mano, la competenza degli specialisti matura laddove siano rispettate due condizioni: a) un ambiente piuttosto regolare da essere prevedibile; b) l'opportunità di imparare queste regolarità attraverso una pratica prolungata.

Come scrive Simon: «*La situazione ha fornito un indizio, questo indizio ha dato all'esperto accesso a informazioni immagazzinate nella memoria e le informazioni forniscono la risposta. L'intuizione non è né più né meno che riconoscimento*»[19].

---

[14] W. M. GROVE, *Clinical Versus Statistical Prediction: The Contribution of Paul E. Meehl, in Journal of Clinical Psychology*, vol. 61, n. 10, 2005, p. 1233-1243.

[15] KAHNEMAN, *Pensieri lenti e veloci*, Milano, 2014, p. 246.

[16] MEEHL, 1986, in KAHNEMAN, *op. cit.*, p. 246.

[17] KAHNEMAN, *op. cit.*.

[18] KAHNEMAN, *op. cit.*, p. 249.

[19] SIMON, 2002, in KAHNEMAN, *op. cit.*, p. 261.

Gli algoritmi statistici, per contro, superano di parecchio gli esseri umani negli ambienti rumorosi per due motivi: a) hanno più probabilità dell'uomo di individuare indizi deboli, ma validi; b) conservano un modesto livello di accuratezza usando gli indizi in maniera coerente.

Da qui la conclusione di Kahneman per cui «*per massimizzare l'accuratezza predittiva, le decisioni finali dovrebbero essere affidate alle formule specie negli ambienti a bassa validità predittiva*»[20]. Lo psicologo israeliano, Nobel per l'economia nel 2002 «*per avere integrato risultati della ricerca psicologica nella scienza economica, specialmente in merito al giudizio umano e alla teoria delle decisioni in condizioni d'incertezza*», riconosce che la causa di un errore conta parecchio: «*Se un bambino muore perché un algoritmo ha commesso un errore è più terribile che se muore a causa di un errore umano e la differenza di intensità emozionale è prontamente tradotta in preferenza morale*»[21], tuttavia, conclude Kaheman «*è immorale affidarsi a giudizi intuitivi per decisioni importanti quando è disponibile un algoritmo che commette meno errori*»[22].

In conclusione, l'ipotesi debole dell'Intelligenza Artificiale è, in una certa misura, già una realtà impiegata in diversi campi e destinata ad essere sempre più adottata nel campo sanitario. I professionisti sanitari saranno affiancati da sistemi esperti e reti neurali che li aiuteranno nell'attività quotidiana nell'elaborazione dei dati, nel riconoscimento di patologie, nella formulazione di diagnosi, nella presa di decisione. Ciò richiederà, da una parte, la formazione di nuovi professionisti sanitari che abbiano sufficienti competenze nella gestione dell'Intelligenza Artificiale, dall'altra l'introduzione di nuove figure provenienti dal campo delle scienze cognitive in grado di fornire il necessario supporto ai decisori. Ciò comporterà, inoltre, una riflessione etica e giurisprudenziale circa l'uso di questi nuovi strumenti, così come già avvenuto tutte le volte che la tecnologia ha in qualche modo inciso sul naturale corso degli eventi.

Nella restante parte del presente documento si intende inquadrare il tema giurisprudenziale alla luce dell'attuale *corpus* normativo nazionale ed internazionale.

2. L'Intelligenza Artificiale nel mondo. – Nel corso degli ultimi anni, si è assistito a un sempre maggiore interesse per lo sviluppo delle Intelligenze Artificiali e per la loro regolamentazione sia da parte di tutti i Paesi industrializzati, sia da parte delle economie emergenti.

In tale prospettiva, è possibile notare come, da un punto di vista internazionale e comparato, i temi principali su cui si è focalizzata l'attenzione dei legislatori siano principalmente due. Da un lato, il loro inquadramento nell'ambito della competitività industriale che, di conseguenza, viene spesso trattata da un punto di vista strategico da parte dei Governi. Dall'altro, il tema del conflitto esistente tra molte attività la cui gestione viene attribuita alle Intelligenze Artificiali e i principi etici fondamentali generalmente riconosciuti: tema, quest'ultimo, che assume particolare rilevanza soprattutto nella tutela degli interessi della società civile.

Nonostante in molti Stati sia presente questa particolare attenzione alla ricerca di un equilibrio socialmente sostenibile tra gli interessi legati alla competizione e quelli inerenti l'eticità, occorre evidenziare che non è così distinguibile in tutti gli ordinamenti: per fare un esempio, l'Europa ha adottato una strategia che è sicuramente finalizzata a tentare di ottenere il primato in materia rispetto agli altri Paesi concorrenti, ma questa finalità risulta controbilanciata da una rigorosa attenzione rivolta al rispetto dei diritti e dei principi etici necessariamente coinvolti.

Questa tensione comporta investimenti importanti e la competizione si gioca, di conseguenza, anche sulla base della capacità di ciascun Stato di impiegare risorse economiche – e non solo – per lo

[20] KAHNEMAN, *op. cit.*, p. 249.
[21] KAHNEMAN, *op. cit.*, p. 253.
[22] KAHNEMAN, *op. cit.*, p.253.

sviluppo del mercato. In particolare, gli Stati più esposti sotto questo profilo sono la Cina e gli Stati Uniti che, di fatto, si stanno contendendo il dominio nel settore. Anche in tal caso, le cifre investite sono particolarmente ingenti: per esempio nella strategia elaborata dalla Cina nel 2015 ("*Made in China 2030*"), l'obiettivo è la costruzione di un settore completamente concentrato sulle Intelligenze Artificiali per un valore di circa 150 miliardi di dollari.

Ci sono poi altre risorse che risultano necessarie per riuscire a rendersi competitivi sul mercato: si pensi alla necessità di possedere complesse strutture che siano in grado di accedere a dati rilevanti e che riescano a lavorare su essi per la ricerca e l'elaborazione; o agli investimenti da introdurre nel sistema educativo e della formazione per favorire la promozione di livelli di competenza sempre più elevati ed adeguati allo sviluppo del settore e alla sua innovazione; o, ancora, alla creazione di un sistema amministrativo che possa adeguarsi rapidamente alla modernizzazione e offrire ai cittadini servizi sempre all'avanguardia.

Anche altri Paesi industrializzati e alcune economie emergenti hanno elaborato dei piani strategici per lo sviluppo della tecnologia e le modalità di competizione sul mercato. Ad esempio, in Canada, il *Canadian Institute for Advanced Research* (CIFAR), in collaborazione con le tre più importanti istituzioni che si occupano di Intelligenza Artificiale (*Amii*, *Mila*, e *Vector Institute*) ha investito 125 milioni di dollari per la creazione di una strategia nazionale per lo sviluppo della tecnologia.

Anche la Corea del Sud ha elaborato il "*Mid-To-Long-Term Plan in Preparation for The Intelligent Information Society*" con l'obiettivo di coinvolgere, nello sviluppo della tecnologia, sia la Pubblica Amministrazione, sia la società privata: infatti, il primo punto della strategia fa proprio riferimento ad una «*public-private partnership*», che vede al primo posto il lavoro delle aziende e dei cittadini in generale – veri protagonisti di questa innovazione – con il supporto del Governo e della comunità scientifica.

Il Giappone ha avviato una prima strategia nel 2017, il "*Artificial Intelligence Technology Strategy*" e, successivamente, la "*Society 5.0 Initiative*". La strategia elaborata nel 2017 è costituita da tre fasi che dovrebbero svilupparsi tra il 2020 e il 2030: la prima fase riguarda l'uso e l'applicazione delle Intelligenze Artificiali che utilizzano dati in diversi ambiti; la seconda, invece, cerca di estendere l'applicazione della tecnologia a settori di pubblico dominio; l'ultima fase, infine, è costruita pensando a un sistema interconnesso che consenta di collegare i settori in cui viene applicata la tecnologia. Nel presentare il documento, il Consiglio ha posto particolare attenzione ai settori della sanità, delle infrastrutture, dei trasporti, dell'educazione, promuovendo anche la costituzione di *start-up*. L'obiettivo della successiva strategia "*Society 5.0 Initiative*" è di cercare di superare alcuni problemi di carattere sociale (invecchiamento della popolazione; difficoltà dal punto di vista ambientale, ecc.) attraverso i vantaggi che le nuove tecnologie possono offrire. Attraverso questi provvedimenti, il Giappone intende rendersi ancora più competitivo, trasformando la società attuale e rendendola più vivace dal punto di vista economico, grazie alla creazione di nuovi mercati e al miglioramento della produttività.

L'India ha elaborato un piano strategico nel 2018, in cui si prendono in particolare considerazione le aree interessate per l'implementazione della tecnologia di Intelligenza Artificiale, tra cui, sanità, agricoltura, educazione, trasporti e infrastrutture – per quest'ultime si prevede la trasformazione delle città in *smart cities*.

In Russia, nel 2017 è stata proposta una specifica modifica del Codice Civile della Federazione Russa che – indipendentemente dall'autonomia riconosciuta ed attribuita al *robot* – dispone che la responsabilità venga sempre ricondotta sullo sviluppatore, produttore o addestratore della macchina: in questo modo si è fornita soluzione anche alle questioni riguardanti la rappresentanza processuale del *robot* dinanzi al giudice e il ruolo delle agenzie di sorveglianza. Inoltre, a questa innovazione si aggiunge anche una "*Model Convention on Robotics and AI*" che introduce nuove disposizioni in merito alla creazione e all'impiego delle Intelligenze Artificiali.

Il Presidente della Federazione Putin ha, altresì, annunciato di voler creare una rete Internet autonoma ed indipendente, per fare in modo di poter aumentare la sicurezza del *cyber*spazio russo ed evitare che gli Stati Uniti possano raccogliere dati provenienti dalle aziende e, più in generale, da residenti sul territorio. L'annuncio è avvenuto durante il *Security Council* del Cremlino e, da allora, sono stati presi dei provvedimenti finalizzati a consentire la realizzazione delle tre priorità nazionali enunciate dal Presidente. In primo luogo, aumentare il livello di protezione delle reti di comunicazione – in particolare, degli organi governativi – in modo tale da proteggere le informazioni della Russia da interferenze indesiderate. In secondo luogo, lavorare per garantire la stabilità e la sicurezza del *Web* russo. Infine, investire risorse idonee allo sviluppo di tecnologie e *software* di origine russa, per rendersi più competitivi a livello globale. Sotto questo ultimo profilo, si evidenzia la particolare arretratezza del Paese che fa quasi esclusivamente affidamento su tecnologie estere, con la conseguenza di dover sottostare – secondo quanto dichiarato Putin – agli interessi geopolitici dei Paesi stranieri e di trovarsi di fronte all'impossibilità di contrastare attacchi effettuati mediante quelle stesse tecnologie. Infatti, Putin ha predisposto la creazione di una strategia nazionale che determini un allineamento degli sviluppi militari, governativi, accademici e di tutti gli altri attori privati, per accelerare l'ammodernamento dello Stato e allinearsi con le altre due superpotenze al momento riconosciute: Cina e Stati Uniti.

Anche all'interno dell'Unione europea sono state presentate molte iniziative nazionali in Paesi come la Germania, il Regno Unito, il Portogallo, la Francia: quest'ultima si è trovata in una situazione peculiare rispetto agli altri Stati, proprio grazie al lavoro della "*Mission Villani*" che ha posto particolare attenzione su settori di politica industriale come la sanità, l'istruzione, l'ambiente, l'agricoltura, la sicurezza, la difesa, i trasporti, l'occupazione, e così via, facendo attenzione alla sostenibilità dello sviluppo del Paese. La "*Mission Villani*", infatti, evidenzia anche la necessità di principi etici che assicurino il rispetto dei diritti umani in ogni fase del progresso tecnologico perseguito. Pertanto, questa "*Mission*" si presenta come proposta per garantire che lo sviluppo e la diffusione delle Intelligenze Artificiali possano consentire una reale riduzione delle diseguaglianze economiche e sociali.

Tuttavia, questa differenza rispetto agli altri Paesi si è assottigliata, dal momento che, nel 2018, il Governo canadese – in collaborazione con quello francese – ha sviluppato un piano intergovernativo per l'Intelligenza Artificiale ponendo particolare attenzione al rispetto dei diritti umani, dell'innovazione, della diversità e della crescita economica.

Questo moltiplicarsi di strategie nazionali ha, di conseguenza, attratto l'attenzione delle istituzioni europee, che hanno cercato di fare in modo che questi piani convergessero verso una direttrice comune.

Gli atti dell'Unione Europea che interessano l'Intelligenza Artificiale sono molteplici, tuttavia uno dei più importanti per l'uniformazione dei progetti nazionali è il "*Digitising European Industry*", un Piano Coordinato presentato nel 2018. Nella Comunicazione che presenta il Piano, la Commissione europea propone un'Intelligenza Artificiale "*made in Europe*", che sia etica, sicura e all'avanguardia e che si concentri anche sugli aspetti scientifici e industriali, punto di forza dell'Europa.

La strategia presentata dalla Commissione si articola su quattro punti fondamentali. In primo luogo, l'aumento delle risorse disponibili, attraverso l'espansione degli investimenti pubblici e privati nel campo delle Intelligenze Artificiali. In secondo luogo, la costruzione di un quadro etico e giuridico adeguati agli sviluppi della tecnologia e che permettano agli ordinamenti di essere pronti ai cambiamenti socioeconomici che si verificheranno. Ancora, l'auspicio della Commissione è la creazione di un coordinamento a livello di Stati membri dell'Unione, che renderà sicuramente efficace la strategia proposta. Infine, la creazione di una Intelligenza Artificiale "*trustworthy*".

2.1. In particolare: in Italia. - Quanto alla strategia adottata in Italia in materia di Intelligenze Artificiali, l'ultimo Rapporto dell'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE) sulla "*Digital Transformation*" ha evidenziato quale sia l'approccio che caratterizza l'Italia in materia di Intelligenza Artificiale.

Da un lato, l'Italia si pone al quinto posto tra le economie dei Paesi G20, per il maggior numero di documenti scientifici pubblicati sui sistemi di *machine learning* – dopo gli Stati Uniti, la Cina, l'India e la Gran Bretagna – e risulta essere una tra i *leader* in Europa per la manifattura dei *robot* industriali (prima dell'Italia ci sono la Germania, la Repubblica Ceca, la Repubblica Slovacca e la Slovenia): per dare un termine comparativo, quanto prodotto in Italia sui *robot* industriali corrisponde a un terzo del numero prodotto dalla Corea del Sud (*leader* mondiale nel 2017).

Dall'altro lato, pur in considerazione dei dati su indicati, in Italia gli investimenti governativi in materia sono stati radicalmente ridotti nel corso degli ultimi anni, impedendole di porsi tra i Paesi più competitivi in termini di innovazione e impiego della tecnologia.

Nel quadro della strategia "*Europa 2020*" approvata nel 2010 dalla Commissione europea – che rientra nell'ambito dell'azione dell'Unione per il raggiungimento di una maggiore competitività su scala mondiale – una delle iniziative proposte è la "*Agenda europea per il Digitale*" che consente agli Stati membri di individuare strategie volte al raggiungimento degli obiettivi comunitari. Infatti, nella presentazione di questo piano decennale di trasformazione dell'Europa, viene affermato: «*L'Europa sta vivendo una fase di trasformazione. La crisi ha vanificato anni di progressi economici e sociali e messo in luce le carenze strutturali dell'economia europea. Nel frattempo, il mondo si sta rapidamente trasformando e le sfide a lungo termine (globalizzazione, pressione sulle risorse, invecchiamento) si accentuano. L'UE deve prendere in mano il proprio futuro. Per ottenere buoni risultati l'Europa deve agire in modo collettivo, in quanto Unione. Abbiamo bisogno di una strategia che ci consenta di uscire più forti dalla crisi e di trasformare l'UE in un'economia intelligente, sostenibile e inclusiva caratterizzata da alti livelli di occupazione, produttività e coesione sociale. Europa 2020 dà un quadro dell'economia di mercato sociale europea per il XXI secolo*».

In realtà, l'innovazione tecnologica è stata più rapida di quanto ipotizzato nella iniziativa indicata e l'Italia è stata costretta a assumere provvedimenti urgenti per far fronte alle nuove sfide che le Intelligenze Artificiali, l'*Internet of Things* e il meccanismo della *Blockchain* ponevano. Per questo motivo, è stato elaborato un "*Piano triennale per l'informatica nella Pubblica Amministrazione*" che delinea un piano programmatico finalizzato allo sviluppo di quattro aree di interesse: ecosistemi digitali (come la sanità, l'istruzione, la giustizia, ecc.); le infrastrutture immateriali (come tutte le piattaforme abilitanti e i dati delle Pubbliche Amministrazioni); le infrastrutture fisiche e, infine, la *cybersecurity*.

Tuttavia, proprio per la difficoltà delle materie trattate e la necessità di competenze tecniche interdisciplinari che consentano di affrontare efficacemente le sfide che si presentano, è sta anche istituita una *Task Force* che ha pubblicato un "*Libro Bianco*" in cui si cerca di definire quali siano i rischi e le opportunità che si presentano con l'utilizzo delle Intelligenze Artificiali nel settore pubblico, muovendo dalla convinzione che la tecnologia, se utilizzata in maniera corretta, sarà in grado di rendere l'Italia molto più competitiva nel prossimo futuro.

Per quanto concerne la modalità di impiego della tecnologia, l'attenzione, pertanto, è rivolta sia alle imprese che si occupano di Intelligenza Artificiale, sia alla Pubblica Amministrazione. Il sistema di intervento pubblico delineato, per i casi in cui queste operazioni possono comportare dei rischi – in particolare, nei settori ritenuti più sensibili – è armonizzato dall'azione dell'Unione Europea.

Nel rispetto delle normative adottate a livello di Unione, pertanto, l'ordinamento italiano si è dovuto adattare, per esempio, alle regole disciplinanti le modalità di partecipazione dello Stato nelle

imprese. Infatti, per quanto concerne quelle che operano in settori strategici, l'intervento dello Stato si è trasformato dal c.d. "*goden share*", al c.d. "*golden power*".

In tal senso, il d.l. 2141 del 2012 ha sostituito il meccanismo del *golden share*, che attribuiva allo Stato una partecipazione azionaria dotata di poteri particolari, con quello del *golden power* con cui il Governo è in grado di esercitare maggiori poteri rispetto a quelli normalmente concessi agli azionisti, in occasione di operazioni specifiche che riguardino settori strategici. In tal senso, la disciplina nel citato d.l. individua gli ambiti di intervento consentiti e che riguardano, in particolare, la difesa e la sicurezza nazionale e i settori di energia, trasporti e telecomunicazioni. Nell'elenco non viene contemplato il settore della Intelligenza Artificiale, che al contrario, dovrebbe essere ritenuto altrettanto strategico, considerando le opportunità di sviluppo che comporta, ma anche i rischi esistenti nel suo utilizzo.

In realtà, il c.d. "*decreto fiscale*" – recante "*disposizioni urgenti in materia finanziaria e per esigenze indifferibili*" – del 2017 prevedeva un'estensione del *golden power* anche ai settori denominati "*ad alta intensità tecnologica*". Infatti, dopo una serie di vicissitudini normative, l'art. 4 del d.l. 105 del 2019 ha esteso l'ambito operativo delle norme in tema di poteri speciali esercitabili dal Governo, comprendendo, appunto, anche questi settori.

Il problema principale che deve essere affrontato dal legislatore in modo chiaro ed espresso è l'individuazione di una modalità per consentire allo Stato italiano di essere al passo con l'incessante sviluppo tecnologico, stanziando le ingenti risorse economiche necessarie allo scopo. Questa sfida deve essere affrontata anche dal un punto di vista della tecnica di produzione del diritto: in un ambiente in costante trasformazione, sono necessari degli interventi finalizzati a rendere più flessibile la disciplina, che consenta la sperimentazione di forme di economia e una costante innovazione del sistema, volto a una sempre maggiore sostenibilità.

Da un punto di vista pratico, l'ordinamento italiano si è concentrato in modo particolare sull'aspetto pubblico, cercando di innovare l'azione – sempre più trasparente – della Pubblica Amministrazione e i servizi offerti al cittadino. Si pensi, per esempio, ad una delle novità introdotte dal d.l. 179/2012, in merito agli *Open Data*. L'idea è che i dati nel settore pubblico appartengano a tutti i cittadini e di conseguenza, in quanto patrimonio dello Stato, debbano essere resi accessibili, nel rispetto della normativa GDPR. Inoltre, la volontà di condivisione dei dati è rappresentata anche dalla creazione della Piattaforma Digitale Nazionale Dati (prevista nel Codice di Amministrazione Digitale: art. 50-*ter* del d.lgs. 82 del 2005) in cui le Pubbliche Amministrazioni dovranno pubblicare i dati in loro possesso, in collaborazione con l'Autorità Garante per la protezione dei dati personali, al quale viene conferito il compito di controllo sull'attività svolta.

Nonostante i progetti e le iniziative volte a rendere il settore pubblico sempre più digitale, sono ancora necessari molti interventi per fare in modo che questi strumenti possano davvero essere efficaci. Inoltre, si avverte l'esigenza della creazione di una maggiore consapevolezza del valore dei dati, sia per poterlo sfruttare sul mercato, sia per evitarne l'uso indiscriminato da parte di altri soggetti operanti nel settore delle Intelligenze Artificiali. Infine, risulta completamente assente – esattamente come nel resto d'Europa – una disciplina che regoli in maniera specifica e dettagliata la responsabilità in caso di danno prodotto dall'algoritmo. Pertanto, l'unica soluzione normativa che, per il momento, si prospetta, nell'ambito della sicurezza dei prodotti all'interno del mercato europeo, è quella di applicare la Direttiva Macchine e di ricondurre alla Direttiva per i diritti e le garanzie riconosciute ai consumatori tutti i casi di vendita di prodotti destinati al consumo.

3. Qualificazione giuridica dell'Intelligenza Artificiale. – Anche quando si desidera affrontare un tema dal punto di vista giuridico, la prima attività richiesta all'interprete è, soprattutto, quella di individuare una definizione del fenomeno che si intende studiare che sia comprensiva delle

diverse realtà che lo compongono: in altre parole, si richiede allo studioso di individuare i singoli elementi più semplici che compongono e costituiscono il fatto complesso che si vuole definire per poi procedere ad una loro ricomposizione secondo categorie note che consentano di comprenderne il significato. In questo modo, attraverso la definizione di un fenomeno, si è in grado di spiegare che cosa esso sia attraverso l'indicazione di quali siano le sue caratteristiche peculiari e la delimitazione dell'ambito in cui esso opera. La stessa attività, pertanto, richiede la semplificazione del fenomeno e la riconduzione di questo a categorie precostituite: solo a questo punto sarà possibile applicare il resto della disciplina.

Tuttavia, si è già avuto modo di evidenziare le difficoltà esistenti nell'individuazione di una definizione complessiva di cosa si intenda per "Intelligenza Artificiale": difficoltà che si ripercuotono anche sul giurista che si trova nella difficoltà di individuare quel catalogo di caratteristiche precise che gli consentono di specificare e, conseguentemente, di applicare, la disciplina corrispondente.

Tale difficoltà è, principalmente, determinata da due fattori. Innanzitutto, dalla particolare interdisciplinarietà che caratterizza le Intelligenze Artificiali che rendono impossibile l'individuazione di una definizione che sia unica e che sia idonea a rappresentarle in ogni ambito di applicazione: giuridico e non. Inoltre, l'incredibile rapidità di innovazione e di evoluzione che caratterizza il settore.

Acclarata l'impossibilità di costruire una definizione unitaria, diviene indispensabile possedere una conoscenza sufficientemente approfondita di ciascun tipo di Intelligenza Artificiale, in modo da poterne individuare le caratteristiche di funzionamento che consentano di qualificarla e, possibilmente, di ricondurla ad una determinata categoria giuridica.

Infatti, i sistemi di Intelligenza Artificiale sono molto diversi tra loro.

Da un lato, vi sono macchine e *robot* in grado di svolgere mansioni specifiche su indicazioni – e sotto il serrato controllo – dell'uomo.

Dall'altro, vi sono Intelligenze Artificiali che presentano una particolare abilità ad imparare in modo dinamico, interattivo e completamente autonomo, dall'ambiente in cui sono inseriti, anche quando le condizioni sono in continuo mutamento: in modo particolare, quelli di *machine learning* e *deep learning*. Ciò comporta che, proprio per la loro dinamicità, queste tecnologie sono in grado di porre in essere comportamenti con modalità imprevedibili nel momento di elaborazione e di sviluppo del programma iniziale. Potrebbe essere, pertanto, che questi sistemi siano caratterizzati da una sorta di "opacità" nel proprio funzionamento e di sostanziale "imprevedibilità" nel comportamento: due elementi che possono avere delle conseguenze importanti sotto il profilo giuridico nel caso in cui la loro condotta sia tale da provocare un danno. In questi casi, il problema che si pone riguarda non solo l'individuazione di chi sia il soggetto che debba rispondere del danno arrecato ma, come si avrà modo di esporre, si creano delle perplessità addirittura, nell'inquadramento (sotto il profilo giuridico) del tipo di responsabilità a cui fare riferimento[23].

[23] Queste sono le forme di Intelligenza Artificiale su cui maggiormente si concentrano gli investimenti delle aziende e che, pertanto, destano il maggiore interesse anche sotto il profilo giuridico. Si possono fare tre esempi.

In primo luogo, a differenza delle forme di Intelligenza Artificiale direttamente collegate con l'azione dell'uomo, questi sistemi di *machine learning* e di *deep learning* utilizzano dei modelli più complessi che possono rendere molto più difficile tracciare i singoli passaggi – e la logica ad essi sottostante – per stabilire il motivo per cui da una data premessa si è raggiunto un determinato risultato. Alcune forme di Intelligenza Artificiale rendono palese il processo logico che consente a chi le utilizza di essere tracciate, in quanto utilizzano un sistema chiamato "albero decisionale": in questo modo, è più semplice stabilire dove si sia verificato l'errore, in quanto lo stesso algoritmo mostra i singoli passaggi che sono stati effettuati per il raggiungimento di quell'esito specifico. Altri modelli, quali, ad esempio, i sistemi di *deep learning* che

Altro aspetto che assume particolare rilevanza nella presente analisi, riguarda l'*input* umano e la sua interazione con le Intelligenze Artificiali. Infatti, nonostante si possa riconoscere che i continui successi nell'innovazione e lo sviluppo di nuove tecnologie sia in larga parte da attribuire alle tecnologie stesse – che consentono di rendere più veloci i processi di calcolo e, spesso, sono in grado di evolvere da sole – non bisogna dimenticare che l'intervento umano è presente in ogni passaggio dello sviluppo di tecnologie guidate da sistemi di Intelligenza Artificiale. Si pensi al momento dell'ideazione di una nuova macchina; alla elaborazione della proposta, alle scelte relative al *design* da applicare, alle analisi di settore, ai *test* funzionali e così via. Inoltre, ci si aspetta che questi stessi sistemi siano in grado di operare nel mondo reale in collaborazione con gli esseri umani, spesso su larga scala.

In particolare, frequentemente, l'Intelligenza Artificiale viene programmata in modo da fornire all'essere umano una semplice raccomandazione, in modo tale da lasciare la discrezionalità della decisione finale da adottare soltanto all'uomo. Tuttavia, si può discutere sull'effettiva libertà di decisione lasciata all'individuo quando è assistito da un algoritmo, soprattutto in ambito sanitario. Quando l'uomo si trova di fronte a una macchina in grado di individuare statisticamente le probabilità di un determinato risultato, egli si sente strettamente vincolato dai numeri che gli vengono proposti, a prima vista, in modo assolutamente neutro. In realtà, i dati raccolti e le modalità di elaborazione delle statistiche stesse possono portare a dei risultati talvolta distorti, la cui devianza è difficilmente riconoscibile all'occhio umano[24].

In questo quadro così eterogeneo e così condizionante per l'uomo, diviene indispensabile comprendere se le norme oggi vigenti siano idonee a disciplinare il fenomeno e se sono in grado di adeguarsi alle nuove sfide che questa realtà concreta presenta e presenterà.

Infatti, i giuristi si chiedono se sia possibile utilizzare per le Intelligenze Artificiali le categorie giuridiche già esistenti, adattandole ai problemi specifici posti dalle nuove tecnologie, oppure se, in alternativa, sia più opportuno orientarsi verso l'individuazione di nuove regole, appositamente emanate per la regolamentazione di questo particolare settore. Considerando che la scelta fra i due

---

utilizzano uno schema di rete neurale, non consentono altrettanta trasparenza nell'analisi del loro comportamento.

In secondo luogo, anche quei sistemi che utilizzano un tipo di processo logico che consente di poterli tracciare potrebbero risultare, in realtà, imperscrutabili, in quanto protetti dal diritto di proprietà intellettuale, che attribuisce al titolare il diritto di mantenere il segreto sulla formula algoritmica utilizzata.

Da ultimo, residua il problema di definire con esattezza la differenza esistente fra il ricevere un'informazione e la reale capacità di comprenderla. Infatti, chiunque non sia un esperto in materia non otterrà alcuna utilità dal ricevere informazioni sui metodi utilizzati per l'addestramento dell'algoritmo posto a fondamento di un sistema. Di conseguenza, la trasparenza si riduce a mera formalità, e risulta di fatto inutile la possibilità di rendere (formalmente) pubbliche specifiche informazioni molto tecniche che riguardano il funzionamento degli algoritmi impiegati in determinati settori. La combinazione della imperscrutabilità e dell'opacità degli algoritmi ha portato molti studiosi a qualificare le Intelligenze Artificiali come "scatole nere" (c.d. *Black Boxes*): una qualificazione che, comunque, viene fortemente discussa.

[24] Semplificando al massimo, la macchina offre il proprio risultato attraverso un procedimento sillogistico: data la situazione "X" concreta (tesi), la si compara con la regola generale "Y" derivante dalla quantità di dati elaborati dalla macchina (antitesi), viene offerta la soluzione da adottare "Z" (sintesi). Facciamo un esempio. Se noi esponiamo il seguente ragionamento: Paolo e Pietro suono uomini (tesi), e gli uomini sono mortali (antitesi), è corretto sostenere che Paolo e Pietro sono mortali (sintesi). In realtà, se per una qualche ragione viene sbagliata la qualificazione della tesi "X" o se c'è un errore nella definizione della antitesi "Y" il risultato finale potrebbe essere sicuramente logico ma aberrante, ad esempio, sostenendo: Paolo e Pietro sono apostoli (tesi); gli apostoli sono dodici (antitesi); Paolo e Pietro sono dodici (antitesi); come detto un risultato aberrante.

orientamenti dipenderà dall'esigenza di disciplinare nel miglior modo possibile la materia, in realtà, occorre evidenziare che entrambe le soluzioni prospettate presentano aspetti molto delicati che necessiteranno di particolare attenzione e di approfondimento anche tenendo presente che la soluzione che verrà adottata dovrà essere giustificata (ed anche legittimata), all'interno dell'ordinamento giuridico e dei principi fondamentali che lo regolamentano.

Considerando la loro natura e il vasto impiego in qualunque settore, l'individuazione di una adeguata disciplina giuridica di riferimento è un problema rilevante, in quanto le Intelligenze Artificiali sono in grado di provocare dei danni, anche notevoli.

Come si avrà modo di rilevare nel prosieguo della trattazione, per le forme di Intelligenza Artificiali più semplici – in cui è evidente il condizionamento (e, pertanto, l'errore) derivante dal comportamento dell'uomo – la dottrina ha fornito diversi tipi di soluzioni. Diverso è il caso delle forme più evolute di Intelligenza Artificiale che non rispondono alla normale logica giuridica. Per poter attribuire ad un determinato soggetto la responsabilità del danno occorre, preliminarmente, dimostrare l'esistenza di un collegamento fra la condotta (attiva o omissiva) tenuta dal soggetto e il danno concreto che si è verificato: quello che i giuristi definiscono provare l'esistenza del c.d. nesso di causalità. Come detto, le Intelligenze Artificiali più evolute e sofisticate sono dotate di sistemi di auto-apprendimento, che rende la loro condotta molto simile a quella che potrebbe essere tenuta da un essere umano. Il che rende molto complicato sotto il profilo giuridico definire l'esistenza del "nesso di causalità": in questi casi, infatti, nemmeno chi ha costruito o addestrato la macchina è in grado di ricostruire esattamente l'esistenza di una correlazione fra l'attività posta in essere dall'uomo (produttore e/o addestratore) e il danno che si è verificato.

Si tenga presente che anche quando gli algoritmi sono "più semplici", la ricostruzione del nesso causale è, comunque, sempre molto complicata, in quanto risulta quasi impossibile ripercorrere tutti i passaggi che hanno comportato la produzione di quell'evento, che si è risolto in un danno: tema talmente rilevante che merita uno specifico approfondimento nella presente trattazione[25].

Ciò posto, non si tratta di un settore produttivo abbandonato a sé stesso. Per fare in modo di ricondurre a principi etici condivisi le condotte dei programmatori, sviluppatori e addestratori di algoritmi, nel corso degli anni sono state adottate molte linee guida da parte di Organizzazioni Internazionali finalizzate al perseguimento delle medesime finalità: la creazione di Intelligenze Artificiali "*trustworthy*" (affidabili), definire gli *standard* di qualità dei dati forniti durante l'addestramento, la sostenibilità del loro sviluppo ed inserimento nella realtà sociale, e così via. Un primo esempio, già ricordato, è la Risoluzione del Parlamento europeo del Febbraio 2017, ma, più di recente, anche i Paesi del G20 hanno adottato delle linee guida che sono dirette a orientare la condotta di coloro che operano nel settore.

Tuttavia, nonostante se ne riconosca il valore, si tratta sempre di indicazioni estremamente generali la cui applicabilità è, prevalentemente, limitata alla sola fase di programmazione, sviluppo e addestramento degli algoritmi che animano le Intelligenze Artificiali. La *ratio* ispiratrice è evidente: per ottenere dei "buoni" *robot* è necessario avere, in primo luogo, dei "buoni" sviluppatori e utilizzatori delle macchine stesse.

Ciò nonostante, pensare che sia sufficiente far riferimento solo alla presenza di "buone intenzioni" nell'uso dell'Intelligenza Artificiale e alla qualità impeccabile della programmazione, potrebbe non risultare sufficiente.

Se si vuole affrontare il problema senza cadere in facili semplificazioni, occorre accettare il fatto che la tecnologia non è, in sé, neutrale. Ci possono essere numerose occasioni in cui una macchina pone in essere una condotta cui consegue un evento dannoso, nonostante siano state prese tutte le precauzioni del caso e rispettate tutte le possibili linee guida esistenti. Per questo motivo, si

[25] Si veda il successivo Par. 3.

ritiene assolutamente necessario trovare una soluzione adeguata alla definizione dei profili di responsabilità.

3.1. Alcune considerazioni di carattere preliminare alla qualificazione giuridica e alla ricostruzione del nesso di causalità. – Come si è avuto modo di esporre, è difficile (se non impossibile) fornire una soluzione unitaria che valga per tutti i tipi di Intelligenza Artificiale e nei prossimi paragrafi si indicheranno gli orientamenti più seguiti nel disciplinare la maggior parte delle macchine presenti sul mercato: in particolare, quelle che dipendono – direttamente o indirettamente – dalle scelte fatte da un uomo. Per le macchine più evolute (di *machine learning* e di *deep learning*), si darà menzione delle soluzioni a tutt'oggi fornite dalla dottrina, prevalentemente straniera, che si è occupata dell'argomento.

Più precisamente, i giuristi si interrogano, in primo luogo, sulla qualificazione giuridica delle Intelligenze Artificiali: una volta qualificate sarà più semplice definire quale disciplina giuridica sarà possibile applicare. Sul punto non vi è unanimità: infatti, alcuni ritengono applicabile per analogia alcune disposizioni specifiche presenti nel Codice civile; altri la riconducono alla disciplina della responsabilità del produttore; altri ancora propendono per riconoscere le Intelligenze Artificiali come "agenti" e propongono il loro riconoscimento come soggetti del diritto (come se fossero degli esseri umani). In ambito sanitario, infine, parrebbe applicabile la disciplina dei dispositivi medici. Tutte impostazioni che meritano uno specifico approfondimento.

La seconda questione che interessa dal punto di vista giuridico è, una volta qualificata l'Intelligenza Artificiale – vale a dire, ricondotto il fenomeno a una categoria giuridica – come ricostruire collegamento esistente fra azione compiuta e danno realizzato (il c.d. nesso causale) e, conseguentemente, poter individuare il responsabile del danno concreto.

In questo caso, come si vedrà, emergono questioni più o meno complesse a seconda che l'Intelligenza Artificiale utilizzata sia più o meno imperscrutabile: si ritorna al tema particolare delle *Black Boxes*.

3.2. L'Intelligenza Artificiale intesa come se fosse un bambino, un lavoratore dipendente, una attività pericolosa, una cosa o un animale. – Come già accennato, parte della dottrina ritiene opportuno qualificare giuridicamente le Intelligenze Artificiali come se fossero delle "cose", ritenendo applicabili alcune fattispecie previste nel Codice civile: in particolare, gli articoli dal 2048 al 2052. La ragione è evidente: al fine di semplificare giuridicamente la situazione e la posizione del consumatore (del paziente), applicando queste norme si prevede una responsabilità oggettiva in capo all'essere umano che ha in custodia e/o che utilizza la macchina[26].

[26] Secondo quanto disposto dall'art. 2048 del cod. civ. (intitolato: "*Responsabilità dei genitori, dei tutori, dei precettori e dei maestri d'arte*"): «*Il padre e la madre, o il tutore, sono responsabili del danno cagionato dal fatto illecito dei figli minori non emancipati o delle persone soggette alla tutela, che abitano con essi. La stessa disposizione si applica all'affiliante.*

I precettori e coloro che insegnano un mestiere o un'arte sono responsabili del danno cagionato dal fatto illecito dei loro allievi e apprendisti nel tempo in cui sono sotto la loro vigilanza.

Le persone indicate dai commi precedenti sono liberate dalla responsabilità soltanto se provano di non avere potuto impedire il fatto».

In questo caso, rilevano le caratteristiche specifiche dell'Intelligenza Artificiale, in quanto ciascuna delle diverse fattispecie presenti nel Codice civile richiedono la presenza di elementi precisi e specifici per poter trovare applicazione nel caso concreto.

Rileverà anche l'ambito di applicazione della macchina stessa. Se si prendesse, ad esempio, un *robot* tele-guidato utilizzato in ambito medico, questo sarà disciplinato in modo differente rispetto ad altri *robot*, sempre comandati a distanza, ma utilizzati, ad esempio, nelle operazioni spaziali. Tale diversità è data dal fatto che, quando possibile, trova applicazione la *lex specialis*: ogni macchina, pertanto, verrà regolamentata nell'ambito di operatività della legislazione che disciplina il loro settore specifico.

Pur potendo essere considerati "simili", i primi saranno sussunti nella disciplina dettata dalla legge 8 marzo 2017, n. 24 (nota anche come Legge "Gelli-Bianco" dal nome dei due relatori alla Camera e al Senato); i secondi, invece, essendo privi di una legislazione speciale, saranno ricondotti alla disciplina generale dell'art. 2050 cod. civ., che regola la responsabilità per l'esercizio di attività pericolose[27].

3.2.1 Le Intelligenze Artificiali utilizzate nell'ambito sanitario e la Legge 8 marzo 2017, n. 24. – Senza alcuna pretesa di completezza, alla fine degli anni '90, la responsabilità sanitaria è stata fortemente condizionata da alcune sentenze delle Sezioni Unite della Corte di Cassazione il cui orientamento, fondato sulla teoria del c.d. "contatto sociale", aveva portato a qualificare come contrattuale (*ex* art. 1218 cod. civ.) ogni forma di responsabilità in materia, comprendendovi non solo quella delle strutture sanitarie (pubbliche o private) ma anche quella dei professionisti sanitari che entrando in contatto con il paziente (il c.d. contatto sociale) venivano automaticamente assoggettati a garantire a quest'ultimo la salute [28].

Questa impostazione è stata successivamente messa in discussione con l'emanazione del D.L. 13/09/2012 (convertito in legge 8 novembre 2012, n. 189: nota come "Legge Balduzzi" dal nome

---

Il successivo art. 2049 (che disciplina la "*Responsabilità dei padroni e dei committenti*") statuisce che «*I padroni e i committenti sono responsabili per i danni arrecati dal fatto illecito dei loro domestici e commessi nell'esercizio delle incombenze a cui sono adibiti*».

Secondo l'art. 2050 ("*Responsabilità per l'esercizio di attività pericolose"*) «*Chiunque cagiona danno ad altri nello svolgimento di una attività pericolosa, per sua natura o per la natura dei mezzi adoperati, è tenuto al risarcimento, se non prova di avere adottato tutte le misure idonee a evitare il danno*».

Ancora, l'art. 2051 ("*Danno cagionato da cosa in custodia*") sostiene che «*Ciascuno è responsabile del danno cagionato dalle cose che ha in custodia, salvo che provi il caso fortuito*».

Infine, l'art. 2052 ("*Danno cagionato da animali*") prevede che «*Il proprietario di un animale o chi se ne serve per il tempo in cui lo ha in uso, è responsabile dei danni cagionati dall'animale, sia che fosse sotto la sua custodia, sia che fosse smarrito o fuggito, salvo che provi il caso fortuito*».

[27] La disciplina dettata dall'art. 2050 cod. civ. si presta, soprattutto, a regolamentare le ipotesi di responsabilità causata dai *robot* che rientrano nella prima macro-categoria che è stata elaborata da *Strategic Research Agenda for robotics in Europe*: i *robot* c.d. teleoperati, macchine utilizzate in settori molto rischiosi per l'uomo, quali, ad esempio, la bonifica delle mine o alcune operazioni spaziali.

[28] Si v. Cass. SS.UU., 11/10/2008, n. 577. L'art. 1218 cod. civ. che disciplina "*La responsabilità del debitore*", stabilisce che «*Il debitore che non esegue esattamente la prestazione dovuta è tenuto al risarcimento del danno, se non prova che l'inadempimento o il ritardo è stato determinato da impossibilità della prestazione derivante da causa a lui non imputabile*».

dell'allora Ministro della salute) in cui, nel disciplinare la responsabilità penale dei professionisti sanitari, all'art. 3 inserisce l'inciso «*In tali casi resta comunque fermo l'obbligo di cui all'art. 2043 cod. civ.*». Tralasciando ogni approfondimento sul punto, basti segnalare che, in conseguenza di tale inciso, alcuni Tribunali di merito hanno ritenuto che il legislatore del 2012 avesse introdotto il c.d. "doppio binario". Secondo questa impostazione, la responsabilità delle strutture sanitarie andrebbe qualificata sempre come responsabilità "contrattuale" (in quanto il paziente stipula con la struttura sanitaria un contratto – atipico – di spedalità); quella dei professionisti dipendenti delle strutture, invece, dovrebbe essere ricondotta nella responsabilità extracontrattuale ai sensi dell'art. 2043 cod. civ. secondo cui: «*Qualunque fatto doloso o colposo, che cagiona ad altri un danno ingiusto, obbliga colui che l'ha commesso a risarcire il danno*».

Il c.d. doppio binario è stato, normativamente, riconosciuto con l'emanazione dell'art. 7 della legge 24 del 17[29] (nota come Legge Gelli dal nome del deputato relatore), in cui si è definitivamente stabilito che, a fronte di un evento dannoso, il professionista sanitario sia responsabile *ex* art. 2043 cod. civ., mentre la struttura ospedaliera (pubblica o privata) continua a rispondere di responsabilità contrattuale *ex* art. 1218 cod. civ..

Quest'ultima impostazione assorbe e, conseguentemente, regola anche le ipotesi di danno provocato da un *robot* chirurgico: se il paziente subisce un danno in conseguenza dell'utilizzo di una macchina mossa da Intelligenza Artificiale, di tale danno risponderà, innanzitutto, a titolo contrattuale, la struttura sanitaria: in tal caso, il paziente avrà un onere della prova più "leggero" limitandosi a provare l'esistenza del danno e che questo dipende dalla prestazione fornita dalla struttura sanitaria e la relativa azione si prescriverà in 10 anni dal momento in cui emerge il danno.

Il paziente potrà agire anche contro il professionista sanitario che controllava la macchina per responsabilità extracontrattuale ma, in tal caso, il suo onere probatorio sarà aggravato dalla

---

[29] L'art. 7 della Legge 8 marzo 2017, n. 24, rubricato "*Responsabilità civile della struttura e dell'esercente la professione sanitaria*", stabilisce: «*1. La struttura sanitaria o sociosanitaria pubblica o privata che, nell'adempimento della propria obbligazione, si avvalga dell'opera di esercenti la professione sanitaria, anche se scelti dal paziente e ancorché' non dipendenti della struttura stessa, risponde, ai sensi degli articoli 1218 e 1228 cod. civ., delle loro condotte dolose o colpose.*

2. La disposizione di cui al comma 1 si applica anche alle prestazioni sanitarie svolte in regime di libera professione intramuraria ovvero nell'ambito di attività di sperimentazione e di ricerca clinica ovvero in regime di convenzione con il Servizio sanitario nazionale nonché attraverso la telemedicina.

*3. L'esercente la professione sanitaria di cui ai commi 1 e 2 risponde del proprio operato ai sensi dell'articolo 2043 cod. civ., salvo che abbia agito nell'adempimento di obbligazione contrattuale assunta con il paziente. Il giudice, nella determinazione del risarcimento del danno, tiene conto della condotta dell'esercente la professione sanitaria ai sensi dell'articolo 5 della presente legge e dell'articolo 590-*sexies *del codice penale, introdotto dall'articolo 6 della presente legge.*

4. Il danno conseguente all'attività della struttura sanitaria o sociosanitaria, pubblica o privata, e dell'esercente la professione sanitaria è risarcito sulla base delle tabelle di cui agli articoli 138 e 139 del codice delle assicurazioni private, di cui al decreto legislativo 7 settembre 2005, n. 209, integrate, ove necessario, con la procedura di cui al comma 1 del predetto articolo 138 e sulla base dei criteri di cui ai citati articoli, per tener conto delle fattispecie da esse non previste, afferenti alle attività di cui al presente articolo.

5. Le disposizioni del presente articolo costituiscono norme imperative ai sensi del cod. civ.».

Per ogni ulteriore approfondimento, per tutti, si v. AA.VV., *Responsabilità sanitaria*, a cura di S. ALEO, P. D'AGOSTINO, R. DE MATTEIS, G. VECCHIO, Milano, 2018.

dimostrazione della sussistenza della colpa del professionista e soggetto ad un termine di prescrizione della relativa azione più breve (5 anni dal momento in cui emerge il danno).

Anche in tal caso, sarà comunque necessario procedere ad una analisi precisa ed approfondita delle caratteristiche specifiche della macchina utilizzata: non è detto che un'Intelligenza Artificiale inserita in un *robot* chirurgico sia equiparabile a quella inserita in una Tomografia Computerizzata (TC) .

In ogni caso, sarà molto difficile individuare una soluzione unitaria che sia adeguata a tutte le ipotesi di applicazione della tecnologia di Intelligenza Artificiale.

3.3. L'Intelligenza Artificiale intesa come un prodotto. – Parte della dottrina ha ritenuto opportuno ricondurre anche tale materia nell'ambito delle disposizioni che regolano la responsabilità del produttore, in quanto proprio quest'ultimo rappresenta colui che – più di chiunque altro fra i vari soggetti coinvolti – è in grado di: individuare meglio quale sia effettivamente il rischio che i *robot* da lui prodotti possano provocare dei danni; adottare le misure necessarie per evitare che si verifichino (*risk management*); collocare assicurativamente il rischio dei danni che non è in grado di evitare, redistribuendo sul prezzo dei singoli prodotti venduti l'onere del premio assicurativo pagato per proteggersi dall'eventualità di un danno indebito a terzi[30].

In particolare, la Direttiva 85/374/CEE[31] – che disciplina la responsabilità per danno da prodotti difettosi, recepita nel nostro ordinamento con il DPR 224 del 1998, poi abrogato dal Codice del consumo (d.lgs. 205/2006) – potrebbe trovare applicazione anche nei casi in esame relativi al risarcimento dei danni provocati da difetti di fabbricazione di Intelligenze Artificiali.

La dottrina si è interrogata sulla applicabilità di tale disciplina anche alle tecnologie emergenti: in proposito, la stessa Commissione europea ha espresso la volontà di riesaminare la direttiva, creando dei gruppi di esperti al fine di valutare l'opportunità di applicazione della stessa a tale settore[32].

Secondo quanto stabilito dall'art. 2, primo comma, della direttiva, nel definire cosa si intenda per "prodotto", si è stabilito: «*prodotto, ai fini delle presenti disposizioni, è ogni bene mobile, anche se incorporato in altro bene mobile o immobile*». Questa definizione fa un chiaro riferimento alla materialità dei beni, mobili o immobili, che sono qualificabili come prodotti. Ma occorre evidenziare quanto disposto nel successivo terzo comma della norma in esame, in cui si dispone che «*per prodotto si intende anche l'elettricità*»: il legislatore europeo, quindi, ha specificamente incluso nella categoria di "prodotto" anche l'elettricità, che è un bene immateriale. Ne consegue che, pertanto, seguendo il metodo interpretativo secondo cui dove la legge lo ha voluto lo ha previsto e quando non

---

[30] Si v. A. CORDIANO, *Sub art. 115*, in E. CAPOBIANCO, L. MEZZASOMA, G. PERLINGIERI (a cura di), *Codice del consumo annotato con la dottrina e giurisprudenza*, Napoli, 2018, p. 633. Sulla responsabilità del produttore v., tra i tanti: AA.VV., *La responsabilità del produttore*, a cura di G. ALPA, M. BIN., P. CENDON, in *Trattato di dir. comm. e diritto pubblico dell'economia*, diretto da F. GALGANO, vol. XIII, Padova, 1989; G. ALPA, U. CARNEVALI, F. DI GIOVANNI, G. GHIDINI, U. RUFFOLO, C. M. VERARDI, *La responsabilità per danno da prodotti difettosi*, Milano, 1990, *passim*; U. CARNEVALI, voce *Responsabilità del produttore*, in *Enc. Dir. Aggiorn*., II, Milano, 1998, p. 936.

[31] https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX%3A31985L0374.

[32] Si v. *Relazione della Commissione al Parlamento europeo, al Consiglio e al Comitato economico e sociale europeo sull'applicazione della direttiva del Consiglio relativa al ravvicinamento delle disposizioni legislative, regolamentari ed amministrative degli Stati Membri in materia di responsabilità per danno da prodotti difettosi (direttiva 85/374/CEE), COM (2018) 246 final*.

Inoltre, si v. *Commission Staff Working Document Liability for emerging digital technologies Accompanying the document Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Artificial intelligence for Europe, SWD (2018) 137 final*.

lo ha voluto lo ha escluso (*ubi lex voluit dixit, ubi noluit taquit*), si potrebbe concludere la Direttiva sia applicabile – solo ed esclusivamente – ai beni immateriali espressamente indicati. Tale ricostruzione, però, è molto dibattuta[33] e, allo stato, risulta ancora impossibile prendere una posizione chiara e definita in merito alla qualificazione del *software* come prodotto o come servizio. Infatti, mentre negli Stati Uniti, questa qualificazione viene risolta applicando il c.d. "*essential nature test*", secondo cui è necessario analizzare la *ratio* del negozio giuridico concluso per determinare nel caso concreto la disciplina applicabile al *software*[34], in Europa non è ancora stata individuata una strategia unica per affrontare la questione.

Le ragioni sin qui esposte evidenziano come risulti auspicabile l'applicazione delle Direttive 85/374/CEE e 2001/95/CE, in quanto consentirebbe di applicare in materia tutti i vantaggi che la disciplina della responsabilità del produttore è in grado di offrire. Inoltre, comporterebbe l'adozione di tutte quelle strategie che l'Unione ha approvato al fine di rendere il mercato sempre più competitivo, da un punto di vista sia interno che internazionale: infatti, la necessità di una maggiore

---

[33] Sul punto si è aperto un dibattito dottrinario. Coloro che propendevano per l'applicazione della normativa in esame anche alle Intelligenze Artificiali si scontravano con il fatto che il loro elemento essenziale è rappresentato dal *software*: un bene immateriale che non sempre è incorporato in un *hardware.* Favorevole all'esclusione dei *softwares* dalla disciplina della Direttiva, J. TRIAILLE, *The EEC Directive on Product Liability and its Application to Databases and Information*, in *Computer Law and Practice*, 1991, p. 219. Si v., inoltre, D. WUYTS, *The Product Liability Directive: More than two decades of defective products in Europe*, in *Journal European Tort Law*, vol. 5, 2014. Ancora, si v. K. ALHEIT, *The Applicability of the EU Product Liability Directive to Software*, in *The Comparative and International Law Journal of Southern Africa*, vol. 34, n. 2, 2001, p. 188–209, che muove dalla considerazione che il *software* possa, di per sé, essere comunque considerato come un prodotto, in quanto la Commissione europea non ha mai posto in dubbio il fatto che un *software* potesse essere qualificato come tale, ritenendo applicabile la Direttiva 85/374/CEE in tutti i casi in cui il *software* sia incorporato in un bene mobile. Il problema che viene evidenziato deriva dalla caratteristica dell'immaterialità del *software* che ne impedirebbe la riconduzione alla disciplina. Infatti, nonostante si condivida che non dovrebbe rilevare la presenza dell'*hardware* per qualificare il programma come "prodotto", tuttavia, si osserva che dalla lettura dell'art. 2 è difficile desumere una diversa conclusione.

A questa impostazione si contrappone quella parte della dottrina (v., per esempio, G. WAGNER, *Robot Liability*, in S. LOHSSE, R. SCHULZE, D. STAUDENMAYER, (a cura di), *Liability for Robotics and in the Internet of Things: Munster Colloquia on Eu Law and the Digital Economy*, West Sussex, 2019, p. 36) che sottolinea l'ampiezza della definizione di "prodotto" offerta dal legislatore comunitario come intesa a comprendervi qualunque tipo di bene. In tal senso, si richiama quanto espresso dal terzo comma dell'art. 2, sostenendo che, in realtà, si tratterebbe di una indicazione meramente esemplificativa (ma non esaustiva) dei beni immateriali che si possono ricondurre alla generale categoria di "prodotto" ai sensi della normativa.

Ciò nonostante, al fine di fornire soluzione al problema della tangibilità del bene, si è proposto di concentrare l'attenzione nella manifestazione fisica che caratterizza il *software* al fine di reperire un elemento tangibile idoneo a qualificarlo come bene materiale, consentendone, pertanto, la riconduzione nella categoria del "prodotto": cfr. A. TETTENBORN, *Product Liability and Consumer Protection*, in *Clerk & Lindsell on Torts*, Fasc. 11, 2010, p. 11-49. Una soluzione anacronistica: in realtà, al giorno d'oggi, i *software* sono oggetto di contratti di compravendita *online*, e il supporto fisico è quasi del tutto scomparso. Pertanto, sarebbe impossibile, secondo questa impostazione dottrinaria, la riconduzione nella materia nella disciplina prevista della Direttiva 85/374/CEE. In merito all'impossibilità di qualificare il *software* come prodotto, si v. P. BORTONE, L. BUFFONI, *La responsabilità per prodotto difettoso e la garanzia di conformità nel codice del consumo*, Torino, 2007, p. 31. La qualificazione del *software* come servizio è accolta dalla dottrina statunitense (si v. S. CHOPRA, L. F. WHITE, *A Legal Theory for Autonomous Artificial Agents*, Michigan, 2011, p. 136).

Inoltre, si tende a presentare un paragone tra i *Softwares as a Service* – c.d. "*SaaS*" e i *Software as a Product* – i "*SaaP*" –, favorendo i primi ai secondi: v. *Saas vs. Saap: A Chillingly Honest Dissection*, disponibile su https://www.canto.com/blog/saas-vs-saap/.

[34] Cfr. K. ALHEIT, *The Applicability of the EU Product Liability Directive to Software*, *cit.*, p. 199. Inoltre, sull'applicabilità del "*essential nature test*", si v. *RRX Industries, Inc. v. Lab-Con, Inc.* (1985).

certezza del quadro normativo di riferimento è stata evidenziata anche dalla *Digital Single Market Strategy* (DSM[35]) e dalla Comunicazione "*Building a European Data Economy*"[36].

Ciò premesso, l'applicazione alle nuove tecnologie della disciplina in esame, presenta dei profili problematici che devono essere analizzati, anche nella prospettiva di un adeguamento alle caratteristiche specifiche tipiche delle Intelligenze Artificiali.

Per approfondire tali profili è interessante lo studio sulla disciplina in vigore svolta da un gruppo di esperti incaricati dalla Comunità europea che ha pubblicato uno studio dedicato alla direttiva e ad una valutazione della sua idoneità a fornire una completa regolazione delle nuove tecnologie[37]. Questo studio si aggiunge a quanto già espresso dalla Commissione europea in merito all'importante evoluzione tecnologica che si è avuta dal 1985 ad oggi e alla affermata necessità di procedere ad una valutazione della Direttiva che sia volta a verificarne l'efficienza, la pertinenza nella regolazione delle sfide che la tecnologia pone, l'efficacia per il raggiungimento degli obiettivi e, infine, far sì che la disciplina in tema di responsabilità ivi contenuta sia in grado di offrire un valore aggiunto sia alle imprese impegnate nel mercato, sia ai terzi indebitamente danneggiati[38].

Alla fine, il parere espresso dalla Commissione è risultato positivo, ritenendo ancora idonea la Direttiva del 1985 allo scopo per cui era stata emanata: ciò nonostante, si evidenzia la necessità di procedere ad un adeguamento della normativa, da realizzarsi attraverso un'interpretazione evolutiva delle nozioni fondamentali ivi previste: quali, ad esempio, quella di "prodotto", "produttore", "difetto", "danno" e "onere della prova". In considerazione della rilevanza che tali concetti rivestono nella analisi dei profili di responsabilità derivante dall'utilizzo di Intelligenze Artificiali, risulta opportuno soffermare brevemente l'attenzione sugli approfondimenti offerti dalla Commissione di esperti su menzionata.

Sulla nozione di "prodotto" si è avuto modo di esporre. Anche la Commissione evidenzia la presenza di alcune difficoltà nella qualificazione della Intelligenza Artificiale nei casi in cui si tratti di un non "*embedded software*"[39]: vale a dire in tutti i casi in cui il *software* viene programmato per funzionare in un sistema che non sia un *computer*.

Quanto, invece, alla nozione di "produttore", da una prima analisi della norma, emerge come la definizione sia molto ampia: infatti, non solo comprende il produttore finale e chiunque ne abbia prodotto un componente o una materia prima, ma anche chi importa materiali all'interno della Comunità, chi appone il marchio sul prodotto e ogni altro fornitore che non riesca a comunicare al consumatore l'identità del produttore in un tempo ragionevole[40]. La nozione evidenzia il *favor* voluto

---

[35] Si v. *Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni. Strategia per il mercato unico digitale in Europa, COM (2015) 192 final*.

[36] Si v. *Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni: "costruire un'economia dei dati europea", COM (2017) 9 final.*

[37] Cfr. *SWD (2018) 137 final, cit.*.

[38] Si v. *COM (2018) 246 final, cit.*.

[39] Un *embedded software* è un programma per *computer* che viene scritto per essere utilizzato e funzionare su macchine e dispositivi che generalmente non sono qualificati come *computer* (i c.d. "*embedded systems*"). Solitamente, il programma viene realizzato per lo specifico *hardware* in cui verrà inserito (come, ad esempio, un aeromobile o una navicella spaziale) e ha dei vincoli di tempo e memoria.

[40] Si v. D. FAIRGRIEVE, G. HOWELLS, P. MØGELVANG-HANSEN, G. STRAETMANS, D. VERHOEVEN, P. MACHNIKOWSKI, A. JANSSEN, R. SCHULZE, *Product Liability Directive*, in P. MACHNIKOWSKI (a cura di), *European Product Liability: An Analysis of the State of the Art in the Era of New Technologies*, Cambridge, 2016, p. 61.

dal legislatore nei confronti dei consumatori, prevedendo, da un lato, una loro più ampia protezione in caso di danno e, dall'altro, la possibilità (la garanzia) di poter citare in causa più soggetti per ottenere il risarcimento del danno[41].

Altro problema inerente alla figura del produttore riguarda la sua identificazione. Sul punto, la relazione del gruppo di esperti che si è occupato dell'analisi della disciplina, ha evidenziato che, se in molti casi il prodotto finale e il produttore possono risultare facili da individuare – indipendentemente dal fatto che si includano i *software* o altri elementi digitali, o dal numero di soggetti coinvolti nel processo –, in altri casi, potrebbe risultare molto più complesso. E ancora, la stessa relazione, evidenzia come un ulteriore elemento di criticità sia rappresentato dal grado di estensione della capacità del produttore di mantenere il controllo sulle caratteristiche di un prodotto in un contesto di tecnologie emergenti. Il produttore, infatti, potrà essere considerato responsabile solo nei limiti in cui la specifica caratteristica del prodotto, che si è rivelata causa del danno, si possa ricondurre al rischio che lo stesso ha assunto iniziando l'attività imprenditoriale[42].

Passando alla nozione di "difetto" (art. 6), dalla lettura della norma emerge come questa consideri il concetto di "difetto" indissolubilmente connesso con la nozione di "sicurezza": in questo caso, infatti, è necessario distinguere tra la responsabilità del prodotto e la sua affidabilità[43].

Occorre, però, evidenziare che quando si deve analizzare la sicurezza delle Intelligenze Artificiali, potrebbe risultare difficile accertare se il difetto risulti determinato dal programma o se, invece, derivi da altri elementi che sono interconnessi in un ecosistema digitale[44]. Sul punto, sarebbe opportuno, pertanto, valutare quali siano gli *standard* di sicurezza richiesti per ciascuno degli elementi coinvolti nell'attività dell'Intelligenza Artificiale, considerando tutti i rischi specifici coinvolti dalle nuove tecnologie.

Inoltre, risulta molto complicato determinare la difettosità di un *software*: in particolare, i sistemi avanzati di *deep learning* e *machine learning* mettono in seria discussione il criterio su indicato secondo cui è indispensabile determinare ciò che il consumatore si potrebbe legittimamente aspettare dall'applicazione in quanto, grazie alle loro elevate capacità cognitive, non consentono una predizione certa di ciò che potrà effettivamente accadere quando saranno utilizzate dallo stesso.

A ciò si aggiunga che manca una definizione di "difetto del *software*" che renda più semplice l'applicazione della disciplina e la determinazione della responsabilità del produttore.

---

[41] Cfr. G. HOWELLS, *Defect in English Law: lesson for the harmonization of European product liability*, Cambridge, 2005.

[42] Cfr. *COM (2018) 246 final*, *cit.*.

[43] Si v. D. C. VLADECK, *Machines without Principals: Liability, Rules and Artificial Intelligence*, in *Washington Law Review*, vol. 89, n. 117, 2014, p. 137. La direttiva 85/374/CEE disciplina il risarcimento *ex post* per danni sofferti da consumatori a causa di un prodotto difettoso. Tuttavia, si sostiene che vi siano altre disposizioni all'interno dell'ordinamento dell'Unione europea (ad esempio, si richiama la direttiva 2001/95/CE), che prevedono una prevenzione *ex ante* del danno, attraverso la determinazione (e l'imposizione) di *standard* produttivi finalizzati a garantire che i prodotti immessi nel mercato europeo siano qualificabili come affidabili. Le due discipline sono complementari, in quanto più un prodotto è conforme agli *standard* individuati dal legislatore, meno il consumatore dovrà (potrà) rivolgersi al giudice per l'ottenimento del risarcimento del danno. Il prodotto, inoltre, non potrà essere considerato come difettoso solo per il semplice fatto che un altro prodotto migliore sia stato messo in circolazione nello stesso mercato.

La disposizione, pertanto, mette in relazione il difetto del prodotto con la sua sicurezza, senza considerare la sua idoneità al raggiungimento dello scopo per cui è stato realizzato: per la disciplina di quest'ultima caratteristica, troveranno applicazione altre disposizioni relative alla regolazione della vendita dei prodotti.

Per la determinazione della sicurezza che ci si può legittimamente attendere, sarà necessario valutare caso per caso ciò che, oggettivamente, il consumatore si sarebbe potuto aspettare dal prodotto acquistato.

[44] Così *SWD (2018) 137 final*, *cit.*, p. 18.

Parte della dottrina ha tentato di colmare questa lacuna, facendo riferimento alla nozione di difetto inteso come errore strutturale che determina il fallimento del sistema quando processa i dati che gli vengono forniti. In tale prospettiva, si afferma, l'errore del *software* sarebbe un errore materiale presente nel codice e commesso dal programmatore dello stesso. Così inteso, l'errore può essere determinato sia da un comportamento commissivo, sia da un'omissione: nel primo caso, si implementeranno nel codice delle parti che non erano comprese nel progetto originario; nel secondo caso, invece, si presenteranno delle mancanze che renderanno il codice incompleto.

Il problema, come rilevato dalla dottrina, è che manca anche un metodo per la individuazione dell'errore che sia certo e che consenta una valutazione oggettiva dello stesso applicabile a tutti i tipi di *software*.

In particolare, come già detto, non esiste un *test* che consenta di stabilire con certezza quando un *software* è difettoso. Questo comporta, in primo luogo, il rischio che si attribuisca un errore all'algoritmo, quando questo, in realtà, non presenta alcun difetto. In secondo luogo, nonostante il produttore possa effettuare numerosi *test* sul *software* per controllare che non ci siano *bug*, tuttavia non ci potrà mai essere la certezza assoluta che in futuro non ci saranno degli errori di elaborazione dei dati[45].

Si rischierebbe, pertanto, di appesantire l'onere della prova richiesta al consumatore, richiedendo una vera e propria *probatio diabolica* finalizzata a dimostrare la responsabilità del produttore[46]. Infatti, un approccio che miri all'eliminazione di qualunque difetto dell'algoritmo appare impossibile, dal momento che è sufficiente un minimo errore per provocare danni anche notevoli. Nel 2015 è stata pubblicata una strategia chiamata "*Horizon 2020*" che si poneva l'obiettivo di creare un settore di produzione in ambito tecnologico che avesse come obiettivo lo "*zero-defect manufacturing*"[47]. Questo programma ha sicuramente dato la spinta per permettere ai produttori di minimizzare il più possibile gli errori di produzione ma ha anche evidenziato che, come detto, l'eliminazione totale di questi è obiettivo, al momento, irraggiungibile.

Infine, in questa situazione di incertezza, anche la qualità dei dati cui l'Intelligenza Artificiale ha accesso assume particolare rilevanza[48]. Infatti, il tipo e qualità dei dati forniti al sistema diventano estremamente rilevanti quando si cerca di dimostrare che il prodotto – pur non essendo difettoso *ex sé* – tuttavia sia stato addestrato in maniera errata e che, per questo motivo, abbia contribuito alla produzione del danno. In tale prospettiva, quando si cerca di determinare l'idoneità dell'allenamento del *software,* i *set* di dati utilizzati per addestrare la macchina, dovrebbero essere disponibili per una valutazione della loro qualità ed adeguatezza.

L'art. 9 della direttiva, fornisce una definizione di "danno": la questione sorta in merito concerne l'opportunità di continuare a includere in queste categorie di danno anche quelli derivanti dall'utilizzo delle Intelligenze Artificiali. Il gruppo di esperti che si è occupato dell'analisi della direttiva, infatti, si è domandato quali siano i danni che debbano essere compensati ponendo, particolare attenzione, al danno morale. Inoltre, sarebbe anche necessario valutare se in caso di evento dannoso provocato da sistemi di Intelligenza Artificiale sia opportuno individuare un tetto massimo al danno risarcibile: in caso affermativo, occorrerà considerare l'ipotesi di adottare una applicazione

---

[45] Si v. F. E. ZOLLERS *et al.*, *No More Soft Landings for Software: Liability for Defects in an Industry That Has Come of Age*, in *Santa Clara High Technology Law Journal*, vol. 21, n. 4, 2005, p. 750-753. Inoltre, si v. I. J. LLOYD, *Information Technology Law*, Oxford, 2008, p. 562.

[46] J. S. BORGHETTI, *How Can Artificial Intelligence Be Defective?*, in S. LOHSSE, R. SCHULZE, D. STAUDENMAYER (a cura di) *Liability For Artificial Intelligence and The Internet Of Things*: *Münster Colloquia on EU Law and the Digital Economy IV*, Baden, 2019, p. 64.

[47] *Horizon 2020, Intelligent Approach to Zero-defect Manufacturing*, 2015.

[48] Si v. A. BACHMANN, A. BERNSTEIN, *Software Process Data Quality and Characteristics: A Historical View on Open and Closed Source Projects*, in *Proceedings of The Joint International and Annual ERCIM Workshops on Principles of Software Evolution*, 2009, p. 119–128.

della stessa cifra per tutti i tipi di danno, indifferentemente dal tipo di tecnologia che l'ha provocato, oppure se operare delle distinzioni e di che tipo. Tali riflessioni sono rilevanti, in quanto le Intelligenze Artificiali possono aumentare il rischio di produzione di determinati danni, soprattutto morali, oppure possono determinare il sorgere di nuovi tipi di rischio, strettamente legati alla loro particolare natura[49]: si pensi, ad esempio, alla *cyber*sicurezza o al pericolo di violazione del diritto alla *privacy*.

Nel momento in cui si verifica un evento dannoso, il danneggiato che intende agire per ottenere il risarcimento del danno, sarà gravato dall'onere della prova nei termini indicato dall'art. 4: dovrà fornire la dimostrazione dell'esistenza del danno, del difetto esistente e del nesso teleologico che collega il difetto al danno subito[50]. Il danno risarcibile, pertanto, è solamente quello prodotto a causa del difetto.

Come già accennato, l'onere della prova grava sul consumatore danneggiato: trattandosi di un'ipotesi di responsabilità oggettiva, non sarà necessario fornire prova del dolo o della colpa del produttore[51] ma occorrerà dimostrare l'esistenza del danno, del difetto e che sussiste un collegamento fra il danno e la condotta posta in essere (nesso causale). In altre parole, l'evento dannoso dev'essere ricondotto al difetto e, quest'ultimo, deve dipendere da un errore commesso dal produttore. L'esistenza di questi collegamenti (eziologici) non è facile da dimostrare, in particolare quando si tratta di Intelligenze Artificiali. Questo comporta che l'unica via perseguibile è quella di chiedere al produttore – nell'ambito della sua "sfera di controllo"[52] – di fare tutto quanto è possibile per prevedere tutte le condizioni esistenti che potrebbero determinare il sorgere di un danno e di agire, conseguentemente, per mettere in atto quanto necessario a evitare (a prevenire) che questi eventi si possano verificare. In questa prospettiva, assume particolare rilevanza definire esattamente che cosa rientri nella "sfera di controllo" del produttore quando mette in commercio un prodotto come quelli

---

[49] Si v. *SWD (2018) 137 final*, *cit.*, p. 21.

L'art. 7 della Direttiva 85/374/CEE, in esame, inoltre, prevede dei casi di esclusione della responsabilità del produttore, il quale avrà la possibilità di provare il ricorrere di una serie di circostanze normativamente indicate che gli permettono di non essere onerato del risarcimento del danno

Ancora, secondo quanto indicato nel successivo art. 8 è irrilevante se al verificarsi dell'evento dannoso, oltre insieme al difetto del prodotto, abbia contribuito anche il concorso da parte di un terzo, fatte salve le disposizioni nazionali in materia di diritto alla rivalsa. Tuttavia, il secondo comma della norma, dispone la rilevanza parziale o esclusiva del concorso di colpa da parte del danneggiato stabilendo che «*la responsabilità del produttore può essere ridotta o soppressa, tenuto conto di tutte le circostanze, quando il danno è provocato congiuntamente da un difetto del prodotto e per colpa del danneggiato o di una persona di cui il danneggiato è responsabile*». Si tratta di una disposizione particolarmente rilevante soprattutto in tutti i casi in cui il consumatore utilizzi un sistema di Intelligenza Artificiale ad alta capacità cognitiva. Infatti, è necessario stabilire il limite entro cui il produttore può, legittimamente, contestare l'uso improprio da parte del consumatore, soprattutto nelle ipotesi in cui il prodotto è realizzato in modo tale da consentire una modificazione da parte di terzi dopo essere stato reso disponibile sul mercato[49]. Sul punto, la giurisprudenza statunitense ritiene che il produttore debba essere considerato responsabile in tutti i casi in cui i danni siano prevedibili, prevenibili o mitigabili.

[50] Si v. art. 4, dir. 85/374/CEE.

[51] Con il termine "produttore" si comprendono tutti i soggetti riconducibili alla definizione ai sensi dell'art. 2 della direttiva 85/374/CEE.

[52] J. M. FISCHER, M. S. J. RAVIZZA, *Responsibility and Control*: *A Theory of Moral Responsibility*, Cambridge, 1998, p. 13.

indicati che sono caratterizzati da una particolare autonomia: parlare di controllo e di autonomia è un ossimoro.

Molti Stati Membri, infatti, hanno adottato la c.d. "*development risk defence*"[53], che consente l'esclusione della responsabilità se lo stato della conoscenza scientifica e tecnologica al tempo in cui il prodotto è stato messo in circolazione non avrebbe consentito al produttore di individuare – ed eliminare – il difetto che ha provocato il danno (cfr. art. 7, dir. 85/374/CEE). Questo tipo di difesa potrebbe diventare sempre più importante, man mano che la tecnologia delle Intelligenze Artificiali si svilupperà: se non era prevedibile non era evitabile e, pertanto, non sussiste responsabilità del produttore.

Tuttavia, il *Report* "*Liability for Artificial Intelligence and Other Emerging Digital Technologies*"[54] del 2019 ha proposto una soluzione ai problemi che emergono dall'analisi della direttiva, ritenendo più opportuno mantenere un regime di responsabilità oggettiva per l'individuazione dei danni causati da prodotti difettosi ed eliminando ogni distinzione tra prodotti materiali e immateriali. Inoltre, sostenendo che il produttore dovrebbe essere oggettivamente responsabile per i danni derivanti da tecnologie emergenti anche se il difetto è emerso dopo che il prodotto era stato messo in circolazione, se si dimostra che il produttore avrebbe ancora potuto esercitare un controllo sul prodotto anche successivamente, attraverso degli aggiornamenti migliorativi.

In ogni caso, nonostante vi fosse il parare favorevole da parte degli Stati Membri, il gruppo di esperti incaricato ritenne inopportuno favorire e sostenere l'utilizzo della c.d. "*development risk defence*".

A questo punto, risulta opportuno analizzare l'ipotesi di ricondurre le Intelligenze Artificiali nell'ambito della disciplina giuridica prevista per i dispositivi medici.

3.3.1 In particolare: l'Intelligenza Artificiale intesa come dispositivo medico. – Nell'ordinamento dell'Unione europea, la definizione di "dispositivo medico" è contenuta nell'art. 1, comma 2, lett. a) della direttiva 93/42/CEE: il termine è applicato a ogni tipo di strumento, incluso il *software* informatico, «*impiegato per il corretto funzionamento e destinato dal fabbricante ad esser impiegato nell'uomo a scopo di: diagnosi, prevenzione, controllo, terapia o attenuazione di una malattia*»[55]. Questa definizione è stata accolta dalle Linee guida sui dispositivi medici che sono state approvate dalla Commissione europea per indirizzare gli *stakeholders* al rispetto della nuova disciplina.

Il quadro normativo europeo ora in vigore deriva da tre direttive regolanti i dispositivi medici (dir. 93/42/CEE, dir. 90/385/CEE e dir. 98/79/CEE) e richiede ai fabbricanti di garantire che il

[53] Si v. K. ALHEIT, *The Applicability of the EU Product Liability Directive to Software*, *cit.*, p. 204.

Infine, si v. *Potter v. Chicago Pneumatic Tool Co.*, Supreme Court of Connecticut, 1997: per dimostrare la responsabilità del produttore, è necessario provare che il prodotto non fosse perfetto secondo lo stato di sviluppo delle conoscenze tecnologiche del tempo, in quanto sarebbe stato possibile apporre ulteriori modifiche o aggiornamenti al sistema.

[54] Si v. *Report* "*Liability for Artificial Intelligence and Other Emerging Digital Technologies*", 2019, p. 42.

[55] Art. 1, c. 2, lett. a): «*dispositivo medico: qualsiasi strumento, apparecchio, impianto, sostanza o altro prodotto, utilizzato da solo o in combinazione, compreso il* software *informatico impiegato per il corretto funzionamento e destinato dal fabbricante ad esser impiegato nell'uomo a scopo di: diagnosi, prevenzione, controllo, terapia o attenuazione di una malattia; diagnosi, controllo, terapia, attenuazione o compensazione di una ferita o di un* handicap*; studio, sostituzione o modifica dell'anatomia o di un processo fisiologico; intervento sul concepimento, la cui azione principale voluta nel o sul corpo umano no sia conseguita con mezzi farmacologici né immunologia né mediante metabolismo, ma la cui funzione possa essere assistita da questi mezzi*».

dispositivo prodotto sia adeguato al raggiungimento dello scopo per cui è fabbricato. Ciò significa che le direttive individuano una serie di requisiti che devono essere rispettati. A seconda della classificazione del rischio del dispositivo, il produttore deve ottenere l'approvazione e la conferma della sussistenza dei requisiti minimi richiesti, da parte di un comitato indipendente nominato dalle autorità degli Stati Membri.

Questa disciplina è stata riformata dal nuovo Regolamento sui dispositivi medici (MDR) e dal Regolamento sui dispositivi diagnostici in vitro (IVDR), entrambi entrati in vigore il 25 maggio del 2017: il primo ha trovato applicazione a partire dal 26 maggio 2020; il secondo, a partire dal 26 maggio 2022. Tale riforma si è resa necessaria in quanto le direttive degli anni '90 non si sono rivelate idonee ad affrontare le nuove tecnologie: comprese i sistemi di Intelligenza Artificiale. Inoltre, trattandosi di Regolamenti, a differenza delle direttive trovano diretta applicazione negli ordinamenti degli Stati Membri, senza che i legislatori debbano recepire la disciplina europea.

Alcune delle caratteristiche principali di questa nuova disciplina sono: l'inclusione di una gamma più ampia di prodotti, l'estensione della responsabilità con riferimento ai prodotti difettosi, il rafforzamento dei requisiti richiesti per i dati clinici e la tracciabilità dei dispositivi, un monitoraggio più rigoroso da parte dei comitati e una maggiore trasparenza attraverso la pubblicazione di tutte le informazioni relative ai dispositivi medici[56].

In particolare, con riferimento ai sistemi di Intelligenza Artificiale, il Considerando 19 precisa che «*è necessario precisare che il* software *specificamente destinato dal fabbricante a essere impiegato per una o più delle destinazioni d'uso mediche indicate nella definizione di dispositivo medico si considera un dispositivo medico, mentre il* software *destinato a finalità generali, anche se utilizzato in un contesto sanitario, o il* software *per fini associati allo stile di vita e al benessere non è un dispositivo medico. La qualifica di* software*, sia come dispositivo sia come accessorio, è indipendente dall'ubicazione del software o dal tipo di interconnessione tra il* software *e un dispositivo*»[57]. Inoltre, lo stesso Regolamento conferma la necessità di individuare requisiti minimi per *software* e *hardware* che rientrino nella categoria dei "dispositivi medici".

Pertanto, nell'ordinamento europeo un *software* di Intelligenza Artificiale è qualificabile come "dispositivo medico"[58] quando esso (i) sia un programma informatico (e non un semplice documento digitale); (ii) svolga una funzione diversa e ulteriore rispetto alla mera conservazione o trasmissione o ricerca di dati; (iii) operi a beneficio di specifici pazienti; e (iv) svolga una delle funzioni incluse nella definizione di dispositivo medico (vale a dire diagnosi, prevenzione, controllo o terapia di una malattia o di una ferita o di un *handicap*; studio, sostituzione o modifica dell'anatomia o di un processo fisiologico; intervento sul concepimento).

Una volta qualificato il *software* come dispositivo medico (c.d. SaMD: "*Software as Medical Device*"), troverà applicazione la disciplina corrispondente, vale a dire il Regolamento 2017/745 (*Medical Device Regulation*), che si applica dal 26 maggio 2020.

Secondo tale disciplina, in primo luogo è necessario stabilire a quale classe appartiene tale applicazione: il documento di riferimento è l'allegato VIII del Regolamento, in cui è previsto che i *software* classificati come "dispositivi medici" rientrano nella prima classe, ad eccezione di quelli destinati a monitorare i processi fisiologici, che rientrano nella classe II a) o II b), nel caso in cui

---

[56] Si v. S. CROSSLEY, L. L. P. EVERSHEDS, *EU Regulation of Health Information Technology, Software and Mobile Apps*, in *Practical Law Global Guide 2016(17)*, 2016, p. 1-14.

[57] Si v. *considerando 19*, Reg. 2017/745.

[58] Le Linee guida approvate dalla Commissione, cui si è già fatto riferimento, distinguono tra *software* e *stand-alone software*: i secondi, a differenza dei primi, non sono incorporati in un dispositivo medico al momento di immissione nel mercato. Alcuni esempi concreti di *stand-alone* si possono trovare nel *Manual on Borderline and Classification in the Community Regulatory Framework for Medical Devices*, disponibile su: https://ec.europa.eu/docsroom/documents/35582.

oggetto di monitoraggio siano parametri vitali la cui variazione può creare un pericolo immediato per il paziente.

Un altro aspetto importante riguarda *l'iter* per l'ottenimento del marchio CE, che varia a seconda della classe di appartenenza. Per esempio, nel caso di un SaMD è necessario produrre una serie di documenti, quali:

- la documentazione tecnica, contenente la descrizione e le specifiche del dispositivo e le specifiche dei requisiti *software* (i requisiti di sicurezza e prestazioni); informazioni sull'uso e l'installazione, che devono essere fornite dal fabbricante; informazioni sulla progettazione e fabbricazione, inclusa l'analisi dei rischi e dei benefici e la gestione del rischio;
- la documentazione della sorveglianza *post* commercializzazione, che comprende il piano di sorveglianza (ad esempio, la gestione dei reclami); i processi e le procedure di azioni correttive, preventive e di gestione delle modifiche e di aggiornamento; i processi e le procedure per garantire identificazione e rintracciabilità del dispositivo, nonché le azioni di sostituzione; infine, un rapporto periodico di aggiornamento sulla sicurezza;
- dichiarazione di conformità europea.

Come ogni dispositivo medico, anche un *"Software as Medical Device*" è soggetto a una procedura di identificazione, comune a livello europeo, che ha lo scopo di garantire la massima rintracciabilità nella catena di fornitura, facilitando – se necessario – le procedure di richiamo dei prodotti che dovessero presentare dei rischi per la sicurezza. Questo tipo di identificazione viene chiamato UDI, ed è caratterizzata da due elementi: UDI-DI e UDI-PI.

Il primo, UDI-DI è un codice numerico o alfanumerico unico, utilizzato anche come chiave di accesso alle informazioni memorizzate in una banca dati UDI: tale codice è specifico per ogni modello di dispositivo.

Il secondo, UDI-PI, invece, è un codice numerico o alfanumerico che indica la specifica unità di produzione del dispositivo.

Per le applicazioni di SaMD, l'identificazione richiede che a ogni nuova versione di *software* sia associato un nuovo codice UDI. Inoltre, ogni volta che viene rilasciata una versione del *software* che presenti una modifica significativa (riguardanti, per esempio, le prestazioni e l'efficacia originali, la sicurezza per il paziente e l'utilizzatore oppure l'uso previsto del SaMD) tale *software* deve essere considerato come "prodotto nuovo" e, pertanto, si richiede l'emissione di un nuovo UDI-DI. Inoltre, le modifiche significative possono includere, tra le altre, algoritmi nuovi o modificati, strutture di *database*, la piattaforma operativa, l'architettura, nuove interfacce utente o nuovi canali per l'interoperabilità. Infine, ogni volta che viene introdotta una modifica "minore" (come la correzione di *bug*, miglioramenti riguardanti l'uso, ma non collegato alla sicurezza fisica, interventi sulla sicurezza delle informazioni, miglioramenti dell'efficienza) la nuova versione deve essere identificata con un nuovo UDI-PI.

Tuttavia, ci sono delle applicazioni *software* che non sono qualificabili come SaMD. In questi casi, trovano applicazione le "*EU Guidelines on Assessment of the Reliability of Mobile Health Applications*", approvate dalla Commissione europea.

Tali Linee guida prevedono una dettagliata analisi di rischi e una valutazione dei criteri dell'applicazione *software* che devono essere rispettati. Tra questi, si ricordano:

- affidabilità: l'applicazione deve essere in grado di funzionare correttamente in diversi ambienti;
- stabilità: valuta le reazioni dell'applicazione in situazioni critiche, come per esempio la perdita di connessione alla rete o perdita di energia;

- efficacia: cerca di identificare prove dell'efficacia dell'applicazione nel raggiungimento degli obiettivi dichiarati;
- usabilità e Accessibilità: valuta se l'applicazione è utilizzabile dalle persone a cui è indirizzata;
- trasparenza: i soggetti che hanno permesso la realizzazione dell'applicazione (chi lo ha finanziato, per quale motivo, chi detiene i dati dell'utente ecc.) devono essere conosciuti;
- credibilità: le caratteristiche funzionali e la metodologia offerta dall'applicazione devono basarsi su appropriata documentazione devono essere convalidate da organismi autorizzati;
- sicurezza: l'applicazione deve essere impostata in modo appropriato rispetto alle aspettative dell'utente per un funzionamento sicuro;
- sicurezza e *privacy*: l'applicazione deve essere programmata nel rispetto della normativa relativa alla protezione dei dati personali (*cfr*. Reg. 2016/679, *General Data Protection Regulation*);
- desiderabilità: l'applicazione, per come è presentata, dovrebbe invogliare gli utilizzatori ad un uso prolungato nel tempo.

Nell'ordinamento statunitense, a partire dalla fine del 2016, il *21st Century Cures Act* ha chiarito lo scopo della giurisdizione regolamentare della *Food and Drug Administration* (FDA) nell'ambito dei *software* utilizzati nel settore sanitario, specificando che un dispositivo medico è uno strumento utilizzato per la diagnosi di una malattia o un'altra condizione, o per la cura, l'attenuazione, o la prevenzione di una malattia, sia dell'essere umano, sia di animali, o, ancora, utilizzato per colpire la struttura o qualunque altra funzione del corpo umano o di un animale.

Pertanto, secondo questa definizione, ogni Intelligenza Artificiale che persegue uno degli scopi indicati sarà regolato dalla FDA, come previsto dal *Federal Food Drug and Cosmetic Act*. La FDA classifica i dispositivi medici in tre categorie, a seconda dell'utilizzo e del rischio che pongono e li disciplina in maniera diversa: maggiore il rischio, maggiore il controllo.

La natura di "scatole nere" e la rapida crescita delle applicazioni dei sistemi di *machine learning* e di *deep learning* renderà più complicato il processo di approvazione dei nuovi dispositivi medici, in quanto ciascuno di questi richiede un elevato numero di *test* e di verifiche che possono essere anche molto complessi. Per fare un esempio, l'introduzione, nel 1998, del *software* di rilevazione assistita per le mammografie[59] ha richiesto molti anni ed enormi pressioni politiche per fare in modo che la FDA ne consentisse l'utilizzo come secondo lettore di *screening*[60].

Invece, la questione si complica notevolmente quando si cerca di ottenere l'approvazione per sistemi di Intelligenza Artificiale che presuppongano una sostituzione del medico: in questo caso, la FDA impone dei requisiti così severi da risultare opprimenti, scoraggiando i produttori. Tuttavia, la ragione per cui i criteri diventano così stringenti non è di carattere prettamente giuridico: infatti, si richiede una discussione sulla responsabilità etica che coinvolge questi sistemi e, per il momento, non è ancora stata individuata una soluzione tale da consentire l'approvazione per l'inserimento nel mercato di macchine di questo tipo.

---

[59] A. J. MÉNDEZ, P. G. TAHOCES, M. J. LADO, M. SOUTO, J. J. VIDAL, *Computer-aided Diagnosis: Automatic Detection of Malignant Masses in Digitized Mammograms,* in *Medical Physics*, vol. 25, 1998, p. 957–964.

[60] E. AZAVEDO, S. ZACKRISSON, I. MEJÀRE, M. HEIBERT ARNLIND, *Is Single Reading with Computer-Aided Detection (CAD) as Good as Double Reading in Mammography Screening? A Systematic Review*, in *BMC Medical Imaging*, vol. 12, 2012, p. 22.

3.4. L'Intelligenza Artificiale più evoluta considerata come una persona. – Come anticipato, le categorie giuridiche esistenti possono risolvere i problemi giuridici riguardanti la maggior parte delle Intelligenze Artificiali ma non riescono ad inquadrare quelle intelligenze più evolute ed autonome: di *machine learning* e di *deep learning.* Per affrontare questa situazione, la dottrina ha anche valutato la possibilità di riconoscere i *robot* più evoluti e con più alte capacità cognitive come soggetti del diritto.

Come è immaginabile, questa proposta pone una serie di problemi.

In primo luogo, si deve capire se, dal punto di vista giuridico, sia più opportuno ricondurre l'Intelligenza Artificiale verso la categoria delle "persone giuridiche" (una finzione giuridica che attribuisce autonoma personalità a enti come, ad esempio, le Società), oppure se – come pare emergere dalle ricerche effettuate dagli studiosi nel settore – l'obiettivo ultimo sia quello di creare un *robot* che presenti le medesime caratteristiche (giuridiche) riconosciute all'essere umano. A sostegno di quest'ultima ipotesi ci sono anche alcuni studi neuroscientifici che escludono l'esistenza del libero arbitrio nell'uomo, ritenendo gli esseri umani determinati esattamente come lo sono gli algoritmi.

Con riferimento alle Società, si è eccepito che comunque, queste non sono in grado di agire autonomamente ma solo (e sempre) attraverso un rappresentante legale: un essere umano. Il che contrasterebbe con le intenzioni di chi vuole rendere completamente autonome (sotto il profilo giuridico) le intelligenze Artificiali più evolute: alla fine sarebbero sempre riconducibili ad un loro rappresentante legale.

In tal senso, si rileva la differenza esistente tra le persone giuridiche (come enti e società) e le Intelligenze Artificiali: al contrario di quanto accade per le società, secondo questa impostazione, le Intelligenze Artificiali verrebbero considerate come veri attori reali nella rete delle relazioni sociali che instaurano e, di conseguenza, definire opzioni più efficaci per ottenere da loro il risarcimento del danno[61].

Già dagli inizi degli anni '90, quando ancora l'idea di macchine così autonome da non poter essere controllate sembrava una mera ipotesi futuristica, si è iniziato a parlare della creazione della "personalità elettronica" da riconoscere alle Intelligenze Artificiali[62]. A quell'epoca, alcuni Autori[63] hanno sostenuto che se il programma informatico è strutturato in modo tale da non poter essere ragionevolmente prevista quale sarà la condotta che porrà in essere, allora si dovrebbe concludere riconoscendogli una autonomia anche dal punto di vista giuridico, liberando in parte il produttore dell'onere della responsabilità dei danni indebiti provocati dal *robot.*

In questa prospettiva, i *robot* verrebbero considerati come unica entità e, conseguentemente, dovrebbero essere dotati di un fondo economico autonomo per poter far fronte agli eventuali risarcimenti. Certamente, si risolverebbero alcuni problemi molto importanti e, primo fra tutti, si eliminerebbe la necessità – da parte dell'utilizzatore – di individuare il difetto che ha provocato il danno: se il *robot* è un soggetto e la sua condotta ha provocato un danno, allora sarà il *robot* in quanto entità unica a dover risarcire il danno, attraverso il fondo di cui viene dotato.

A supporto della impostazione, parte della dottrina[64], evidenzia il tema del c.d. *responsibilty gap*: cioè, la mancanza di potere di controllo da parte del produttore sull'attività futura posta in essere dall'Intelligenza Artificiale. Quando un algoritmo è in grado di imparare da *set* di dati e prendere delle decisioni partendo da delle informazioni che non sono predefinite dal programmatore, allora lo

---

[61] Per un quadro approfondito sul punto, si v. B. BROZEK, M. JAKUBIEC, *On the Legal Responsibility of Autonomous Machines*, in *Artificial Intelligence Law*, vol. 25, 2017, p. 301.

[62] Si v. L. B. SOLUM, *Legal Personhood for Artificial Intelligences*, in *North Carolina Law Review*, vol. 70, n. 4, 1992, p. 1231-1288.

[63] *Id.*.

[64] A. MATTHIAS, *The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata*, in *Ethics and Information Technology*, vol. 6, 2004, p. 175 ss.

stesso può essere qualificato come un sistema autonomo. Questa caratteristica del programma rende, da un lato, imprevedibile la sua condotta e, dall'altro, difficilmente comprensibili i processi decisionali che hanno portato l'algoritmo a concludere il ragionamento e ad adottare quella determinata soluzione: si pensi al problema della c.d. "*Black Box*"[65]. Per questo motivo, si ritiene opportuno orientarsi verso il riconoscimento dell'Intelligenza Artificiale come soggetto del diritto, attribuendole una personalità giuridica.

Inoltre, come detto, il denaro necessario per far fronte alla riparazione dei danni potenziali a terzi, potrebbe essere depositato in un apposito fondo, alimentato dai produttori delle singole parti della macchina e dagli investitori nella ricerca per lo sviluppo di quella tecnologia. In questo modo, si potrebbe ottenere, non solo la garanzia risarcitoria per i terzi danneggiati, ma anche la possibilità di riconoscere ai produttori una protezione, in quanto il deposito di quanto necessario alla riparazione dei danni potenziali consente loro una separazione (una autonomia) patrimoniale necessaria a evitare un'eccessiva (e incontrollata) esposizione a rischio.

Per tali ragioni, si è affermato[66], che il riconoscimento della personalità giuridica servirebbe principalmente come metodo per consentire una limitazione della responsabilità (c.d. *capping*). Inoltre, potrebbe anche non cambiare la persona che sopporta i costi del funzionamento di questo meccanismo e nemmeno i casi in cui il risarcimento sia aggiudicato. Infatti, a meno che il *robot* non sia in grado di ricevere un compenso dalla sua attività, il suo capitale dovrebbe essere conferito da un umano, o da una società, che necessariamente sostiene economicamente l'operazione e trae profitto dall'utilizzo della macchina. Ne consegue che, pertanto, pur riconoscendo personalità giuridica alla macchina, non ci sarebbero sostanziali modificazioni per quanto attiene il peso dell'onere della prova che opererebbe secondo le regole attualmente vigenti: essendo una ipotesi di responsabilità oggettiva, l'obbligo di risarcimento del danno verrebbe attribuito al creatore del *robot*, che lo rappresenterebbe legalmente.

Una delle questioni che sorgono in merito all'opportunità di riconoscimento della personalità giuridica in capo ai sistemi autonomi riguarda il fatto che, per essere considerato responsabile – e onerato del risarcimento del danno – il *robot* dovrebbe poter guadagnare dall'esercizio della sua attività ed essere (autonomamente) titolare di un patrimonio[67].

Ad ogni modo, un risultato molto simile potrebbe essere raggiunto prevedendo l'obbligo di assicurare le Intelligenze Artificiali a beneficio di terzi, magari prevedendo un tetto massimo all'entità del danno risarcibile (c.d. "*stop-loss*"), come ammesso dall'art. 16 della direttiva.

Un altro vantaggio risulta dal fatto che per il consumatore sarà più facile dimostrare l'esistenza del nesso causale tra il danno e la condotta posta in essere dal *robot*: infatti, per accedere al fondo e ottenere il risarcimento si potrà ritenere sufficiente la prova dell'esistenza di un collegamento esistente fra il danno arrecato e l'azione (o l'omissione) realizzata dal *robot*.

Una volta effettuato il pagamento, ci potrà essere una azione di rivalsa fra tutti coloro che hanno contribuito alla realizzazione della macchina e/o del programma e l'onere della prova ricadrà su di loro, i quali saranno anche tenuti (secondo le diverse responsabilità) a contribuire alla ricostituzione del fondo attribuito al *robot*: sarà interesse dei vari produttori individuare il vero e

---

[65] F. PASQUALE, *The Black Box Society: The Secret Algorithms that Control Money and Information*, Boston, 2015.

[66] A. BERTOLINI, *Robots as Products: The Case for a realistic Analysis of Robotic Application and Liability Rules*, in *Law, Innovation and Technology*, vol. 5, n. 2, 2013, p 213.

[67] Sulle conseguenze derivanti dalla possibilità della macchina di essere autonomamente retribuita per l'attività svolta, si v. A. BERTOLINI, *op. cit.*. Per un'analisi più approfondita sul punto, si v. G. COMANDÈ, *Risarcimento del danno alla persona e alternative istituzionali*, Torino, 1999, p. 333 ss..

diretto responsabile all'interno della catena di produzione e obbligarlo al versamento di quanto dovuto a favore del fondo.

Ciò che rileva, pertanto, è la volontà di individuare una via adeguata a sostenere uno sviluppo controllato della tecnologia: se si consente la separazione del rischio attraverso un meccanismo simile a quello appena descritto, si incentiva la produzione e la ricerca nell'ambito delle nuove tecnologie, mantenendo, tuttavia, adeguati livelli di sicurezza e di protezione per i consumatori.

L'idea è sostenuta anche dal Parlamento europeo, nella Risoluzione del 2017. Tale documento immagina due scenari di responsabilità: il primo è costituito dalla responsabilità oggettiva, di cui si è già parlato, in cui è sufficiente dimostrare il danno, il difetto e il nesso causale tra i due elementi. Il secondo, invece, è rappresentato dalla necessità di una più articolata valutazione dei rischi e della effettiva capacità di controllo di colui che, nella situazione corrente, è in grado di minimizzare gli esiti negativi della condotta dell'Intelligenza Artificiale[68].

Nel primo caso, si richiama quanto esposto in merito alle difficoltà di ricostruzione del nesso causale. Inoltre, quando si parla di responsabilità oggettiva c'è sempre la possibilità da parte di chi ha determinato il sorgere del danno di liberarsi da ogni responsabilità sostenendo che questo è stato determinato da un "caso fortuito" (avvenimento imprevedibile ed eccezionale che si inserisce d'improvviso nella azione del soggetto) o da una "forza maggiore" (una forza esterna che costringa la persona a tenere una determinata condotta senza avere la possibilità di opporsi): in tal caso, sono evidenti le perplessità che potrebbero sorgere se si considera che si tratta di Intelligenze Artificiali che sono autonome nella decisione di come agire. Il che apre una discussione in merito alla possibile estensione del concetto di controllo da parte del produttore.

Ad ogni modo, al Punto 59, lett. f, della Risoluzione invita la Commissione a valutare l'opportunità della «*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*». Il Parlamento europeo, pertanto, ritiene che sia possibile riconoscere tale *status* ai *robot* con capacità cognitive più elevate, in considerazione del fatto che riescono ad interagire in modo indipendente con parti terze: per tali ragioni, si sostiene una loro qualificazione come "persone elettroniche" o "*ePersons*"[69].

Inoltre, occorre ricordare che esiste anche un'altra comunicazione alla Commissione – la "*Building an European Data Economy*"[70] – in cui il Parlamento europeo affronta il tema in modo differente, facendo riferimento ai criteri tradizionali sulla responsabilità per i soggetti che hanno immesso un *robot* nel mercato o che controllano i rischi associati al suo utilizzo, i quali sarebbero direttamente responsabili per qualsiasi danno creato dalla macchina[71].

La *ratio* sottesa all'istituto sembrerebbe quella di fornire certezza legale ai produttori e utilizzatori di *robot* autonomi[72]: se attraverso le "*ePersons*" i produttori fossero maggiormente protetti

---

[68] Cfr. Punto 55 della Risoluzione (2015/2103(INL)), in cui si afferma che il Parlamento «*constata che l'approccio di gestione dei rischi non si concentra sulla persona «che ha agito con negligenza» in quanto responsabile a livello individuale bensì sulla persona che, in determinate circostanze, è in grado di minimizzare i rischi e affrontare l'impatto negativo*».

[69] Si v. G. COMANDÉ, *Intelligenza Artificiale e responsabilità tra* liability *e* accountability. *Il carattere trasformativo dell'IA e il problema della responsabilità*, in *Analisi Giuridica dell'Economia*, vol. 1, 2019, p. 180 e la bibliografia citata alla nota n. 31.

[70] Cfr. *COM (2017) 9 final*.

[71] Si v. Punto 53, 2015/2103(INL): «*ritiene che il futuro strumento legislativo debba essere fondato su una valutazione approfondita della Commissione che stabilisca se applicare l'approccio della responsabilità oggettiva o della gestione dei rischi*».

[72] Si v. *COM (2017) 9 final.*

dalle responsabilità, questo potrebbe costituire un elemento di favore per l'introduzione di nuovi prodotti digitali nel mercato.

Questa proposta dovrà, però, trovare applicazione attraverso la soluzione di problemi concreti ad essa connessi: ad esempio, individuando le modalità più opportune per poter identificare i *robot*[73] o come si possa consentire al *robot* di risarcire effettivamente il danno che ha provocato. In proposito, come si è detto, una soluzione potrebbe essere quella di costituire un apposito fondo[74] da attribuire alle Intelligenze Artificiali persone giuridiche, in modo tale che possano rispondere direttamente delle obbligazioni che sorgono in capo a loro.

3.4.1 Le intelligenze Artificiali più evolute e il c.d. *dilemma situation*. – Altro tema da affrontare riguarda il tema del c.d. "*dilemma situation*" [75]. Quando si tratta di *robot* ad alte capacità cognitive, molti hanno suggerito che le macchine siano dotate di un codice etico che consenta loro di prendere delle decisioni basandosi su valori generalmente riconosciuti che possano controbilanciare e condizionare il *robot* quando si trova di fronte a un problema da risolvere[76]. Sul punto, ci si chiede quali siano i criteri decisionali che l'algoritmo deve applicare quando si trova di fronte a una situazione complessa da decidere. La soluzione al dilemma non è automatica, ma deve poter essere risolta: a maggior ragione se si decide che sia opportuno attribuire personalità giuridica a tali sistemi.

In effetti, da più parti si suggerisce, in generale, la necessità di attribuire un codice etico alle Intelligenze Artificiali, ma non è ancora stato elaborata una modalità attuativa che permetta di codificare i principi morali da adottare per addestrare gli algoritmi al loro rispetto.

Con riferimento all'opportunità di attribuire personalità giuridica alle Intelligenze Artificiali, alcuni Autori[77] hanno evidenziato che, così operando, anche gli algoritmi potrebbero godere di principi costituzionalmente garantiti: per questo motivo, nel definire quale regime giuridico si intende attribuire alle nuove tecnologie, è necessario valutare con estrema attenzione le conseguenze, oltre che da un punto di vista strettamente giuridico, anche sotto il profilo filosofico, morale e politico.

Tuttavia, anche in ambito europeo non tutti sono d'accordo con la proposta avanzata dal Parlamento nel 2017: per esempio, il Comitato Economico e Sociale europeo ha espresso la sua posizione contraria sostenendo che riconoscere le Intelligenze Artificiali come soggetti del diritto sarebbe un «*rischio inaccettabile di azzardo morale. Dal diritto in materia di responsabilità civile deriva una funzione preventiva di correzione del comportamento che potrebbe venir meno una volta che la responsabilità non ricade più sul costruttore perché trasferita sul* robot»[78].

3.4.2 *EPersons*: rilievi critici e prospettive. – Come detto, parte della dottrina è contraria al riconoscimento della personalità giuridica ai *robot*[79].

---

[73] A. BENSOUSSAN, *Plaidoyer pour un droit des robots: de la «personne morale» à la «personne robot»*, in *La lettre des juristes d'affairs*, n. 1134 del 28 ottobre 2013.

[74] E. PALMERINI, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, in *Responsabilità civile e previdenza*, vol. 6, 2016, p. 1835-1836.

[75] M. ANDERSON, S. ANDERSON, *Il buon* robot, in *Le Scienze*, 2010, p. 90-95.

[76] *Id.*.

[77] Si v. S. CHOPRA, L. WHITE, *A Legal Theory, cit.*.

[78] Si v. Parere del Comitato economico e sociale europeo su «*L'Intelligenza Artificiale: Le ricadute dell'Intelligenza Artificiale sul mercato unico (digitale), sulla produzione, sul consumo, sull'occupazione e sulla società*», (2017/C 288/01).

[79] A. AMIDEI, *Politica intelligente e responsabilità: profili e prospettive evolutive del quadro normativo europeo*, in U. RUFFOLO (a cura di), *Intelligenza Artificiale e responsabilità*, Milano, 2017, p. 98

Altra, ancora, sostiene le "*ePersons*" rischiano di essere superflue[80]. È vero, si rileva, che attraverso questo riconoscimento si potrebbero eliminare tutte le complicazioni derivanti dal fatto che le Intelligenze Artificiali sono composte da prodotti digitali *ab origine* disgregati. Essendo l'Intelligenza Artificiale composta da una serie di elementi di fatto separati e poi assemblati tra loro, i terzi danneggiati da un *robot* potrebbero trovarsi di fronte a notevoli difficoltà nella determinazione di quale sia l'elemento assemblato da cui origina il difetto e, conseguentemente, il produttore da ritenere responsabile del danno creato dalla macchina[81]. Come già detto nel paragrafo dedicato al tema, il malfunzionamento del *robot* non chiarisce se questo difetto è da imputare all'*hardware* messo in circolazione da un determinato produttore, o al *software* che è stato scaricato (a sua volta prodotto da un altro soggetto). Allo stesso modo, la ricostruzione della responsabilità dell'utilizzatore potrebbe essere altrettanto complicata.

Pertanto, in un mercato di prodotti disaggregati, la "qualificazione" di un *robot* come soggetto autonomamente responsabile potrebbe servire per "aggregare" i profili di responsabilità e ad attribuire l'onere del risarcimento del danno a un'unica entità. Vi sarebbe una inversione dell'onere della prova per l'identificazione della parte della macchina responsabile del malfunzionamento che ricadrebbe sul *robot*, su tutti i soggetti in esso interessati e, alla fine, sul sistema assicurativo posto a garanzia dei rischi da utilizzo.

Ciò nonostante, questa dottrina sostiene che quando si considera la possibilità di qualificare i *robot* come entità capaci di sopportare la responsabilità della loro condotta, la risposta che ci si dovrebbe dare deve essere negativa. Infatti, si afferma, i *robot* difettano del capitale necessario per estinguere (autonomamente) le obbligazioni derivanti dalle richieste di risarcimento dei danni da essi provocati. Inoltre, se venisse comunque riconosciuta loro una personalità elettronica, la posizione giuridica del danneggiato non subirebbe sostanziali modifiche e il produttore sarebbe, comunque, vincolato a mantenere gli *standard* produttivi indicati dalla normativa europea: in tal senso, alla fine, le "*ePersons*" rischiano di essere superflue.

La qualificazione dei *robot* come enti legali, pertanto, risulterebbe in una esternalizzazione del rischio da parte dei produttori[82]. Per evitare questa situazione si potrebbero ipotizzare due soluzioni alternative.

La prima consiste nel prevedere che il *robot*, per essere qualificato come "*ePerson*" debba essere dotato di un capitale minimo: questo capitale obbligherebbe altri attori a disporre il finanziamento necessario per soddisfare i danni potenziali e le possibili richieste di risarcimento del danno. Il capitale così costituito potrebbe poi essere conferito al *robot* e mantenuto in suo nome, per fare in modo che possa sopportare l'onere di future richieste di risarcimento.

La seconda alternativa, invece, potrebbe assoggettare tale qualificazione del *robot* come "*ePerson*" alla stipula obbligatoria di un'assicurazione. Sotto questo profilo occorre rilevare che la proposta per l'adozione di un regime assicurativo obbligatorio è anche sostenuta dal Parlamento europeo, che suggerisce «*l'istituzione di un numero di immatricolazione individuale, iscritto in un registro specifico dell'Unione, al fine di associare in modo evidente il robot al suo fondo, onde consentire a chiunque interagisce con il robot di essere informato sulla natura del fondo, sui limiti*

---

ss.; F. DI GIOVANNI, *Intelligenza Artificiale e rapporti contrattuali*, in U. RUFFOLO (a cura di), *Intelligenza Artificiale e responsabilità*, *cit.*, p. 127 ss.: si contesta il fatto che l'algoritmo possa ritenersi dotato di volontà propria e che questa caratteristica gli debba essere riconosciuta attraverso l'attribuzione di una soggettività di diritto diversa dall'umano che lo controlla. E. PALMERINI, *Robotica e diritto*, *cit.*.

[80] G. WAGNER, *Robot, Inc.: Personhood for Autonomous Systems?*, in *Fordham Law Review*, vol. 88, n. 2, 2019, p. 608 ss..

[81] *Id.*, p. 602.

[82] *Id.*, p. 609 ss..

della responsabilità in caso di danni alle cose, sui nomi e sulle funzioni dei contributori e su tutte le altre informazioni pertinenti»[83].

Da ultimo, se in futuro si riterrà opportuno riconoscere le Intelligenze Artificiali come soggetti del diritto, è plausibile che non tutte acquisteranno automaticamente questo *status*. Infatti, è verosimile ipotizzare che si renderà necessaria una classificazione dei vari tipi di Intelligenze esistenti, in modo tale da poter trattare ciascuna con un regime giuridico differente a seconda delle effettive capacità.

QUALIFICAZIONE GIURIDICA IA

Non tutte le Intelligenze Artificiali sono uguali: servono soluzioni diverse per esigenze diverse.
Manca una definizione unica applicabile a tutte le IA in generale

CODICE CIVILE	PRODOTTO	AGENTE	DISPOSITIVO MEDICO
Si applica la legislazione di settore (L. Gelli-Bianco → artt. 1218/1228 – art. 2043). Devono sussistere caratteristiche specifiche del *software* affinché la disciplina del Codice civile sia applicabile.	Si applica la Dir. 85/374/CEE, recepita nel Codice del consumo (d.lgs. 205/2006). **Problemi**: 1) il *software* non è per forza *embedded* in un *hardware* (problema nella definizione *ex* art. 2). 2) il *software* potrebbe essere un prodotto o un servizio. 3) definizione "produttore". 4) definizione di "difetto": a. *standard* b. aspettativa legittima del consumatore c. manca *test* per stabilire il difetto d. *Horizon 2020* fa emergere che non esiste un *software* privo di errori. 5) onere della prova sull'attore (art. 4).	Si considera una *fictio iuris* → persona giuridica o Si considera come un essere umano → persona fisica ? 1) esternalizzazione del rischio. 2) necessità di costituire un fondo/capitale minimo. 3) facilità di dimostrazione del danno. 4) facilità di dimostrazione del danno. 5) potrebbero essere riconosciute *e-persons* solo quelle complesse. 6) non tutte le Istituzioni europee sono d'accordo.	Dir. 93/42/CEE + Reg. 745/2017. 1) *software* è incluso nella definizione; 2) "*stand-alone software*" deve rispondere a una serie di requisiti; 3) ogni *software* deve essere inserito in una classe e l'*iter* per l'ottenimento del marchio CE è diverso. 4) disciplina di codici di identificazione UDI-DI e UDI-PI. 5) *software* non definibili come SaMD devono comunque rispettare dei criteri fissati dalle Linee guida pubblicate dalla Commissione europea.

4. Le difficoltà nella ricostruzione del nesso causale: "*in dubio pro machina*". – Precedentemente, si è fatto cenno alle difficoltà esistenti in materia di Intelligenza Artificiale a ricostruire, oggettivamente, il nesso teleologico che collega il danno alla condotta posta in essere. Pur non essendo possibile in questa trattazione affrontare in modo compiuto il tema della causalità soprattutto in un ambito così particolare, ciò che rileva qui evidenziare è che in questo settore la complessità assume una tale rilevanza – in generale e, in special modo nelle sue forme più evolute – da mettere in crisi le regole giuridiche fondamentali per impostare i giudizi di responsabilità.

[83] Si v. Punto 59, lett. e), 2015/2103(INL).

Nel l'ordinamento giuridico italiano si può fare riferimento all'art. 1223 cod. civ. che, in materia di risarcimento del danno, stabilisce «*Il risarcimento del danno per l'inadempimento o per il ritardo deve comprendere così la perdita subita dal creditore come il mancato guadagno, in quanto ne siano conseguenza immediata e diretta*». Ne consegue che, nel disciplinare la relazione esistente tra gli eventi accaduti e il danno subito, questa norma conferisce rilevanza ai soli i danni subiti dal creditore che siano "*diretta e immediata conseguenza*" dell'inadempimento o del ritardo: requisiti che hanno trovato chiarificazione nell'orientamento giurisprudenziale che, per stabilire quando una conseguenza possa essere ritenuta "*immediata*" e "*diretta*" dell'illecito, fanno riferimento ai criteri di "*normalità*" e, soprattutto, di "*prevedibilità*".

Quest'ultimo criterio lo si ritrova sia negli ordinamenti europei che negli Stati Uniti. Il concetto di "prevedibilità" è suscettibile di cambiamenti nel tempo ed è condizionato dal campo concreto di indagine: come detto precedentemente, mentre per le Intelligenze Artificiali più semplici il rapporto esistente fra il comportamento dell'uomo (produttore/programmatore) e danno creato dalla macchina è possibile (pur con notevoli difficoltà) definire se quello specifico danno era "prevedibile", con gli algoritmi più complessi tale accertamento (in merito alle ragioni per cui la macchina ha preso una determinata decisione e/o ha assunto quella specifica condotta), potrebbe risultare non prevedibile né dal creatore dell'algoritmo né dal suo utilizzatore[84].

Una complessità di accertamento tale da qualificare questo tipo di algoritmi come delle "*Black Box*"[85]. Le conseguenze sono evidenti. Come detto, affinché il produttore o il programmatore della macchina possa rispondere del danno arrecato, occorre sia dimostrato che al momento della sua costruzione, o del suo allenamento, un produttore medio o un programmatore medio avrebbe potuto prevedere che si sarebbe potuto verificare un danno e che avrebbe potuto evitarne il verificarsi. In tal caso, tutto entra in crisi in quanto nessuno – nemmeno colui che ha creato l'Intelligenza Artificiale o chi l'ha programmata – è in grado di prevederne i comportamenti e gli effetti.

Quanto esposto evidenzia come, nei confronti delle forme più evolute (e che si evolveranno sempre di più soprattutto nella autonomia decisionale) di Intelligenza Artificiale, la *ratio* della causalità legale entra in crisi: si incrina l'idea secondo cui, la ragione fondante della responsabilità dev'essere ravvisabile nel tipo di misure precauzionali adottate (e da adottare) o nella valutazione del rischio della macchina che la legge si aspetta (ed esige) da una persona mediamente ragionevole che svolge quella attività.

Tuttavia, queste considerazioni rilevano quando si deve valutare la responsabilità del produttore che ha messo in commercio il dispositivo che ha provocato il danno: dimostrare la sua responsabilità sarà più o meno semplice a seconda che il sistema di Intelligenza Artificiale sia più o meno opaco.

---

[84] Si è avuto modo di rilevare come all'interno della categoria "Intelligenze Artificiali" vi siano realtà differenti, con livelli di complessità e di curve di apprendimento fra loro incomparabili. Ci sono sistemi semplici che definiscono le proprie decisioni sulla base di regole che sono state per loro predefinite dal creatore: in tal caso, l'individuazione della correlazione esistente fra il danno e la condotta tenuta dalla macchina è più semplice. Ma, come detto, vi sono anche algoritmi più sofisticati, connotati di autonomo apprendimento dai dati e non da regole predefinite a cui fare riferimento per risolvere uno specifico quesito, ma solo istruzioni su come utilizzare i dati per migliorare le proprie capacità di apprendimento: sono questi i sistemi di Intelligenza Artificiale che creano i maggiori problemi sotto il profilo della ricostruzione del collegamento causale; problemi, ulteriormente, aggravati dalla carenza di trasparenza sui meccanismi che li caratterizzano.

[85] L'algoritmo può risultare talmente complesso da rendere molto difficile (per non dire impossibile) chiarire oggettivamente ogni passaggio svolto dalla macchina nell'esprimere la propria abilità tecnica o, in altre ipotesi, l'incomprensibilità è data da mere esigenze di protezione di segreti industriali: il risultato finale è che nessuno è in grado di comprendere come venga presa una decisione o una predizione, sulla base dei dati raccolti, all'interno di quella Intelligenza Artificiale.

Quando si parla di responsabilità del professionista sanitario, invece, bisogna prendere in considerazione la stretta relazione tra il professionista, la macchina e il paziente. In questo caso, risultando difficile applicare la dottrina che riconoscerebbe personalità elettronica alle Intelligenze Artificiali (in quanto quelle presenti sul mercato non sono ancora sufficientemente autonome), il dispositivo sarà inteso come uno strumento nelle mani, nel nostro caso, del TSRM che sta conducendo l'esame.

Si immagini una situazione in cui il tecnico sanitario di radiologia medica utilizzi una macchina dotata di un sistema di Intelligenza Artificiale che, previa l'analisi di specifici parametri e caratteristiche del paziente (peso, altezza, età, etnia, GFR, comorbilità), raccomandi un determinato e specifico protocollo di somministrazione del mezzo di contrasto. In questo caso, il TSRM avrà di fronte a sé due opzioni: nel primo caso, seguire quanto suggerito dalla macchina; nel secondo, discostarsi da quella indicazione e fare riferimento al medico radiologo per una diversa prescrizione.

Cominciando dalla prima ipotesi, si immagini che la macchina abbia commesso un errore e che, di conseguenza, si produca un danno.

Il difetto della macchina può essere conseguenza dell'errato inserimento dei dati da parte del TSRM. In tal caso operano le regole generali sulla responsabilità professionale del professionista sanitario ex L. 24 del 2017 (nota come "Legge Gelli-Bianco"). Se il TSRM non aveva alcun rapporto contrattuale diretto con il paziente (in quanto operante a qualsiasi titolo all'interno di una struttura sanitaria pubblica o privata), del danno da lui arrecato risponderà la Struttura sanitaria presso cui opera. Il TSRM potrà essere chiamato in causa direttamente dal paziente per responsabilità extracontrattuale (ex art. 7, terzo comma, L.24/17) – o dalla sua azienda "a manleva" nei confronti della richiesta ricevuta dal paziente – oppure, potrà essere soggetto ad una azione di "rivalsa" o di responsabilità contabile amministrativa (avanti alla Corte dei conti): in ogni caso, risponderà del danno arrecato qualora venga accertato giudizialmente che la sua condotta è connotata da "dolo" (volontà di recare danno) o da "colpa grave" (grave negligenza nello svolgere la sua attività).

Diverso è il caso in cui il TSRM abbia svolto correttamente (con scienza e coscienza) la propria attività professionale: in tale ipotesi si potrebbe presumere che il *software* utilizzato presenti un difetto di programmazione, di addestramento o di manutenzione. In questi casi, del danno al paziente risponderà direttamente (a titolo "contrattuale": *ex* art. 7, primo comma, L. 24/17) la struttura sanitaria la quale potrà esimersi dal risarcimento del danno qualora riesca a dimostrare l'esistenza dl difetto della macchina. Se riesce a dimostrare l'esistenza del difetto: allora, la Struttura sanitaria che usufruisce della macchina potrà chiamare nel giudizio instaurato dal paziente – o rivalersi in un giudizio successivo al risarcimento effettuato – il produttore che ha messo in commercio il prodotto difettoso (ai sensi della direttiva 85/374/CEE).

Se, però, nonostante sia presente, il difetto risulti non dimostrabile[86], tutto diviene molto più complesso soprattutto per la Struttura sanitaria che usufruisce della macchina la quale sarà tenuta al risarcimento del danno esistente, ma faticherà a rivalersi nei confronti del produttore/addestratore.

---

[86] Un'altra ragione per cui si ritiene opportuna una modifica della disciplina è data dall'alta percentuale di fallimento delle cause avviate dai consumatori, per impossibilità di dimostrazione del difetto (nel 32% dei casi), o del nesso causale (21% dei casi): il 53% delle domande di risarcimento viene rigettato perché l'attore non è riuscito a provare il fatto o il nesso eziologico. Si v. in merito *SWD (2018) 157 final, cit.*, p. 67.

Tuttavia, nonostante la Direttiva che regola la responsabilità del produttore non ammetta l'inversione dell'onere della prova, la Corte di Giustizia talvolta ammette un'estensione della nozione di "difetto". Per esempio, si v. *Boston Scientific Medizintechnik GmbH* contro *AOK Sachsen-Anhalt – Die Gesundheitskasse (C-503/13)* e *Betriebskrankenkasse RWE (C-504/13).* Altre volte, ha ritenuto compatibili con la Direttiva le disposizioni nazionali che imponevano al produttore, a fronte della richiesta del consumatore, di fornire ogni elemento necessario su possibili conseguenze negative del prodotto (Si v. *Novo Nordisk Pharma GmbH v. S.* (*C-310/13*), 20/11/2014).

Un'alternativa si ha nel caso in cui la macchina abbia proposto un protocollo – poi rivelatosi dannoso – ma il *software* non sia, di per sé, difettoso. Si tratta, in questo caso, di ipotesi meramente marginali, trattandosi di *software* dotati di una capacità di apprendimento fondata sulla loro esperienza. Anche qui, la Struttura che usufruisce della macchina sarà responsabile del danno rispetto al paziente, ma potrà rivalersi non verso il produttore (in quanto il *software* non risulta difettoso) bensì nei confronti del programmatore/addestratore se riesce a dimostrare che il danno è conseguenza di un errore compiuto nel corso dell'addestramento della macchina.

In tutti questi casi, in cui il TSRM ha seguito le indicazioni fornite dalla macchina, questi sarà comunque ritenuto sempre responsabile se il danno arrecato risulta essere frutto di un errore grossolano da essere prevedibile ed evitabile da parte di qualsiasi altro tecnico sanitario di radiologia: secondo la formula dello "*homo eiusdem professionis et condicionis*" quale agente modello per valutare i presupposti di "prevedibilità" ed "evitabilità" dell'evento per accertare la sussistenza della colpa.

Altra ipotesi riguarda il caso in cui il TSRM, a fronte del protocollo suggerito dalla macchina, decide di rivolgersi al medico radiologo che, a sua volta, opta per la prescrizione di una dose di mezzo di contrasto differente rispetto a quella raccomandata dalla macchina[87]. In questa situazione, si determina sempre una sostanziale inversione dell'onere della prova: sarà il professionista a dover giustificare la propria condotta, spiegando le ragioni per cui ha deciso di discostarsi dalla indicazione fornita dalla macchina.

Se con tale condotta si evita il verificarsi del danno (che sarebbe stato provocato seguendo l'indicazione), sorge in capo al professionista un obbligo di comunicazione del fatto alla struttura sanitaria che utilizza la macchina in modo che si possano evitare eventuali danni futuri: se omette la segnalazione potrebbe risponderne lui degli eventuali danni futuri.

Se, invece, il medico radiologo che si è discostato dalla indicazione fornita dalla macchina, somministra una dose di mezzo di contrasto differente e determina un danno al paziente, questi sarà responsabile del danno arrecato. Unica via per esimersi da questa responsabilità sarà riuscire a dimostrare «*che una causa imprevedibile ed inevitabile ha reso impossibile l'esatta esecuzione della prestazione*»[88]: prova non facile considerando che ha volontariamente disatteso l'indicazione (esatta) fornita dalla macchina.

Ciò che emerge dalle considerazioni sopra esposte è che si crei di fatto un principio per cui "*in dubio, pro machina*" (sul punto, si v. oltre Par. 6) che renderebbe molto difficile per il professionista distanziarsi dai "suggerimenti" proposti dal *software* con cui si trova a lavorare.

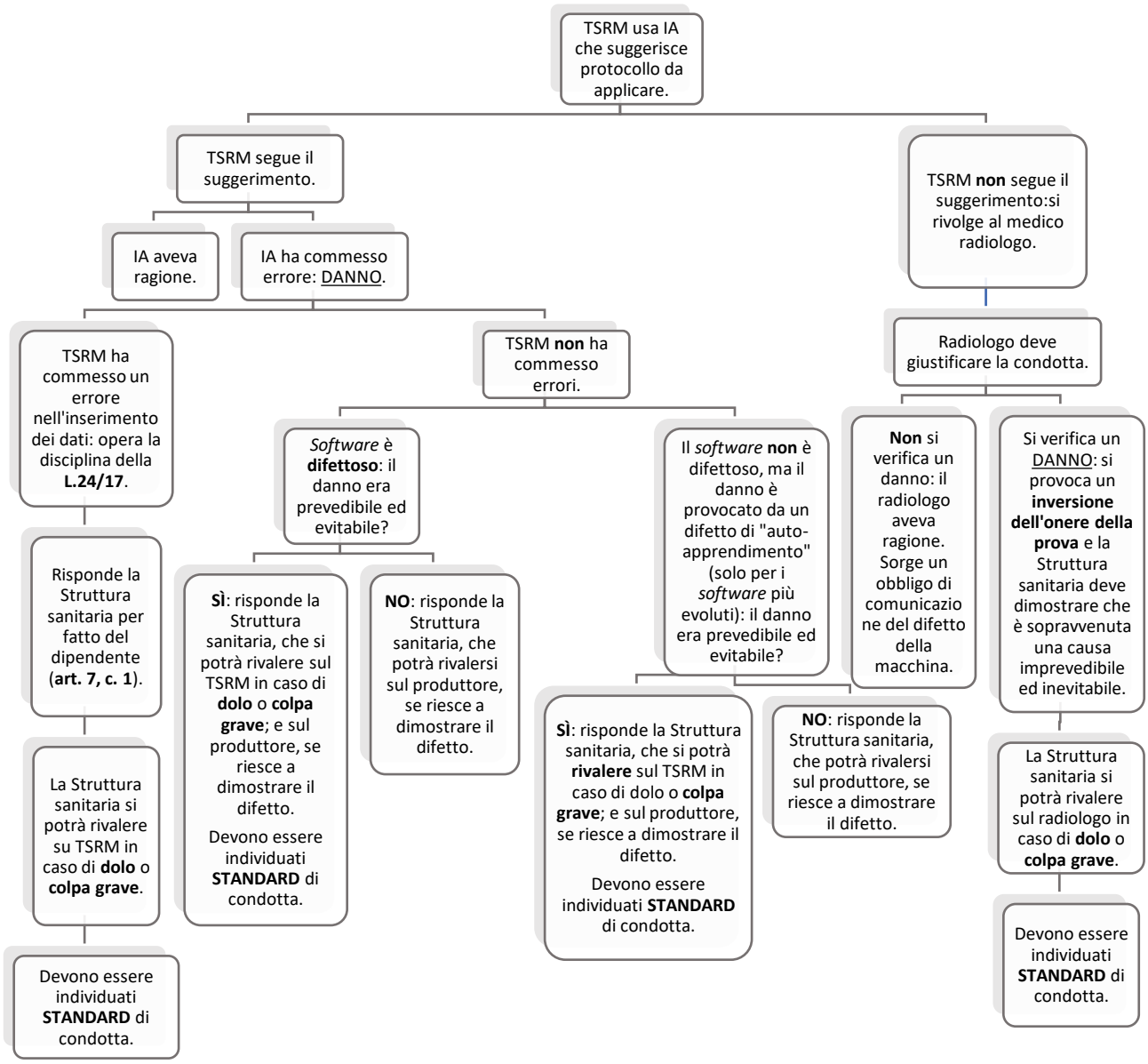
Allora, a questo punto, come si può costruire il rapporto tra l'essere umano e l'Intelligenza Artificiale?

Ancora, in alcune occasioni la Corte ha ammesso che alcune prove di fatto possono essere considerate dalle Corti nazionali come costituenti serie, specifiche e consistenti prove di un difetto e del suo nesso causale con il danno, anche in mancanza di prove scientifiche concludenti. La Corte stessa, tuttavia, ha affermato che queste ammissioni non devono essere interpretate come una modalità di inversione dell'onere della prova (Si v. *C-621/15*).

[87] Al fine di una esposizione più completa, è necessario precisare che la quantità di mezzo di contrasto è solitamente legata al tipo di esame che deve essere effettuato. Tuttavia, essendo un farmaco, la prescrizione della posologia è prettamente medica. Gli attori coinvolti in questo processo sono: l'infermiere, che canula la vena, il TSRM, che prepara l'iniettore, e il medico responsabile del volume (calcolato in ml) e del flusso (calcolato in ml/sec), che coordina l'iniezione. In tal senso, si v. http://www.tsrm.org/wp-content/uploads/2020/07/Parere-Emodinamica-TSRM.pdf.

[88] Si v. Cass., 11/11/19, n. 28991.

Di certo, sembra necessario l'individuazione di *standard* di condotta che permettano di rendere la prova in giudizio più semplice (v. schema 1).



Schema 1.

5. I dati trattati ed utilizzati nell'addestramento delle Intelligenze Artificiali. – Quando si ipotizza l'utilizzo di sistemi di Intelligenza Artificiale, un aspetto rilevante riguarda il trattamento dei dati che servono per l'addestramento degli algoritmi.

La questione è rilevante per due ordini di ragioni: in primo luogo, in quanto sia il *software* necessita di una grande quantità di dati, sia il tipo di dati di cui si tratta in ambito sanitario è classificato come "dato particolare". In secondo luogo, perché nella ricerca scientifica e nella sua

applicazione al settore salute l'innovazione è da sempre evidente: in questo caso, più di altri si impara osservando e testando nuove soluzioni condividendo dati ed esperienze.

Tuttavia, è tutt'ora in corso un dibattito sull'equilibrio che dovrebbe trovarsi tra il diritto alla *privacy* e l'utilizzo dei dati per il miglioramento delle esperienze dell'utilizzatore e il migliore addestramento dei sistemi di Intelligenza Artificiale. Questi ultimi, come detto, richiedono un'enorme quantità di dati e *server* molto potenti, e di solito necessitano un lungo periodo di tempo per essere addestrati, in quanto bisogna prendere in considerazione un numero elevato di parametri. Pertanto, l'assenza di ben documentati *dataset* per l'addestramento degli algoritmi rappresenta uno degli ostacoli principali a una più importante introduzione di tali sistemi in ambito radiologico. L'accesso a *big data* di *imaging* medico, infatti, sarebbe necessario per offrire materiale per l'addestramento di Intelligenze Artificiali, in modo tale che imparino sempre di più a riconoscere anomalie di *imaging*.

Il problema principale è che tali dati sensibili potrebbero essere raccolti in modo illecito, oppure ottenuti da fonti sconosciute, a causa di mancanza di chiare e precise disposizioni in merito.

Inoltre, si pone un'altra questione, relativa alla *cyber*sicurezza: se, da un lato, i dati dovranno essere raccolti e condivisi, in modo tale che l'Intelligenza Artificiale li possa analizzare, dall'altro si renderà sempre più necessario una re-definizione del concetto di "riservatezza" e di altri principi etici fondamentali.

Il Regolamento 2016/679/UE[89] – noto come *General Data Protection Regulation* ("GDPR") – rappresenta il riferimento legislativo in materia. Dalla sua entrata in vigore è stata abrogata la Direttiva 95/46/CE[90], che perseguiva l'obiettivo di armonizzare la protezione dei diritti fondamentali e delle libertà delle persone fisiche con riferimento alle attività di trattamento dei dati, cercando, contemporaneamente, di assicurare la libera circolazione di dati personali tra gli Stati membri.

Nei *Considerando* del Regolamento 2016/679/UE[91] viene approfondita la ragione per cui si è reso necessario un mutamento di disciplina. In tal senso si evidenzia che i vari Stati membri hanno recepito la Direttiva 95/46/CE con modalità molto diverse tra loro, favorendo l'incertezza del diritto in merito alla protezione dei dati, e diffondendo la convinzione che l'attività *online* fosse rischiosa e non tutelata.

L'Unione, pertanto, ritiene che le differenze nelle discipline che assicurano la protezione dei dati personali potrebbero costituire un ostacolo allo sviluppo dell'economia digitale, alterare la concorrenza nel mercato e impedire alle autorità di liberarsi dalla responsabilità facendo riferimento alla disciplina dell'Unione. Di conseguenza, per assicurare un più alto livello di protezione delle persone fisiche e rimuovere gli ostacoli alla libera circolazione dei dati, è necessario che venga adottato un provvedimento che sia direttamente applicabile e che uniformi la disciplina in tutti gli Stati. Ad ogni modo, si richiede comunque che i legislatori nazionali adottino delle disposizioni che consentano di specificare ulteriormente la disciplina: considerato che l'art. 9 del GDPR[92] classifica rigidamente le categorie di "*dati sensibili*", rendendo inammissibili discipline diverse nei vari Stati Membri, l'aspetto problematico riguarda le regole che gli stessi adottano relativamente al trattamento di questi dati, come ben espresso dal *Considerando 10* del Regolamento[93].

---

[89] Si v. Reg. 2016/679/UE.

[90] Si v. dir. 95/46/CE.

[91] Reg. 2016/679/UE, *Considerando* 9 e 10.

[92] Art. 9, Reg. 2016/679/UE.

[93] *Considerando* 10, Reg. 2016/679/UE: «*Il presente regolamento prevede anche un margine di manovra degli Stati membri per precisarne le norme, anche con riguardo al trattamento di categorie particolari di dati personali*». In tal senso, il presente Regolamento non esclude che il diritto degli Stati Membri stabilisca le condizioni per specifiche situazioni di trattamento, anche determinando con maggiore precisione le condizioni alle quali il trattamento di dati personali è "lecito". Ad esempio, il Codice della *Privacy* italiano, novellato dal d.lgs. 101/2018, prevede una deroga alle regole stabilite dall'art. 9 del GDPR, attraverso,

Anche la *ratio* sottesa al nuovo Regolamento rispecchia l'esigenza di individuare un equilibrio tra il bisogno del *software* di ricevere sempre maggiori informazioni e il rischio di pregiudizio che ne è collegato. L'obiettivo del GDPR, pertanto, segue due linee direttrici: da un lato, mantiene ferma l'esigenza di tutela dei diritti fondamentali, nel rispetto dell'art. 8 della Carta dei Diritti Fondamentali dell'Unione Europea[94]; dall'altro, con esso l'Unione esprime la volontà di incentivare lo sviluppo dell'economia digitale garantendo la libera circolazione dei dati personali, come indicato nell'articolo 1 del GDPR. Pertanto, il Regolamento si presenta come una disciplina flessibile idonea a conferire dinamicità al ruolo svolto dalla tutela dei dati personali con riferimento alla continua evoluzione cui sono soggette le Intelligenze Artificiali.

A questo proposito, i *Considerando* 6 e 7 del Regolamento, enunciano le ragioni di carattere strettamente politico che hanno determinato l'adozione del provvedimento.

Partendo dal *Considerando* 6, si può fare riferimento alla parte in cui evidenzia che «*la tecnologia ha trasformato l'economia e le relazioni sociali e dovrebbe facilitare ancora di più la libera circolazione dei dati personali all'interno dell'Unione e il loro trasferimento verso Paesi terzi e organizzazioni internazionali, garantendo al tempo stesso un elevato livello di protezione dei dati personali*».

E il successivo *Considerando* 7, ricorda: «*tale evoluzione richiede un quadro più solido e coerente in materia di protezione dei dati nell'Unione, affiancato da efficaci misure di attuazione, data l'importanza di creare il clima di fiducia che consentirà lo sviluppo dell'economia digitale in tutto il mercato interno*».

Queste indicazioni rientrano pienamente nella strategia europea di creare un'Intelligenza Artificiale "*Made in Europe*" e di rendere l'Unione più competitiva a livello mondiale. Se non si incentivasse il più possibile la circolazione dei dati, il raggiungimento dell'obiettivo sarebbe decisamente più complicato.

Alla luce del Regolamento 2016/679/UE, sono "dati particolari" i dati personali che rivelino l'origine razziale o etnica, le opinioni politiche, le convinzioni religiose o filosofiche, o l'appartenenza sindacale, nonché i dati genetici, dati biometrici intesi a identificare in modo univoco una persona fisica, dati relativi alla salute o alla vita sessuale o all'orientamento sessuale della persona.

Il GDPR, all'art. 9 ne vieta il trattamento, a meno che non ricorra uno dei casi indicati al comma secondo dello stesso articolo, vale a dire: «*a) l'interessato ha prestato il proprio consenso esplicito al trattamento di tali dati personali per una o più finalità specifiche, salvo nei casi in cui il diritto dell'Unione o degli Stati membri dispone che l'interessato non possa revocare il divieto di cui al paragrafo 1;*

b) il trattamento è necessario per assolvere gli obblighi ed esercitare i diritti specifici del titolare del trattamento o dell'interessato in materia di diritto del lavoro e della sicurezza sociale e protezione sociale, nella misura in cui sia autorizzato dal diritto dell'Unione o degli Stati membri o da un contratto collettivo ai sensi del diritto degli Stati membri, in presenza di garanzie appropriate per i diritti fondamentali e gli interessi dell'interessato;

c) il trattamento è necessario per tutelare un interesse vitale dell'interessato o di un'altra persona fisica qualora l'interessato si trovi nell'incapacità fisica o giuridica di prestare il proprio consenso;

d) il trattamento è effettuato, nell'ambito delle sue legittime attività e con adeguate garanzie, da una fondazione, associazione o altro organismo senza scopo di lucro che persegua finalità

---

ad esempio, l'art. Art. 2-*sexies*, rubricato «*Trattamento di categorie particolari di dati personali necessario per motivi di interesse pubblico rilevante*».

[94] Si v. art. 8, Carta dei diritti Fondamentali dell'Unione europea.

politiche, filosofiche, religiose o sindacali, a condizione che il trattamento riguardi unicamente i membri, gli ex membri o le persone che hanno regolari contatti con la fondazione, l'associazione o l'organismo a motivo delle sue finalità e che i dati personali non siano comunicati all'esterno senza il consenso dell'interessato;

e) il trattamento riguarda dati personali resi manifestamente pubblici dall'interessato;

f) il trattamento è necessario per accertare, esercitare o difendere un diritto in sede giudiziaria o ogniqualvolta le autorità giurisdizionali esercitino le loro funzioni giurisdizionali;

g) il trattamento è necessario per motivi di interesse pubblico rilevante sulla base del diritto dell'Unione o degli Stati membri, che deve essere proporzionato alla finalità perseguita, rispettare l'essenza del diritto alla protezione dei dati e prevedere misure appropriate e specifiche per tutelare i diritti fondamentali e gli interessi dell'interessato;

h) il trattamento è necessario per finalità di medicina preventiva o di medicina del lavoro, valutazione della capacità lavorativa del dipendente, diagnosi, assistenza o terapia sanitaria o sociale ovvero gestione dei sistemi e servizi sanitari o sociali sulla base del diritto dell'Unione o degli Stati membri o conformemente al contratto con un professionista della sanità, fatte salve le condizioni e le garanzie di cui al paragrafo 3;

i) il trattamento è necessario per motivi di interesse pubblico nel settore della sanità pubblica, quali la protezione da gravi minacce per la salute a carattere transfrontaliero o la garanzia di parametri elevati di qualità e sicurezza dell'assistenza sanitaria e dei medicinali e dei dispositivi medici, sulla base del diritto dell'Unione o degli Stati membri che prevede misure appropriate e specifiche per tutelare i diritti e le libertà dell'interessato, in particolare il segreto professionale;

j) il trattamento è necessario a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici in conformità dell'art. 89, par. 1, sulla base del diritto dell'Unione o nazionale, che è proporzionato alla finalità perseguita, rispetta l'essenza del diritto alla protezione dei dati e prevede misure appropriate e specifiche per tutelare i diritti fondamentali e gli interessi dell'interessato»[95].

Il patrimonio a cui gli scienziati fanno riferimento è tipicamente informativo, con vocaboli e linguaggi propri e i dati rilevati servono anche per individuare nuovi Percorsi Diagnostici, Terapeutici ed Assistenziali (c.d. PDTA) che favoriscano la scoperta ed il trattamento di patologie per un grande numero di pazienti o anche per particolari individui soggetti a malattie rare. Tuttavia, per quanto sia nobile questo fine, esso si scontra con una pericolosa tendenza all'uso improprio delle informazioni ricavate dai dati personali, che una volta che siano raccolti in grande quantità (*Big Data*) e strutturati (*Structured Data*) secondo adeguati criteri di ricerca possono essere profilati, ovvero organizzati in modo da diventare importanti anche per fini tendenti all'interesse privato di qualche soggetto economico non necessariamente coincidente con quello di chi è in cura o vi presta assistenza.

Si pensi, per esempio, al vantaggio che ne trarrebbe una società che si occupa di profilazione a fini pubblicitari.

Questo tipo di dati, infatti, ha un valore non solo di tipo informativo, ma anche di tipo economico: molti soggetti sono interessati a *dataset* di questo genere, e sono anche disposti a corrispondere un'ingente somma di denaro. Per fare un esempio, si può fare riferimento alla vicenda che ha fatto molto discutere in Inghilterra, vale a dire il caso che ha coinvolto il Sistema Sanitario Nazionale britannico (NHS) che avrebbe venduto milioni di dati ad aziende farmaceutiche allo scopo di fare ricerca. L'inchiesta è nata all'inizio del mese di febbraio 2020, ed è stata condotta dalla Rivista "*The Guardian*": i dati forniti erano anonimi, ma i pazienti avrebbero potuto essere identificati e rintracciati attraverso i riferimenti dei loro medici di base. Un procedimento che sarebbe stato

[95] Art. 9, c. 2, Reg. 2017/745/UE.

effettuato proprio dalle case farmaceutiche interessate a rintracciare alcuni pazienti che presentavano casi clinici di interesse. Alcune fonti rivelano come questa vendita di dati abbia fruttato al Governo inglese oltre dieci milioni di sterline. È rilevante, in questo caso, determinare chiaramente quale sia il confine tra *privacy*, entrate pubbliche e l'importanza della ricerca scientifica[96].

In Italia, nel corso degli ultimi anni, già prima dell'adozione del GDPR il Ministero della Salute e il Garante per la Protezione dei Dati personali hanno a più riprese tratteggiato scenari e vincoli riguardanti la conservazione e il trattamento dei dati clinici e sanitari del paziente. Inoltre, anche le singole Regioni, hanno emesso normative aggiuntive, in particolare con riferimento alla visibilità dei documenti nel Fascicolo Sanitario a livello regionale e tentando di individuare anche norme tecniche di interoperabilità tra i diversi sistemi esistenti.

Per questo motivo, sul mercato si sono sviluppate diverse soluzioni informatiche, sovente trasversali alle funzioni clinico-sanitarie e interfacciate con applicativi esterni e regionali. In effetti in questo scenario, la protezione dei dati personali nel settore salute, è particolarmente critico e complesso, in quanto è necessario adeguare i flussi di lavoro alle linee guida e normative regionali, ma anche continuare a garantire ad ogni utente autorizzato la visibilità dei dati clinici che assicurano un appropriato ed ottimale trattamento del paziente. Le linee guida normative, al di là delle deroghe e ridefinizioni locali, si basano essenzialmente su alcuni concetti fondamentali:

- in primo luogo, la raccolta del consenso del paziente alla raccolta dei documenti clinici, che deve essere registrato in maniera chiara e integrata con sistemi terzi;
- il paziente deve poter esercitare il proprio diritto all'oscuramento di episodi e singoli documenti sanitari;
- ogni operatore sanitario deve avere piena visibilità dei dati clinici del paziente che ha in carico – fatti salvi i dati oscurati – per il periodo durante il quale si articola la cura: pertanto, terminato tale periodo, l'operatore non vi avrà più accesso.
- devono essere resi disponibili percorsi di accessi ai dati che consentano di aggirare i vincoli normativi, rendendo disponibile all'operatore il dossier di pazienti che non siano attualmente in carico. Ogni accesso di questo tipo dovrà però obbligatoriamente essere motivato, e la motivazione registrata sul sistema;
- infine, il sistema deve garantire i principi fondamentali di sicurezza e riservatezza.

Ancora, alcune soluzioni presenti sul mercato mettono a disposizione una serie di strumenti che consentono di rispondere alle esigenze normative in maniera flessibile e modulabile sulle esigenze specifiche. Ciò accade perché la normativa lascia margini di interpretabilità: inoltre – essendo diffuso che le Aziende Sanitarie definiscano le loro proprie regole in deroga – molte delle funzioni relative alla protezione dei dati devono poter essere parametrizzate localmente, per ottenere il miglior equilibrio possibile tra l'aderenza alla normativa e le esigenze locali e cliniche di visibilità dei dati del paziente.

Un ulteriore elemento da prendere in considerazione riguarda il fatto che il GDPR richiede che tutti i dati personali siano "anonimizzati" o "pseudonomizzati" prima di essere processati. L'anonimizzazione è quel processo che prevede la modifica irreversibile del dato, in modo da impedire l'identificazione. La pseudonimizzazione, invece, comporta che le informazioni che permetterebbero di riconoscere il paziente siano rimosse, ma conservate in un *database* sicuro, in modo tale che possano essere recuperate in qualsiasi momento.

Per l'utilizzo delle Intelligenze Artificiali sarebbe necessario determinare con assoluta chiarezza quali siano i dati che devono essere anonimizzati e quali, invece, potrebbero contenere

[96] Per maggiori informazioni, si v. https://www.theguardian.com/world/2020/apr/12/uk-government-using-confidential-patient-data-in-coronavirus-response.

informazioni molto importanti e, per tale ragione, dovrebbero essere comunque custoditi. Per esempio, difficilmente la data di nascita, il nome e il codice fiscale di un paziente potrebbero essere d'aiuto. Invece, altri elementi come l'età, il genere, la etnia e le comorbidità potrebbero essere importanti per un'analisi più approfondita dell'immagine.

In realtà, se si prendesse in considerazione un algoritmo di *deep learning*, ci si renderebbe conto che l'indicazione del cognome potrebbe essere utile al *software* per l'individuazione dell'etnia, e questa informazione sarebbe d'aiuto per migliorare l'accuratezza della diagnosi di quello specifico gruppo di pazienti. Tuttavia, nonostante il vantaggio che se ricaverebbe, non si può negare la presenza del rischio di *bias* e di violazioni della riservatezza.

Alcuni[97] sostengono che per ottenere un sistema di predizioni perfettamente informato, sarebbe necessario che gli *outcomes* siano collegati a una storia del paziente. Infatti, l'utilizzo della pseudonimizzazione – o della non-anonimizzazione – consentirebbe un addestramento impareggiabile della tecnologia. Tuttavia, porrebbe il rischio di una maggiore vulnerabilità. Per esempio, una TAC al volto di un paziente, potrebbe essere ricostruita per creare una superficie di immagini renderizzate che, se date a un *software* di riconoscimento facciale, consentirebbe l'identificazione dell'individuo[98].

Per evitare che tali situazioni possano verificarsi, è necessario che i detentori dei dati siano dotati di sistemi di *cyber*sicurezza molto sofisticati, così che gli accessi ai dati dei pazienti effettuati dalle società private siano tracciabili in ogni momento e che i dati siano resi invulnerabili a una successiva manipolazione. Una soluzione alla necessità di un più elevato livello di riservatezza potrebbe essere rappresentata dalla tecnologia di *block-chain*[99].

Un ultimo problema che merita di essere affrontato riguarda la titolarità del diritto di proprietà sulle immagini acquisite dalla macchina. Infatti, la situazione è diversa, a seconda che si consideri titolare l'Azienda Sanitaria o il paziente, in quanto nel secondo caso l'uso e la diffusione delle immagini a fini di ricerca e, soprattutto, addestramento dei *software* sarebbe complicato dalla necessità di ottenere l'approvazione da parte del paziente. La sfida è chiaramente quella di trovare un bilanciamento tra l'utilizzo efficace dell'Intelligenza Artificiale a servizio dei pazienti e il rispetto del loro diritto alla riservatezza, dando a questi ultimi la possibilità di esprimere il proprio consenso informato al trattamento di dati da parte dei sistemi intelligenti.

Sicuramente, la conservazione dei documenti è un obbligo delle Aziende Sanitarie e la normativa[100] disciplina le modalità e i tempi di conservazione. Infatti, ci sono dei documenti che devono essere conservati per un tempo illimitato e altri che, invece, hanno un termine (40 anni, 30 anni, 20 anni, 10 anni).

Con riferimento ai dati sanitari, è necessario precisare che la conservazione di questi si differenzia dalla custodia dei documenti sanitari. I tempi di conservazione dei documenti sanitari, così come riportati nel prontuario, sono inoppugnabili. Quindi, anche i dati personali in esso contenuti andranno necessariamente conservati per i tempi indicati.

---

[97] M. KOHLI, R. GEIS, *Ethics, Artificial Intelligence, and Radiology*, in *Journal of the American College of Radiology*, vol. 15, 2018, p. 1317.

[98] *Id.*.

[99] R. JAMES, *Creating an Immutable Audit Trail on the Blockchain with Xero & Tierion*, disponibile su: https://devblog.xero.com/creating-an-immutable-audit-trail-on-the-blockchain-with-xero-tierion-be423d39380b.

[100] Circolare del Ministero della sanità del 19 dicembre 1986 n. 900; Art. 5, d.m. 18 febbraio 1982; Art. 4, d.m. 14 febbraio 1997.

In tutti gli altri casi, vige quanto stabilito nell'art. 5, c.1, lett. e) del GDPR: chi richiede il permesso di trattare i dati personali deve specificare anche per quanto tempo ha bisogno di tali dati, e si impegna a cancellarli in maniera definitiva alla decorrenza del termine.

Il Garante per la protezione dei dati personali ha chiarito che i professionisti sanitari che sono soggetti al segreto professionale possono trattare i dati relativi alla salute delle persone – necessari per l'esecuzione della prestazione sanitaria richiesta dall'interessato – anche senza il consenso di quest'ultimo, sia che essi lavorino all'interno di una struttura sanitaria pubblica o privata, sia che essi siano dei liberi professionisti.

In secondo luogo, il Garante ha chiarito che tale regola vale per i trattamenti di dati personali relativi alla salute che servono per raggiungere una specifica finalità e che sono necessari per la cura della salute del paziente interessato. Al contrario, per il trattamento di dati, che, sebbene connessi alla cura della salute del paziente, non sono necessari a tal fine, la legittimità sussiste solo se è presente un'altra base giuridica tra quelle previste dal regolamento europeo (ad esempio, il consenso dell'interessato).

Dalla lettura della norma, pertanto, sembra emergere che l'Azienda sanitaria detenga il possesso della cartella clinica contenente i dati sanitari, ma il paziente sia, comunque, titolare del diritto di proprietà sui dati.

In particolare, anche le fotografie scattate ai fini di interventi chirurgici rientrano nella categoria dei dati personali ed è, pertanto, legittima la richiesta da parte del paziente di acquisire tali immagini[101]. Inoltre, il Garante ha ribadito il diritto del paziente all'accesso a dati personali anche relativi a un esame, o un intervento, indipendentemente dal supporto fisico su cui sono conservati[102].

A riprova di ciò, mentre la conservazione delle cartelle cliniche ospedaliere e delle case di cura è a tempo illimitato, nessuna norma prevede che il libero-professionista debba conservare (e per quanto tempo) la scheda clinica dei propri pazienti e la documentazione allegata. Va tenuto presente che per conservare dati sensibili il libero professionista deve richiedere l'autorizzazione al paziente, in caso contrario deve, nel rispetto degli artt. 15 e 16 del GDPR, distruggere ogni documento compilato, a meno che sia di proprietà del paziente: in questo caso deve restituirlo[103].

Pertanto, da quanto emerge, sembrerebbe che i pazienti debbano fornire il consenso all'utilizzo dei propri dati personali per l'addestramento delle macchine, determinando un livello di difficoltà maggiore per le società che si occupano dello sviluppo e del *testing* dei *software* di Intelligenza Artificiale.

6. Conclusioni. La "Competenza Artificiale": monitoraggio, formazione e informazione dei TSRM quali strumenti principali per fronteggiare il fenomeno. – Terminata l'analisi degli aspetti problematici che emergono quando si considera l'utilizzo di sistemi di Intelligenza Artificiale, è necessario prendere in considerazione quali siano le esigenze specifiche dei tecnici sanitari di radiologia e, conseguentemente, definire quale sia il ruolo che in questo processo evolutivo può essere riconosciuto alla Federazione nazionale che li rappresenta.

Si sta assistendo a un processo evolutivo finalizzato al miglioramento delle prestazioni sanitarie sia da un punto di vista qualitativo, sia in termini di efficienza ed ottimizzazione: si pensi,

[101] Il diritto di accesso ai dati personali è disciplinato all'art. 15, Reg. 679/2018.

[102] Si v. decisione 20 settembre 2006 sul Bollettino n. 75.

[103] Il Tribunale di Foggia (sent. 11/09/2011) ha precisato che, ai sensi dell'art. 2235 cod. civ., il prestatore d'opera non può ritenere le cose e i documenti ricevuti, se non per il periodo strettamente necessario alla tutela dei propri diritti secondo le leggi professionali.

per esempio, al caso in cui il medico radiologo debba effettuare tutte le operazioni per calcolare la quantità di mezzo di contrasto necessaria ai fini di un esame per ogni singolo paziente. Con l'utilizzo dei sistemi di Intelligenza Artificiale tale operazione non solo risulterà semplificata, ma anche migliorata, riducendo ulteriormente sia i tempi di calcolo che il margine di errore.

Tuttavia, bisogna prendere coscienza di che cosa si sta utilizzando, in modo tale da assumere un atteggiamento proattivo nei confronti della macchina e dell'innovazione tecnologica.

Nella parte introduttiva si è evidenziato quali siano le difficoltà esistenti nel proporre una definizione unitaria di "Intelligenza Artificiale": tuttavia, indipendentemente da tale difficoltà, è possibile individuare le caratteristiche fondamentali di quelle utilizzate in un ambito specifico (come, ad esempio, il settore dell'*imaging*) e comprenderne le modalità di funzionamento. Ad esempio, in questo settore, è molto rilevante determinare la competenza richiesta nello svolgimento di una determinata attività: se si considerano i compiti attribuiti all'Intelligenza Artificiale e l'*interplay* che essa realizza con il medico radiologo e con il tecnico sanitario di radiologia che la assiste, allora si può osservare che la caratteristica che connota l'algoritmo che viene utilizzato non è tanto ravvisabile nell'Intelligenza, quanto nella competenza.

Per questo motivo potrebbe essere più opportuno parlare di "Competenza Artificiale".

Ciò posto, allora il quesito diventa: quale competenza dovrà avere il TSRM per poter gestire nel miglior modo possibile questa "Competenza Artificiale"?

Un tema molto importante, anche considerando che, quando l'evoluzione della tecnologia produce semplificazione nella propria operatività, questa determina sempre – in modo inversamente proporzionale – una equivalente complicazione sotto il profilo tecnico che richiede conoscenze e culture specifiche. Si pensi, ad esempio, all'evoluzione della telefonia. Gli apparecchi telefonici in uso cinquant'anni fa avevano una struttura molto semplice sotto il profilo meccanico, consentivano solo di telefonare e richiedevano manutenzioni di facile realizzazione. Oggi, gli *smartphones* sono dei piccoli ma potenti *computer*, svolgono migliaia di funzioni richiedendo pochissima competenza tecnica ma non sono, sostanzialmente, riparabili dal comune consumatore: in altre parole, sono "*smart*" nel loro utilizzo ma, certamente, non nella loro comprensione sotto il profilo tecnico.

Innanzitutto, in questo processo, il TSRM ha (e potrebbe avere sempre più) un ruolo fondamentale di interazione fra il paziente e la macchina. Il che evidenzia alcuni principali profili peculiari che necessitano di approfondimento.

In questo quadro, nella realizzazione dell'esame diagnostico sarà indispensabile fornire al professionista sanitario una competenza specifica finalizzata a consentire alla macchina di poter operare al meglio delle sue possibilità eliminando, il più possibile, quelle variabili soggettive che possano ridurne (o eliminarne) l'operatività, permettendo al processo di concludersi nel modo più efficiente. Questo richiederà la definizione di linee guida a cui fare riferimento, che possano tradursi in protocolli assistenziali: tali protocolli, a loro volta saranno anche condivisi con l'industria. Infatti, nel tempo, ogni singola macchina presenterà caratteristiche specifiche che richiederanno al TSRM una specifica formazione ed informazione che consentano di utilizzarla nel modo più adeguato ed avere il pieno controllo sulla situazione.

Tutta una serie di *standard* di condotta del TSRM che andranno definite dalle società tecnico-scientifiche abilitate alla emanazione di linee guida (*Faster*), tenendo presente il ruolo primario svolto dall'industria, nonché il quadro giuridico in cui tutto questo processo si colloca. Un bagaglio formativo ed informativo che dovrà essere trasferito ai tecnici sanitari di radiologia medica attraverso percorsi di formazione che consentano loro una preparazione adeguata ad affrontare il fenomeno.

Sempre con riferimento al paziente e al rapporto con l'industria, occorre sottolineare anche il rapporto fondamentale esistente tra il tecnico sanitario di radiologia medica, i dati del paziente e il materiale diagnostico prodotto: dati fondamentali per l'industria, dalla cui acquisizione dipende (e

dipenderà sempre più) l'addestramento degli algoritmi. Un addestramento la cui realizzazione richiede dati che siano uniformi, "puliti" e fungibili.

Diviene, pertanto, rilevante comprendere come gestire il rapporto con l'industria che lavora per l'introduzione di macchine sempre più innovative. Infatti, per l'addestramento degli algoritmi servono, come già ricordato, grandi quantità di dati: tuttavia, tali dati acquistano lo *status* di "dati sensibili" (c.d. "particolari" *ex* art. 9 del GDPR) e devono essere trattati con cautela. Forse, nel tempo, verrà ridefinito il concetto di "riservatezza", ma, per ora, è necessario trovare un equilibrio tra il diritto alla *privacy* (e il divieto di trattamento *ex* art. 9) da un lato, e la "fame" – più che giustificabile – di dati da parte dell'industria.

Sotto questo profilo, il TSRM ha una posizione privilegiata di accesso immediato a questi dati ed è anche l'unico soggetto ad avere un rapporto così stretto con la macchina. Pertanto, in primo luogo sarà necessario che il TSRM consenta un'acquisizione dei dati che sia condivisa ed efficiente: ma, ancora una volta, si tratta di attività che richiede (e richiederà sempre più) una specifica formazione del professionista sanitario sulla base di protocolli che dovranno essere condivisi con l'industria nel rispetto della normativa in vigore.

In secondo luogo, bisogna prestare attenzione alla relazione fra TSRM e macchina. In tal senso, sotto il profilo strettamente informatico, il rapporto che il professionista dovrà avere con quella specifica Intelligenza Artificiale e la maggiore conoscenza che sarà necessariamente acquisita nello studio del suo funzionamento, conferirà allo stesso tecnico un ruolo di monitoraggio e di rilevazione degli eventuali problemi ed anomalie che, inevitabilmente, potrebbero sorgere nell'uso del *software*: problemi che potrebbero essere sia "ordinari" che "straordinari" e che, in ogni caso, dovranno essere tempestivamente individuati e comunicati per assicurare un'ottima manutenzione della macchina.

Infine, vi è un aspetto particolarmente interessante e delicato. Nella esposizione si sono individuate le ragioni sia psicologiche sia (e, forse, soprattutto) giuridiche, che inducono a sostenere il rischio di consentire una indiscriminata operatività al principio (comportamentale) "*in dubio pro machina*".

In questo specifico ambito il TSRM deve essere in grado di avere un rapporto consapevole e responsabile (*accountable*) non solo rispetto ai dati che sta utilizzando, ma deve essere in grado di monitorare costantemente i risultati tecnici offerti dalla macchina, avendo in mente i rischi di *data bias* e di *automation bias*. Come detto, sotto quest'ultimo profilo, vi è la tendenza degli esseri umani di preferire le decisioni prese da un sistema artificiale, a volte anche ignorando informazioni contrastanti o decisioni "umane" confliggenti.

L'*Automation bias* determina una maggiore frequenza nell'occorrere di errori omissivi e commissivi. I primi sono errori che si presentano quando l'essere umano non riconosce, o sottovaluta un errore dell'Intelligenza Artificiale con cui sta lavorando. Inoltre, spesso questi errori sono determinati dal fatto che l'Intelligenza Artificiale è in grado di identificare elementi che sfuggono all'occhio umano. I secondi, invece, si presentano quando l'essere umano erroneamente accetta o implementa la decisione della macchina, nonostante ci siano prove che dimostrino il contrario.

Ora, considerando lo sviluppo che questo tipo di macchine possono produrre tenendo presente anche la loro applicazione nell'ambito della tele-radiologia, è prospettabile che il primo *fronting* critico – la prima giustificazione – dell'esame proposto e del risultato offerto dalla macchina dotata di Intelligenza Artificiale sia affidato al TSRM, il quale dovrà effettuare una giustificazione sia in "entrata" che in "uscita" del rapporto paziente-macchina. Il che, ancora una volta, richiederà l'acquisizione da parte del tecnico sanitario di radiologia medica di una competenza nella rilevazione di un'anomalia segnalata della macchina "intelligente" e "competente" che sta utilizzando.

Concludendo, da quanto esposto emerge, con estrema chiarezza, quale sia il ruolo che la Federazione Nazionale può (*rectius*: deve) assumere, ponendosi quale stanza di compensazione e di confronto fra i diversi *stakeholders* (società scientifiche, industria e professionisti sanitari) al fine di

mediarne gli interessi attraverso un gruppo di studio permanente che segua passo a passo l'evoluzione dell'applicazione delle Intelligenze Artificiali nel settore dell'*imaging* supportandolo con gli strumenti tecnici, scientifici e giuridici indispensabili al suo corretto funzionamento ed applicazione.

Un supporto culturale indispensabile per fornire risposta a tutta una serie di domande, in modo tale da poter avere un approccio responsabile all'Intelligenza Artificiale, ad esempio:

- Si è in grado di spiegare come l'Intelligenza Artificiale sia giunta a un determinato risultato, o, quanto meno, è possibile ragionevolmente predire i risultati dell'analisi del *software* se i *set* di dati sono conosciuti?
- Quali sono le strategie più opportune da applicare per la protezione dei sistemi e dei dati da attacchi *cyber*?
- Come si potrebbe creare una versione di controllo dei dati, degli algoritmi e dei prodotti che sia più sostenibile?
- Come si potrebbe minimizzare il rischio di danno al paziente derivante da attacchi *cyber* o violazioni della *privacy*?
- Qual è il metodo migliore per verificare modelli già addestrati, prima che questi vengano utilizzati? Quali sono i criteri per determinarne sicurezza ed eticità?
- Come si possono monitorare i modelli di Intelligenza Artificiale nei processi clinici, così da verificare che si comportino come previsto e che la loro *performance* non subisca un declino nel corso del tempo?

Si tratta di alcune delle possibili domande la cui risposta può aiutare a definire una linea direttrice, un orientamento alla analisi del fenomeno: infatti, la risposta non è unica e definitiva. La materia è in tale evoluzione che solo attraverso un costante monitoraggio del fenomeno, dei problemi emersi e delle relative frequenze sarà possibile ricevere dalla scienza e dall'industria insieme la individuazione di risposte che dovranno esprimersi e specificarsi attraverso la definizione di percorsi formativi ed informativi finalizzati a tendere, con il tempo e l'esperienza, ad ottenere un rapporto con l'Intelligenza Artificiale che sia sempre più etico, trasparente, sicuro e performante.

BIBLIOGRAFIA

A

AA.VV., *La responsabilità del produttore*, a cura di G. ALPA, M. BIN., P. CENDON, in *Trattato di dir. comm. e diritto pubblico dell'economia*, diretto da F. GALGANO, vol. XIII, Padova, 1989.

AA.VV., *Responsabilità sanitaria*, a cura di S. ALEO, P. D'AGOSTINO, R. DE MATTEIS, G. VECCHIO, Milano, 2018.

K. ALHEIT, *The Applicability of the EU Product Liability Directive to Software*, in *The Comparative and International Law Journal of Southern Africa*, vol. 34, n. 2, 2001, p. 188 ss..

G. ALPA, U. CARNEVALI, F. DI GIOVANNI, G. GHIDINI, U. RUFFOLO, C. M. VERARDI, *La responsabilità per danno da prodotti difettosi*, Milano, 1990.

A. AMIDEI, *Politica intelligente e responsabilità: profili e prospettive evolutive del quadro normativo europeo*, in U. RUFFOLO (a cura di), *Intelligenza Artificiale e responsabilità*, Milano, 2017, p. 98 ss..

M. ANDERSON, S. ANDERSON, *Il buon* robot, in *Le Scienze*, 2010, p. 90 ss..

E. AZAVEDO, S. ZACKRISSON, I. MEJÀRE, M. HEIBERT ARNLIND, *Is Single Reading with Computer-Aided Detection (CAD) as Good as Double Reading in Mammography Screening? A Systematic Review*, in *BMC Medical Imaging*, vol. 12, 2012, p. 22 ss..

B

A. BACHMANN, A. BERNSTEIN, *Software Process Data Quality and Characteristics: A Historical View on Open and Closed Source Projects*, in *Proceedings of The Joint International and Annual ERCIM Workshops on Principles of Software Evolution*, 2009, p. 119 ss..

A. BENSOUSSAN, *Plaidoyer pour un droit des robots: de la «personne morale» à la «personne robot»*, in *La lettre des juristes d'affairs*, n. 1134 del 28 ottobre 2013.

A. BERTOLINI, *Robots as Products: The Case for a realistic Analysis of Robotic Application and Liability Rules*, in *Law, Innovation and Technology*, vol. 5, n. 2, 2013, p 213 ss..

J. S. BORGHETTI, *How Can Artificial Intelligence Be Defective?*, in S. LOHSSE, R. SCHULZE, D. STAUDENMAYER (a cura di) *Liability For Artificial Intelligence and The Internet Of Things*: *Münster Colloquia on EU Law and the Digital Economy IV*, Baden, 2019, p. 64 ss..

P. BORTONE, L. BUFFONI, *La responsabilità per prodotto difettoso e la garanzia di conformità nel codice del consumo*, Torino, 2007.

B. BROZEK, M. JAKUBIEC, *On the Legal Responsibility of Autonomous Machines*, in *Artificial Intelligence Law*, vol. 25, 2017, p. 301 ss..

C

U. CARNEVALI, voce *Responsabilità del produttore*, in *Enc. Dir. Aggiorn*., II, Milano, 1998, p. 936 ss..

S. CHOPRA, L. F. WHITE, *A Legal Theory for Autonomous Artificial Agents*, Michigan, 2011.

G. COMANDÈ, *Risarcimento del danno alla persona e alternative istituzionali*, Torino, 1999, p. 333 ss..

G. COMANDÉ, *Intelligenza Artificiale e responsabilità tra* liability *e* accountability. *Il carattere trasformativo dell'IA e il problema della responsabilità*, in *Analisi Giuridica dell'Economia*, vol. 1, 2019, p. 180 ss..

A. CORDIANO, *Sub art. 115*, in E. CAPOBIANCO, L. MEZZASOMA, G. PERLINGIERI (a cura di), *Codice del consumo annotato con la dottrina e giurisprudenza*, Napoli, 2018, p. 633 ss..

S. CROSSLEY, L. L. P. EVERSHEDS, *EU Regulation of Health Information Technology, Software and Mobile Apps*, in *Practical Law Global Guide*, 2016(17), 2016, p. 1 ss..

D

D. C. DENNETT, *Dai batteri a Bach. Come evolve la mente*, Milano, 2018.

F. DI GIOVANNI, *Intelligenza Artificiale e rapporti contrattuali*, in U. RUFFOLO (a cura di), *Intelligenza Artificiale e responsabilità*, Milano, 2017, p. 127 ss..

F

D. FAIRGRIEVE, G. HOWELLS, P. MØGELVANG-HANSEN, G. STRAETMANS, D. VERHOEVEN, P. MACHNIKOWSKI, A. JANSSEN, R. SCHULZE, *Product Liability Directive*, in P. MACHNIKOWSKI (a cura di), *European Product Liability: An Analysis of the State of the Art in the Era of New Technologies*, Cambridge, 2016, p. 61 ss..

J. M. FISCHER, M. S. J. RAVIZZA, *Responsibility and Control*: *A Theory of Moral Responsibility*, Cambridge, 1998.

G

W. M. GROVE, *Clinical Versus Statistical Prediction: The Contribution of Paul E. Meehl, in Journal of Clinical Psychology*, vol. 61, n. 10, 2005, p. 1233 ss..

H

G. HOWELLS, *Defect in English Law: lesson for the harmonization of European product liability*, Cambridge, 2005.

J

R. JAMES, *Creating an Immutable Audit Trail on the Blockchain with Xero & Tierion*, disponibile su: https://devblog.xero.com/creating-an-immutable-audit-trail-on-the-blockchai n-with-xero-tierion-be423d39380b.

K

KAHNEMAN, *Pensieri lenti e veloci*, Milano, 2014.

J. KAPLAN, *Intelligenza artificiale. Guida al futuro prossimo*, Roma, 2017.

M. KOHLI, R. GEIS, *Ethics, Artificial Intelligence, and Radiology*, in *Journal of the American College of Radiology*, vol. 15, 2018, p. 1317 ss..

L

I. J. LLOYD, *Information Technology Law*, Oxford, 2008.

M

R. MARMO, *Algoritmi per intelligenza artificiale*, Milano, 2020.

A. MATTHIAS, *The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata*, in *Ethics and Information Technology*, vol. 6, 2004, p. 175 ss..

J. MCCARTHY *et al., A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, August 31, 1955, in *AI magazine*, vol. 27, n. 4, 2006, p. 12 ss..

A. J. MÉNDEZ, P. G. TAHOCES, M. J. LADO, M. SOUTO, J. J. VIDAL, *Computer-aided Diagnosis: Automatic Detection of Malignant Masses in Digitized Mammograms,* in *Medical Physics*, vol. 25, 1998, p. 957 ss..

T. M. MITCHELL, *Machine Learning*, Milano, 1997.

P

E. PALMERINI, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea,* in *Responsabilità civile e previdenza*, vol. 6, 2016, p. 1835 ss..

S. PARRA, M. TORRENS, *Intelligenza Artificiale. La strada verso la superintelligenza*, Milano, 2017.

F. PASQUALE, *The Black Box Society: The Secret Algorithms that Control Money and Information*, Boston, 2015.

S

J. R. SEARLE, *Mind, Brains and Programs. A Debate on Artificial Intelligence*, in *The Behavioral and Brain Science*, vol. 3, 1980, p. 128 ss..

L. B. SOLUM, *Legal Personhood for Artificial Intelligences*, in *North Carolina Law Review*, vol. 70, n. 4, 1992, p. 1231 ss..

L.M. SPENCER, M. S. SPENCER, *Competenza nel lavoro*, Milano, 1995.

T

A. TETTENBORN, *Product Liability and Consumer Protection*, in *Clerk & Lindsell on Torts*, Fasc. 11, 2010, p. 11 ss..

J. TRIAILLE, *The EEC Directive on Product Liability and its Application to Databases and Information*, in *Computer Law and Practice*, 1991, p. 219 ss..

A. M. TURING, *Computing Machinery and Intelligence*, in *Mind*, vol. 49, 1950, p. 433 ss..

V

D. C. VLADECK, *Machines without Principals: Liability, Rules and Artificial Intelligence*, in *Washington Law Review*, vol. 89, n. 117, 2014, p. 117 ss..

W

G. WAGNER, *Robot Liability*, in S. LOHSSE, R. SCHULZE, D. STAUDENMAYER, (a cura di), *Liability for Robotics and in the Internet of Things: Munster Colloquia on Eu Law and the Digital Economy*, West Sussex, 2019, p. 36 ss..

G. WAGNER, *Robot, Inc.: Personhood for Autonomous Systems?*, in *Fordham Law Review*, vol. 88, n. 2, 2019, p. 608 ss..

D. WUYTS, *The Product Liability Directive: More than two decades of defective products in Europe*, in *Journal European Tort Law*, vol. 5, 2014, p. 1 ss..

Z

F. E. ZOLLERS *et al.*, *No More Soft Landings for Software: Liability for Defects in an Industry That Has Come of Age*, in *Santa Clara High Technology Law Journal*, vol. 21, n. 4, 2005, p. 750 ss..