

## INTRODUZIONE

L'impiego su larga scala di sistemi di Intelligenza Artificiale (d'ora in poi, per semplicità, IA) pone all'attenzione del giurista una serie di questioni rilevanti che obbligano ad osservare con uno sguardo nuovo le tradizionali categorie concettuali dell'ordinamento.

L'avvento di dispositivi dotati di *software* e basati su complessi modelli matematici capaci di svolgere specifiche funzioni in modo autonomo – apprendendo dai dati ricavabili dall'ambiente in cui operano – produce già oggi effetti dirompenti e, per certi aspetti, inediti sotto il profilo etico, sociale e, in special modo, giuridico.

In relazione a quest'ultimo aspetto, infatti, si pongono rilevanti problemi di mediazione tra il dato normativo esistente e una tecnologia, come quella in parola, che trova applicazione in scenari assai eterogenei e che si evolve di continuo con modi e forme non sempre prevedibili; tecnologia che “parla” il linguaggio proprio delle scienze informatiche e che impone anzitutto, allo studioso del diritto che ad essa si accosti, la conoscenza della dimensione teorica del fenomeno, premessa pregiudiziale a qualsiasi valutazione ermeneutica.

“Intelligenza Artificiale”, “algoritmo”, “rete neurale” sono espressioni di assai frequente richiamo nella letteratura giuridica degli ultimi anni – e, ancor prima e ancor più, nel linguaggio quotidiano dei *media* – ma appaiono, in larga parte, utilizzati in modo a-tecnico, stanno semplicemente a significare che una certa operazione (o un insieme di operazioni) è, in tutto o in parte, sottratta al controllo diretto dell'uomo e rimessa invece a sistemi di calcolo, operativi, decisionali, gestiti da un algoritmo, anche molto sofisticato, che rappresenta l'elemento fondante di qualsiasi sistema di IA.

Proprio nelle pieghe di una tale considerazione si cela l'insidia, perché l'automazione di certi processi non sempre e non necessariamente altera o incide sull'applicazione dei tradizionali principi normativi (e, in particolare, sulle regole della responsabilità civile, oggetto del presente lavoro); né, prima ancora, consente di per sé di distinguere agevolmente l'operato della macchina da quello del suo sviluppatore o dell'ideatore dell'algoritmo.

La questione centrale riguarda, piuttosto, l'individuazione del punto di non ritorno, vale a dire della fase in cui il grado di automazione del sistema diventa tale da renderlo “altro” rispetto a colui (o a coloro) che lo ha progettato.

Il tema è di grande interesse non solo sul piano tecnico, respira le suggestioni della letteratura di fantascienza che da decine di anni predice un mondo di automi e robot.

Calata nel contesto giuridico la questione – ed i suoi profili problematici – appare chiara: l'autonomia operativa del sistema di IA pone il tema dell'ascrizione degli effetti dell'agire; ma a chi? Al sistema c.d. intelligente, a chi lo ha creato, o a chi beneficia del prodotto, comunque configurato, che lo incorpora?

A questo riguardo, si darà conto, per cenni, dei tentativi di riconoscere una “soggettività” al sistema intelligente e delle difficoltà teoriche (e ancor più applicative) di una tale impostazione.

La questione si rivela di particolare complessità quando l'operato del sistema algoritmico determina la causazione di un danno. Al fine di affrontare efficacemente tale circostanza, il richiamo alla possibile soggettività del sistema non appare risolutivo: se la responsabilità implica, infatti, l'individuazione di un “ente” cui imputare l'obbligazione risarcitoria, l'attribuzione di una soggettività al sistema non è ancora la risposta, quantomeno finché non gli venga riconosciuta anche una capacità patrimoniale. In questa prospettiva, sarà allora necessario interrogarsi su chi debba concretamente assicurare tale capacità e se un simile espediente possa costituire uno strumento realmente efficace ai fini della tutela dei danneggiati.

Vero è che l'autonomia del sistema di IA è sempre il punto di arrivo di un percorso complesso e in più fasi, che di quel sistema vede lo sviluppo, la programmazione per le funzioni alle quali è stato destinato e per le quali è stato istruito – vale a dire “reso intelligente” – e, da ultimo, l'integrazione in un sistema operativo più ampio, di cui esso è la “mente”.

Non vi è bisogno di scomodare la soggettività per capire che se il sistema è potuto diventare intelligente ciò dipende dal fatto che è stato ideato e programmato in tal senso; e che, anche nell'ipotesi in cui esso fuoriesca dalla sfera di controllo del suo sviluppatore, autoalimentandosi ed istruendosi grazie ad algoritmi di autoapprendimento, non per questo vengono necessariamente meno le questioni relative all'identificazione del responsabile delle azioni di un simile apparato. Tali questioni si collocano lungo una scala che muove dall'ipotesi più elementare, nella quale l'operato della macchina coincide, in realtà, con quello dell'uomo che la governa, sino allo scenario, più incerto e controverso, in cui la macchina, apparentemente slegatasi dal suo ideatore, sarebbe davvero in grado di prendere decisioni (*rectius* di fare previsioni) autonome, in un processo deliberativo che non apparterrebbe – e non sarebbe riferibile – se non a sé stessa.

È in questo contesto che si colloca l'indagine che ci si propone di condurre: un contesto nel quale si è chiamati a interrogarsi su chi debba essere ritenuto responsabile dell'operato del sistema intelligente. La complessità della domanda

è acuita dalla circostanza che il punto di arrivo – l'intelligenza della macchina – è il prodotto dell'intervento di una molteplicità di attori – dall'analista al programmatore, dall'addestratore al manutentore, sino a colui che ne cura gli aggiornamenti – la cui azione congiunta ha reso possibile il risultato.

L'emergere dell'autonomia e dell'autosufficienza del sistema intelligente pone quesiti nuovi, fino ad ora impensabili in un contesto nel quale l'operato di qualsiasi apparecchiatura, macchinario, *software*, era, per quanto articolato e sofisticato, sempre riferibile all'attività dell'uomo che l'aveva ideato e messo in condizioni di operare.

Con particolare riguardo all'ambito del diritto privato e al tema della responsabilità per danni, l'innovativo scenario determinato dalla diffusione dei sistemi di Intelligenza Artificiale, tra cui, in particolare, quelli dotati di capacità di autoapprendimento, solleva una serie di questioni problematiche, in ragione delle possibili conseguenze negative derivanti dall'impiego di tali strumenti.

L'IA offre, invero, grandi potenzialità ma comporta al tempo stesso il venire ad esistenza di nuovi rischi, conseguenti, in special modo, alla sua natura *self-learning* che consente a tali sistemi di interagire con l'ambiente circostante e di alterarlo in modo significativo, con la possibilità – come si osservava – di causare danno.

È lecito allora chiedersi se a fronte del ricorso a sistemi algoritmici si debbano immaginare nuovi strumenti che ne regolamentino le possibili, conseguenti, ipotesi di responsabilità; e se vadano previsti meccanismi diversi da quelli tradizionali per l'ascrizione di responsabilità a coloro che si debbano fare carico del malfunzionamento o dell'inefficienza del sistema.

La questione non è trascurabile, né secondaria, e pone profili in crescente ordine di difficoltà.

Il piano di più immediata percezione riguarda l'applicabilità delle disposizioni tradizionali in materia di responsabilità: è sufficiente il richiamo ai principi contenuti negli artt. 1218 e 2043 c.c.? E quale margine di applicabilità riconoscere alle più specifiche disposizioni in materia di fatto illecito e alle norme in tema di responsabilità da prodotto difettoso?

L'interrogativo è lecito perché, come si rilevava, il danno derivante dal ricorso ad un sistema di Intelligenza Artificiale raramente costituisce l'esito di un meccanismo causale elementare, nel quale la responsabilità è agevolmente ascrivibile a chi quel sistema è tenuto a controllare e a gestire (ad esempio, come si vedrà, il conducente di un veicolo parzialmente automatizzato, che non per questo è esonerato dalla vigilanza sul funzionamento della tecnologia di bordo).

Una situazione diversa – e più complessa – si realizza quando il prodotto intelligente è sottratto al controllo dell'utilizzatore, il quale non assume neppure funzioni correlate; si pensi ad un'apparecchiatura, basata su modelli

algoritmici, utilizzata in una struttura sanitaria per la somministrazione di trattamenti terapeutici; o, non molto diversamente, ad un sistema di IA facente parte di un'entità ugualmente complessa nella quale elemento algoritmico e componente meccanica si integrano a vicenda (è il caso di un sistema frenante intelligente che interagisce con le pinze del freno che esercitano l'azione meccanica nell'arresto del veicolo).

In queste ipotesi il problema si amplia sino a investire due profili di grande rilievo: da un lato, la responsabilità di chi si avvale di un sistema di IA – non direttamente progettato e controllato – nell'esercizio della propria attività d'impresa; dall'altro, la responsabilità solidale di coloro che, a vario titolo, hanno concorso allo sviluppo di un prodotto che implichi anche l'utilizzo di avanzati modelli algoritmici.

Da qui la necessità di esaminare le principali questioni attinenti alle responsabilità da impiego di sistemi intelligenti, e in particolar modo quelle relative alla verifica della compatibilità e della tenuta dell'attuale impianto normativo, tanto di matrice nazionale quanto di derivazione europea. Un secondo snodo dell'indagine riguarda le possibili lacune del quadro regolatorio, che spingono ad interrogarsi sulla necessità di introdurre nuove disposizioni che soddisfino le mutate esigenze del sistema e possano disciplinare efficacemente i fenomeni tecnologici emergenti. Un terzo passaggio è dedicato all'analisi dei possibili correttivi, funzionali ad adeguare la legislazione esistente e a valorizzarne la forza adattiva, con l'obiettivo non solo di assicurarne l'adeguatezza rispetto alle sfide poste dall'IA, ma anche di sfruttarne appieno le potenzialità attraverso interpretazioni evolutive e interventi normativi mirati. Da ultimo, sono prese in esame le nuove regole introdotte dal legislatore europeo sul fronte preventivo, con il regolamento (UE) 2024/1689 (c.d. AI Act), e risarcitorio, con la direttiva (UE) 2024/2853 sulla responsabilità da prodotto (non ancora in vigore). L'obiettivo è verificare se il sistema a “doppio binario” delineato dal diritto dell'Unione – da un lato norme preventive di sicurezza e conformità, dall'altro strumenti di tutela risarcitoria – sia effettivamente idoneo a garantire la circolazione sul mercato di prodotti più sicuri e, al contempo, a rafforzare la tutela del danneggiato. In questa prospettiva, sarà, altresì, utile accertare se – e in quale misura – l'intervento eurounitario, con l'adozione della recente normativa in materia di prodotti difettosi, abbia superato i profili problematici connessi all'adeguamento della direttiva 85/374/CEE ai nuovi scenari produttivi. Occorrerà, inoltre, valutare se tale risultato sia stato conseguito mediante la cristallizzazione, a livello di diritto positivo, delle soluzioni già elaborate in via interpretativa da dottrina e giurisprudenza.

Ai fini della sistematicità dell'indagine verranno dapprima affrontati alcuni profili tecnici e definatori relativi ai sistemi di IA, cui farà seguito l'analisi critica – nei limiti necessari al presente studio – di alcune delle tesi sostenute in dottrina e dalle stesse istituzioni europee allo scopo di

giungere al riconoscimento di una soggettività di diritto dei dispositivi intelligenti.

Nei capitoli successivi (II e III) sarà oggetto di verifica l'applicabilità della disciplina in materia di responsabilità civile – in particolare di quella codicistica e di *product liability* – ai profili problematici generati dall'applicazione di sistemi di IA. Occorre, tuttavia, considerare che tali sistemi, pur accomunati dall'impiego di una medesima tecnologia, trovano applicazione in settori assai eterogenei, la cui disciplina è demandata a plessi normativi non sempre coincidenti. A seconda dell'ambito applicativo e dei soggetti coinvolti, infatti, possono emergere esigenze regolatorie diverse e variare i possibili centri di imputazione della responsabilità: se in tema di mobilità vengono in gioco, accanto al produttore, anche il conducente e il proprietario del veicolo, nell'ambito della salute occorre interrogarsi sul ruolo e sulle responsabilità del medico e della struttura sanitaria. Ne deriva che ciascun settore impone specifiche valutazioni non solo giuridiche, ma anche etiche e di politica del diritto, che non possono essere automaticamente trasposte da un contesto all'altro. Ciò rende ancor più rilevante l'analisi delle regole della responsabilità civile alla luce delle peculiarità proprie di ciascun ambito.

Sarà, dunque, oggetto di esame l'applicazione dell'IA al settore dei sistemi di trasporto intelligenti, con particolare riguardo ai veicoli *self-driving* e, in prospettiva, *driverless*. Attraverso lo studio di questo *case study* verranno, da un lato, analizzate le clausole di garanzia e di esonero da responsabilità contenute nei contratti standard formulati da Tesla e, dall'altro, affrontate alcune delle questioni sollevate, in punto di responsabilità civile, dalle tecnologie in parola. A queste ultime si offrirà risposta attraverso l'interpretazione sistematica dell'attuale normativa civilistica e di quella in materia di danno da prodotto difettoso, opportunamente adattate.

Il prosieguo dell'indagine sarà dedicato alle applicazioni di sistemi di Intelligenza Artificiale nel campo dell'assistenza sanitaria, settore tra i più interessati dalla pervasività degli strumenti di IA. Oggetto di approfondimento, in particolare, è il problema delle implicazioni che la diffusione di tali tecnologie esercita sulle regole di responsabilità, tanto nei rapporti tra medico e paziente, quanto rispetto alla Struttura sanitaria che si avvale di beni intelligenti nell'erogazione della prestazione. Sarà, inoltre, oggetto di analisi – con specifico riferimento ai dispositivi medici *AI-driven* – la portata estensiva della nozione di “prodotto” (e di sua componente) delineata dalla direttiva 85/374/CEE, al fine di accertare se essa possa, in via interpretativa, essere applicata anche al *software*, all'algoritmo e al sistema che li incorpora. Tale verifica si rende necessaria anche alla luce del progressivo assottigliarsi della tradizionale distinzione tra “prodotto” e “servizio”, conseguenza diretta dell'avvento sul mercato di prodotti digitali integrati.

L'indagine si concluderà, infine, con l'analisi dell'architettura preventiva

costruita dall'AI Act e dei riflessi che la stessa comporta sul fronte della responsabilità dei soggetti coinvolti nello sviluppo e nell'utilizzo dei sistemi di IA. Invero, la mancata osservanza dei requisiti di sicurezza previsti dal regolamento (UE) 2024/1689 potrebbe incidere sia sulla nozione di difetto del prodotto ai sensi dell'art. 117 c. cons. (art. 7 della nuova direttiva), sia – nella filiera di rapporti tra professionisti o nei confronti dell'utilizzatore del sistema – sulla determinazione della colpa ai sensi dell'art. 2043 c.c. o, in generale, del grado di diligenza esigibile dal soggetto responsabile.

L'indagine cercherà, altresì, di verificare se la nuova direttiva in materia di prodotti difettosi risolva attraverso la *regola iuris* le criticità sinora affrontate dalle corti e dalla dottrina in via ermeneutica, e se consenta una efficace ripartizione dei rischi e tutela del danneggiato. L'analisi della direttiva 85/374/CEE consente anche di interrogarsi sulle più recenti linee di tendenza che sembrano caratterizzare la responsabilità da prodotto difettoso, alla luce degli sviluppi connessi alle nuove tecnologie e ai nuovi modelli dell'economia circolare e del commercio elettronico.

Riflettere sul modello europeo significa coglierne l'ambizione di coniugare innovazione e sicurezza, progresso tecnologico e responsabilità, in un equilibrio che rappresenta al tempo stesso la cifra distintiva e il principale banco di prova del "nuovo" diritto dell'Intelligenza Artificiale. L'IA, infatti, si caratterizza per una intrinseca ambivalenza: da un lato, promette vantaggi considerevoli in termini di efficienza, rapidità e accuratezza delle decisioni, con ricadute positive in molti ambiti (tra cui quelli della mobilità e della sanità) e, più in generale, sulla qualità della vita delle persone; dall'altro, genera rischi inediti che spaziano dall'opacità algoritmica alla difficoltà di imputazione delle responsabilità, fino a nuove forme di discriminazione e dipendenza tecnologica e a costi ambientali crescenti.

In tale quadro non va trascurata la capacità delle discipline nazionali in materia di r.c. di adattarsi, attraverso un'interpretazione evolutiva, alla regolazione di fenomeni inediti e alla gestione delle nuove occasioni di danno, così come di misurarsi con i mutati scenari produttivi e di offrire strumenti idonei a disciplinare i pregiudizi derivanti da prestazioni tradizionali che, oggi, si avvalgono dell'Intelligenza Artificiale per la loro erogazione.

È nel ricercare un punto di equilibrio dinamico in un contesto così articolato e mutevole che si annida la vera sfida del regolatore e dell'interprete, affinché promozione dell'innovazione e protezione dei diritti fondamentali non si escludano, ma si rafforzino reciprocamente, dando vita a un ordinamento giuridico che sia al tempo stesso equo, efficiente e sostenibile.