

CAPITOLO 8

Responsabilità per danno derivante dall'uso dell'intelligenza artificiale

SOMMARIO: §1. La nozione di intelligenza artificiale. - §2. La tutela preventiva contro i danni derivanti dall'impiego dell'intelligenza artificiale: le pratiche vietate previste dal Regolamento UE e specificate dalle linee guida del febbraio 2025. - §3. La tutela risarcitoria. Alla ricerca della fonte normativa: artt. 2043, 2048, 2047 c.c.: 3.1 Segue: ... artt. 2049, 2051, 2052 c.c.; 3.2 Segue: ... art. 2050 c.c. - §4. La abortita proposta di direttiva sulla responsabilità extracontrattuale dell'IA lascia un vuoto normativo a livello comunitario. - §5. La nuova disciplina per la responsabilità per i danni prodotti dall'intelligenza artificiale introdotta dal D.Lgs. 9 settembre 2026, n. 160: 5.1 L'accesso alle prove; 5.2 La presunzione del nesso di causalità tra violazione degli obblighi ex Reg. UE 2024/1689 e il danno; 5.3 Foro del consumatore e azione diretta nei confronti dell'impresa di assicurazione. - §6. Conclusioni.

1. La nozione di intelligenza artificiale

Prima di procedere all'analisi della responsabilità per i danni cagionati dall'uso dell'intelligenza artificiale, riteniamo propedeutico renderne preliminarmente la **nozione**.

Il **nuovo Regolamento (UE) del 13 giugno 2024, n. 1689**, definisce all'**art. 3, n. 1**, l'intelligenza artificiale come “un **sistema basato su macchina** progettato per operare con diversi livelli di **autonomia** e che può mostrare **adattabilità** dopo il suo impiego, e che, per **obiettivi** espliciti o impliciti, deduce, **in base agli input** che riceve, come **generare output** come previsioni, contenuti, raccomandazioni o decisioni che possono **influenzare ambienti fisici o virtuali**”.

Dalla lettera della norma emerge che la definizione comprende **sette elementi principali**; deve, infatti, trattarsi di un sistema: 1) basato su una macchina; 2) progettato per operare con vari livelli di autonomia; 3) che può mostrare capacità di adattamento dopo l'implementazione 4) e che, per obiettivi espliciti o impliciti; 5) deduce, dagli input che riceve, come generare output; 6) come previsioni, contenuti, raccomandazioni o decisioni 7) che possono influenzare ambienti fisici o virtuali.

Al riguardo, è stato correttamente osservato (SIROTTI GAUDENZI) come la **nozione** sia **tutt'altro che chiara**, e che sarebbe stato più utile spiegare prima cosa si intendesse con l'espressione «intelligenza artificiale» con poche parole, come fatto dall'*Executive Order on Artificial Intelligence* statunitense (da ultimo ritirato dall'amministrazione americana).

La non immediata comprensione della definizione ha indotto la **Commissione europea** a dettare delle **linee guida** che **esplicassero la nozione** riportata nell'*intelligence act*.

Rispetto al primo elemento, nelle *Guidelines* si chiarisce che **il termine «macchina»** vada inteso come comprensivo sia dei componenti *hardware* sia di quelli *software* che consentono al sistema di intelligenza artificiale di funzionare. L'espressione «basato su macchine» deve, pertanto, intendersi **“omni-comprendensiva”** ed è in grado di abbracciare un'ampia varietà di «sistemi computazionali».

In riferimento al secondo elemento, si segnala che il considerando 12 del regolamento 2024/1689 dispone che «i sistemi di IA sono progettati per funzionare con livelli di autonomia variabili, il che significa che dispongono di un certo grado di autonomia di azione rispetto al coinvolgimento umano e di capacità di funzionare senza l'intervento umano». Ebbene, secondo la Commissione, «le **nozioni di autonomia e di inferenza vanno di pari passo**: la capacità di inferenza di un sistema di IA (cioè la sua capacità di generare *output* come previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali) è fondamentale per la sua autonomia» (punto 15 delle *Guidelines*).

Il punto 22 delle linee guida in merito **al terzo elemento** della definizione di cui all'art. 3, §1, del regolamento (il sistema «può mostrare **capacità di adattamento** dopo l'implementazione») specifica che i concetti di **autonomia e adattabilità sono distinti**, ma strettamente correlati e «rappresentano dimensioni diverse della funzionalità di un sistema di IA» (come si legge nelle Linee).

Come visto, il quarto elemento della definizione è costituito dagli **obiettivi del sistema di intelligenza artificiale**, che possono derivare dai dati di addestramento o dall'interazione del sistema con il suo ambiente. Il documento della Commissione precisa che gli obiettivi del sistema possono essere **definiti in modo esplicito o implicito**: «gli obiettivi espliciti si riferiscono a obiettivi chiaramente dichiarati che sono direttamente codificati dallo sviluppatore nel sistema». Invece, si legge che «gli obiettivi impliciti si riferiscono a obiettivi che non sono dichiarati in modo esplicito, ma possono essere dedotti dal comportamento o dalle ipotesi sottostanti del sistema». Questi obiettivi possono derivare dai dati di addestramento o dall'interazione del sistema con il suo ambiente.

Il quinto elemento di un sistema di IA è che deve essere **in grado di generare, dagli input che riceve, gli output**.

Le linee guida specificano le **modalità di questo apprendimento**.

In particolare, la prima categoria di tecniche di IA indicate dalla Commissione è quella relativa all'apprendimento automatico, categoria che contempla un'ampia varietà di approcci che consentono a un sistema di “apprendere”. La seconda categoria di tecniche è costituita da «approcci basati sulla logica e sulla conoscenza che traggono inferenze dalla conoscenza codificata o dalla rappresentazione simbolica del compito da risolvere». Invece di apprendere dai dati, questi sistemi di intelligenza artificiale imparano dalla conoscenza, tra cui regole, fatti e relazioni codificate da esperti umani. Sulla base di queste informazioni, questi sistemi possono “ragionare” tramite motori deduttivi o induttivi, oppure utilizzando operazioni specifiche.

Il **sesto elemento** della definizione di sistema di IA di cui all'art. 3, §1, del regolamento prevede che il sistema sia in grado di generare **output** quali previsioni,

contenuti, raccomandazioni o decisioni [...]». Ebbene, le linee guida specificano queste esternalizzazioni:

- a) le **previsioni** sono uno degli output più comuni che i sistemi di intelligenza artificiale producono e che richiedono il minor coinvolgimento umano;
- b) per **contenuto** si intende la generazione di nuovo materiale da parte di un sistema di intelligenza artificiale: «può trattarsi di testo, immagini, video, musica e altre forme di output»;
- c) per **raccomandazioni** ci si riferisce a suggerimenti di azioni, prodotti o servizi specifici agli utenti sulla base delle loro preferenze, dei loro comportamenti o di altri dati;
- d) le **decisioni** attengono a «conclusioni o scelte fatte da un sistema»; «un sistema di questo genere implica un processo completamente automatizzato tramite cui un certo risultato viene prodotto nell'ambiente circostante il sistema senza alcun intervento umano».

Il **settimo elemento** della definizione di un sistema di IA, infine, è che gli **output** del sistema «possono **influenzare ambienti fisici o virtuali**». Tale elemento dovrebbe essere inteso per sottolineare il fatto che i sistemi di IA non sono “passivi”, ma hanno un vero e proprio “impatto attivo” sugli ambienti in cui vengono distribuiti. Il riferimento ad «ambienti fisici o virtuali» indica che **l'influenza** di un sistema di IA può essere sia **su oggetti fisici tangibili** (ad esempio, un braccio robotico) sia su ambienti **virtuali**, inclusi spazi digitali, flussi di dati ed ecosistemi *software*.

De jure condendo, infine, va riferito che anche nel nostro paese è stato avviato un **iter legislativo per l'adozione di una normativa in materia**. In particolare, **l'art. 2 del D.D.L. AS 1146**, approvato dal Senato in data 20 marzo 2025 e in discussione alla Camera (AC 2316), nell'offrire la nozione di intelligenza artificiale ha **riprodotto l'identico testo contenuto nell'art. 3, n. 1) del Regolamento UE**.

Infine, sentiamo di condividere sostanzialmente la conclusioni delle linee guida della Commissione UE, secondo cui **non è possibile una determinazione automatica o un elenco esaustivo di sistemi** che rientrano o esulano dalla **definizione** di sistema di IA.

Ciò *a fortiori* alla luce delle ultime decisioni politiche della Commissione in materia che (come si vedrà *infra* nel §4), contrariamente all'idea iniziale, ha scelto di ricorrere a una legislazione meno dettagliata per evitare ostacoli alla piena divulgazione dell'intelligenza artificiale. Per le linee guida, infatti, **solo alcuni sistemi di IA sono soggetti a obblighi normativi e di vigilanza** ai sensi della legge sull'IA. L'approccio basato sul rischio della legge sull'IA implica, infatti, che **solo i sistemi che presentano i rischi più significativi per i diritti e le libertà** fondamentali saranno **soggetti ai divieti** stabiliti dall'art. 5 della legge sull'IA, **al regime normativo** per i sistemi di IA ad alto rischio di cui all'art. 6 della legge sull'IA e ai **requisiti di trasparenza** per un numero limitato di sistemi di IA predefiniti di cui all'art. 50 della legge sull'IA. La **stragrande maggioranza dei sistemi**, anche se qualificati come sistemi di IA ai sensi dell'art. 3 della legge sull'IA, infatti, **non è soggetta ad alcun requisito normativo** ai sensi della medesima legge.

2. La tutela preventiva contro i danni derivanti dall'impiego dell'intelligenza artificiale: le pratiche vietate previste dal Regolamento UE e specificate dalle linee guida del febbraio 2025

Una **regolamentazione dell'IA** può avere **due scopi: prevenzione dei danni e disciplina degli atti**, leciti e illeciti. Oggetto del primo è neutralizzare i rischi; del secondo allocare i diritti e le responsabilità nascenti dall'attività tramite sistemi esperti (GENTILI).

Al primo tipo si ispirano l'**AI Act** e le **linee guida** adottate nel febbraio 2025.

La prevenzione ha senso se vi è la percezione di un rischio. Nessuno dubita che l'IA possa avere innumerevoli applicazioni utili, ma è piuttosto comune, almeno nell'area europea, la convinzione che a molti, se non a tutti, i sistemi di IA sia connessa la possibilità di un rischio. Come autorevolmente rilevato (GENTILI), **riferita all'IA la prevenzione** significa da un lato **proibirne gli usi** che siano potenzialmente materialmente o socialmente nocivi, dall'altro **predisporre cautele** proporzionate al differente grado di rischio proprio dei diversi sistemi, o di certe loro applicazioni.

Di qui **nell'impianto dell'AI Act** una **quadripartizione**: pratiche vietate, sistemi ad alto rischio, sistemi a basso rischio, sistemi senza rischio.

Sono vietate le pratiche di specifica ed elevata nocività sociale, cioè lesive della libertà di decisione, della vulnerabilità, della dignità. Il regolamento contempla otto tipologie con ristrette eccezioni (art. 5). Le comminatorie di divieto sono molto dettagliate e dunque si può prevedere che non sorgeranno troppi problemi a riconoscere nel concreto i casi che vi ricadano (GENTILI).

Nonostante la chiarezza della lettera dell'art.5 del Regolamento UE 2024, la Commissione Ue ha adottato, il **4 febbraio 2025**, delle **Linee guida sulle pratiche vietate** che rappresentano «una prima interpretazione con esempi pratici dei divieti indicati» nella disposizione, costituendone, quindi, una chiave di lettura “autentica” da parte dell'organismo comunitario.

Trattasi, in modo particolare, di una **guida operativa** utilizzabile di fronte ai **dubbi** legati alla recente operatività dei divieti previsti per alcuni sistemi di intelligenza artificiale.

Viene ribadita la centralità del c.d. *risk-based approach* dell'AI Act, in virtù del quale, come visto, i sistemi di intelligenza artificiale sono classificati in ragione di quattro diverse categorie di rischio: *i) rischio inaccettabile (unacceptable risk)*, in applicazione del quale i sistemi che presentano rischi inaccettabili per i diritti fondamentali e i valori dell'Unione sono vietati ai sensi dell'art. 5 del regolamento; *ii) rischio elevato (high risk)*, ravvisabile nei sistemi che presentano rischi elevati per la salute, la sicurezza e i diritti fondamentali; tali sistemi sono soggetti a una serie di requisiti e obblighi e sono classificati come «ad alto rischio» ai sensi dell'art. 6 del regolamento e degli allegati I e III; *iii) rischio di trasparenza (transparency risk)*, che richiama i sistemi che presentano un rischio di trasparenza limitato e sono soggetti alle disposizioni dettate dall'art. 50 del regolamento; *iv) rischio minimo o nullo (minimal to no risk)*, riferibile ai sistemi che comportano scarse o inesistenti

probabilità del verificarsi di un danno e che non sono regolamentati (ma i fornitori e i distributori possono aderire volontariamente a codici di condotta volontari).

Qui le norme sulla prevenzione incidono di riflesso sulla repressione: è ragionevole pensare che tali obblighi forniscano la base giuridica per una corresponsabilità civile di tutte queste figure di operatori (SIROTTI GAUDENZI).

3. La tutela risarcitoria. Alla ricerca della fonte normativa: artt. 2043, 2048, 2047 c.c.

Nel programma legislativo delle istituzioni europee compariva anche l'adozione di una **direttiva** tesa a disciplinare la **responsabilità extracontrattuale** per il danno causato dall'utilizzo dell'intelligenza artificiale.

Come vedremo più approfonditamente nel prossimo paragrafo, tuttavia, tale progetto normativo è stato **ritirato nel 2025**. Prima della nuova disciplina introdotta il 9 settembre 2026 dal D.Lgs. n. 160 (di cui si riferirà *amplius* nei §5 e ss.), si è cercata nel codice civile la norma che potesse offrire tutela ai soggetti danneggiati dall'impiego di un sistema di IA.

Lo sforzo del legislatore e degli interpreti, al di là della diversità degli esiti, di cui si dirà, è unico, nobile e meravigliosamente utopistico: fare in modo che la tecnologia sia un mezzo per migliorare la vita degli uomini e non un mostro capace di disumanizzare la nostra anima.

La **prima norma** che potrebbe astrattamente operare è quella contenuta nella **clausola generale** di cui all'art. 2043 c.c.

Si è tuttavia, correttamente obiettato (GENTILI, SCOGNAMIGLIO, SCOTTO DI CARLO, IASELLI) che la **responsabilità** civile connessa all'utilizzo **dell'IA non** può essere ricondotta all'istituto aquiliano disciplinato dall'art. 2043 c.c., **che necessita della colpevolezza** per traslare le conseguenze negative dell'evento dannoso dalla sfera giuridica del danneggiato a quella del danneggiante. Quest'ultimo, peraltro, non va individuato nell'elaboratore, in quanto non rientra tra i soggetti di diritto. Infatti, sebbene non manchino opinioni discordanti, si tende ad assimilare la macchina munita di intelligenza artificiale a un oggetto, cui attribuire le tutele tipiche del diritto di proprietà o di quello di autore, perché nella stessa sono assenti i requisiti caratteristici di un soggetto giuridico. Tra l'altro, il riconoscimento di una soggettività elettronica imporrebbe, da un lato, una significativa rivisitazione del sistema normativo, dall'altro, la costituzione di un patrimonio del soggetto *de quo* (SCOGNAMIGLIO), poiché, in difetto di quest'ultimo, la responsabilità dell'elaboratore non sarebbe di alcuna utilità per chi agisca in giudizio per ottenere il risarcimento del danno, subito a causa della condotta intelligente della macchina (FINOCCHIARO).

Le considerazioni che precedono inducono a **escludere anche il ricorso a forme di responsabilità vicaria** incentrate sul **paradigma genitoriale**, poiché l'elaboratore non potrebbe assumere l'obbligazione risarcitoria di cui all'art. 2048 c.c. (SCOTTO DI CARLO), così come non potrebbe essere condannato a corrispondere l'indennità eventualmente dovuta dall'incapace alle condizioni richieste dall'art. 2047, **comma 2, c.c.** Rispetto a quest'ultima disposizione, infatti, si potrebbe immaginare l'agente

digitale come una sorta di soggetto solo parzialmente capace, così da estendere ai danni cagionati dal medesimo il criterio di imputazione dell'obbligo risarcitorio al sorvegliante dell'incapace. Al riguardo, si è condivisibilmente obiettato (SCOGNAMILGIO) che, alla forzatura logico-sistematica consistente nel qualificare come soggetto l'agente digitale, si somma l'ulteriore forzatura che applica al medesimo una nozione, quale quella di incapacità parziale – o, specularmente, di capacità limitata – che, nel disegno del codice civile, non può proprio prescindere dal sostrato naturalistico della persona umana. E questo anche senza volersi soffermare sulle difficoltà alle quali condurrebbe un'ipotesi di estensione alla materia della responsabilità dell'agente digitale della previsione dell'art. 2047, comma 2, c.c.

3.1 *Segue: ... artt. 2049, 2051, 2052 c.c.*

Per altri (MONTEROSSÌ), invece, la responsabilità *de qua* potrebbe essere regolata dalla norma di cui all'art. **2049 c.c.**, considerata regola non solo utile a risolvere, sul piano tecnico, i problemi posti dall'intensificarsi della delega di attività ai sistemi di IA, ma capace, al contempo, di garantire una coerenza, sul piano storico-evolutivo, con la sua tradizionale funzione in chiave di politica del diritto.

Diverse **le ragioni** che, secondo questa parte della dottrina, riposano **a favore** di questa soluzione.

Sul **piano storico-evolutivo**, innanzitutto, quella regola costituisce l'esito di un processo di **rielaborazione** sostenuto dalla necessità di **rafforzare la tutela dei consociati** contro il maggior **pericolo di danno associato** all'affermarsi di una società sempre più complessa, che imponeva agli **attori economici** di servirsi con maggiore frequenza del **contributo di lavoratori e ausiliari**, per espletare le proprie attività artigianali, produttive e di commercio. Proprio per soddisfare questa finalità di tutela del danneggiato secondo l'attuale interpretazione ormai pacifica, **l'addebito di responsabilità poggia sul "solo fatto della altrui utilizzazione strumentale"**, per il riscontro del quale – secondo la giurisprudenza di legittimità – non è necessario qualificare giuridicamente il rapporto tra l'autore materiale dell'illecito e il soggetto "fatto responsabile" (ad es. in termini di lavoro subordinato), essendo sufficiente dimostrare che per volontà di un soggetto (committente), un altro (commesso) esprima un'attività per conto del primo. Parimenti estensiva è l'interpretazione che si dà alla locuzione «nell'esercizio delle incombenze», essendo detto requisito soddisfatto ogniqualvolta sia possibile registrare un rapporto di "occasionalità necessaria" tra il fatto illecito e l'esercizio delle mansioni (rapporto che si esclude laddove il preposto agisca per finalità private). Quanto, infine, al presupposto del «fatto illecito», si evidenzia che, di regola, la responsabilità del committente è sancita anche qualora l'ausiliario non sia personalmente imputabile (art. 2046 c.c.) e finanche laddove non sia stato possibile individuare a quale soggetto, tra i plurimi ausiliari di cui si sia servito il committente, il danno fosse materialmente addebitabile.

La **norma**, quindi, **mira a "fare responsabile" colui che** si trova nella posizione migliore per **confrontare i benefici derivanti dall'avvalersi di energie e capacità altrui** per svolgere un'attività che presenti un minimo di continuità e dal quale intende trarre un qualche vantaggio, **con i costi (tra cui quelli dei possibili**

danni a terzi) che tale contributo può comportare; sì da decidere, sulla scorta di tale analisi, secondo quali modalità, con quali cautele, per quali incombenze e con quale frequenza avvalersi di un determinato ausiliario e/o lavoratore. D'altra parte, il soggetto candidato responsabile coincide, di regola, con quello che direttamente – o tramite un'organizzazione d'impresa – dovrebbe meglio conoscere il tenore delle incombenze da assegnare e le caratteristiche dell'ausiliario o del lavoratore a cui intende affidarle.

Sulla scorta di questi preliminari rilievi la dottrina in esame trae la conclusione per la quale tanto l'esigenza di *policy*, che sta dietro il funzionamento dell'art. 2049 c.c., tanto il dato tecnico, che presiede alla struttura della norma, offrono significativi argomenti nel senso di provare ad **adeguare il sistema risarcitorio in relazione al danno da IA lavorando sulla regola di responsabilità “vicaria”**, nel presupposto che il **ricorso all'assistenza di una macchina autonoma e ad autoapprendimento possa essere assimilato all'impiego di un ausiliario umano**, laddove tale assistenza comporti un danno a terzi.

A questa prospettazione si è, **in senso contrario**, contestata (GENTILI, SCOGNAMIGLIO), *in primis*, la **qualificazione come soggetto di un software** che sembra pensare e agire; in secondo luogo, si è obiettato che se è vero che **la norma** accolla al *dominus* una responsabilità oggettiva sulla quale nel caso di uso dell'IA non ci sono molti dubbi, tuttavia, **richiede che la condotta dell'ausiliario** sia accompagnata **da colpa o dolo, inconcepibili per il sistema di IA** (come si è rilevato allorché si è esclusa l'operatività dell'art. 2043 c.c.).

Si potrebbe replicare, a questa critica, che è stato ormai da tempo consumato, anche nell'elaborazione giurisprudenziale della materia, il passaggio che attribuisce all'espressione «fatto illecito», contenuta nell'art. 2049 c.c., semplicemente il significato di «fatto produttivo di responsabilità» (così DE MENECH): passaggio che potrebbe accreditare l'assunto che il criterio di ascrizione della responsabilità al preposto, necessario per affermarne la propagazione al preponente, possa ormai integralmente prescindere da una valutazione in termini di colpevolezza della posizione del medesimo preposto. Riteniamo, tuttavia, condivisibile l'**obiezione** di fondo mossa, da parte di autorevole dottrina (GENTILI, SCOGNAMIGLIO), al tentativo di **costruire come «soggetto» il dispositivo**, ancorché «animato», per così dire, dalla applicazione di intelligenza artificiale.

Altri, ancora, hanno suggerito di richiamare la norma di cui l'**art. 2051 c.c.** per applicare al caso il principio della responsabilità dell'utilizzatore per il sistema di IA che egli ha «in custodia», salvo che esso produca danno «per caso fortuito».

Anche questa **ricostruzione** è stata **criticata** (MONATERI). Si è, infatti, rilevato che l'art. 2051 presuppone una **cosa** dotata di una **propria attitudine** materiale a **generare danno**, una potenzialità insidiosa che il custode può prevenire mediante un controllo ordinario. La **ratio** dell'istituto si fonda sull'idea che la **cosa** sia **stabile**, fisicamente determinata, e che il **custode** possa effettivamente **vigilare sul suo stato**, eliminandone vizi e condizioni di pericolo. Orbene – secondo questa impostazione – **l'intelligenza artificiale non presenta alcuna di queste caratteristiche**. Non è una «cosa» nel senso tradizionale del termine, perché non esprime un comportamento fisico né un'insidia materiale. È un insieme di processi informatici che producono

risultati variabili in funzione dei dati introdotti e del contesto operativo. Senza contare che il **presupposto** della **custodia** richiede un **potere di governo continuativo** che consenta al **custode** di **prevenire** e **dominare** l'**insidia** della cosa. Tale potere è **assente** nei **produttori** o **progettisti** dell'**AI**, i quali, una volta realizzato il sistema, non dispongono più di alcun controllo effettivo sugli usi cui esso verrà destinato. L'attività di vigilanza necessaria per fondare la custodia non può essere esercitata a distanza, né può estendersi alle infinite modalità con cui l'**AI** può essere impiegata nella pratica. Il produttore non può prevedere i dati immessi, l'ambiente operativo, gli scopi perseguiti, né gli eventuali errori dell'utilizzatore. La **potenzialità dannosa** non risiede, pertanto, **nello strumento, ma nell'uso** che ne viene **fatto**. Non si dà, in altre parole, un autonomo "stato di pericolo" imputabile al sistema in sé, né una condizione oggettiva che il custode potrebbe eliminare con un intervento materiale. Secondo questo pensiero, in sostanza, attribuire all'**AI** una responsabilità oggettiva o pseudo-oggettiva non apparirebbe conforme alle norme codicistiche. Al contrario, rischierebbe di introdurre una distorsione: presume che la macchina abbia un grado di autonomia tale da potere "volere" il danno. Ma, in realtà, l'**AI** è priva di qualsiasi volontà, producendo solo *output*. L'autonomia che le viene attribuita è considerata una finzione retorica, necessaria per descrivere fenomeni complessi, ma non utilizzabile per fondare una responsabilità giuridica. Nel momento in cui la responsabilità diventa finzione – si contesta –, si ottiene una doppia perdita: non si incentiva la prevenzione e si rallenta lo sviluppo tecnologico.

Altri, ancora, hanno proposto di ricorrere alla responsabilità disegnata dall'art. 2052 c.c. equiparando l'agente digitale all'animale: anche tale ricostruzione non è andata esente da **critiche** (SCOGNAMIGLIO), laddove vorrebbe **sovrapporre un'intelligenza** che, come quella «artificiale», è comunque il **frutto della creazione dell'uomo, a un'intelligenza di origine naturale**: tenendo conto, infatti, di questa radicale differenza di referenti fattuali dell'una e dell'altra questione la via dell'interpretazione analogica, o anche soltanto di quella estensiva, appare senz'altro preclusa.

3.2 *Segue: ... art. 2050 c.c.*

Infine, un'ipotesi di inquadramento plausibile delle questioni di responsabilità civile derivanti dall'esercizio di sistemi di intelligenza artificiale è stata, invece, giudicata dalla dottrina prevalente (*ex pluris* SCOGNAMIGLIO, GENTILI, SCOTTO DI CARLO) quella che valorizzi una serie di elementi. Si richiama (SCOGNAMIGLIO), innanzi tutto, l'elemento del **rischio particolarmente qualificato** che si ricollega all'uso delle applicazioni di intelligenza artificiale, tanto che – come si è visto – è proprio sul profilo del rischio che si fonda l'approccio del legislatore europeo in materia; questa notazione è già di per sé sufficiente a indirizzare verso la disposizione che, nel disegno del codice civile, esprime appunto un criterio di imputazione basato sulla particolare rischiosità di una certa attività, e cioè l'**art. 2050 c.c.** In questa prospettiva, è considerato del tutto lineare il passaggio argomentativo che consideri **pericolosa**, quanto meno **per la natura dei mezzi adoperati, l'utilizzazione di sistemi di autoapprendimento** del dispositivo di intelligenza artificiale, recanti le caratteristiche, evidenziate in precedenza, di **imprevedibilità** nella produzione degli *output* origina-

ti dal corrispondente algoritmo. È, del resto, ormai consolidato da decenni l'indirizzo giurisprudenziale che seleziona le attività pericolose, ai fini dell'applicazione del criterio di imputazione appena richiamato, sulla base di una valutazione calibrata, assai opportunamente, sull'evoluzione della scienza e della tecnica, con il corollario che un'attività in un dato momento storico addirittura inconfigurabile – come sono state appunto, fino ad anni recenti, le applicazioni dei dispositivi di intelligenza artificiale – possa poi in effetti venire ad esistenza e svolgersi, assumendo le caratteristiche di pericolosità tali da fondare il criterio di imputazione dell'art. 2050 c.c., che potrebbe successivamente perdere nel momento in cui un ulteriore progresso della tecnica consenta di reputare superato il presupposto della pericolosità.

Un **altro elemento** ritenuto a favore dell'applicazione della e che, a sua volta, pare confermare la condivisibilità dell'impostazione appena ipotizzata, è quello che si ricollega al fatto che qui i **profili di rischio**, e dunque di pericolo di danno, si ricollegano **non tanto a un oggetto** – sia pure «animato» dall'algoritmo – **ma all'attività stessa di impostazione dell'algoritmo ovvero di sua «manutenzione»**, se così si può dire, in relazione a eventuali evoluzioni del medesimo che richiedano interventi, appunto al fine di prevenire il concretizzarsi di accadimenti dannosi. Proprio per questa ragione, **il criterio di imputazione previsto dall'art. 2050 c.c. è ritenuto più calibrato sulle** caratteristiche delle **applicazioni di intelligenza artificiale**, e dei **rischi di danno** che le stesse traggono con sé.

Ancora, è individuata **un'altra ragione sistematica**: la Proposta di direttiva (come vedremo nel paragrafo successivo) da ultimo richiamata si muoveva dichiaratamente all'interno dell'orizzonte di una **responsabilità per colpa**, in modo non diverso da quella disegnata dall'art. 82, §3, Regolamento *privacy* UE 679/2016. Si aggiunge (SCOGNAMIGLIO), inoltre, a sostegno della soluzione condivisa, che se è certamente vero che anche quello contenuto nell'art. 2050 c.c. è ritenuto prevalentemente un criterio di imputazione di responsabilità oggettiva, è anche vero che questo esito applicativo risulta essere assai più il frutto del modo, particolarmente rigoroso, in cui viene inteso, nella elaborazione giurisprudenziale della materia, l'onere della prova liberatoria a carico dell'esercente l'attività pericolosa, che non la conseguenza di una scelta del legislatore, immediatamente desumibile dall'enunciato della disposizione: e infatti il riferimento alla prova di avere adottato tutte le misure idonee a evitare il danno sembra effettivamente evocare più un modello di responsabilità dove la colpa del candidato responsabile comunque rilevi, sia pure con un onere della prova invertito rispetto a quello della norma generale dell'art. 2043 c.c. (su questo, tuttavia, come visto *supra*, in questa Parte, Sez. I, Cap. 4, §3, riteniamo di non condividere questa ricostruzione in riferimento alla natura giuridica della responsabilità ex art. 2050 c.c.).

Ancora, secondo altra parte della dottrina (SCOTTO DI CARLO), **la norma trova applicazione** in quanto consente di ritenere **pericolosa**, tra le varie attività, quella di **produzione di macchine a elevata automazione, in ragione del rischio di fallibilità** del prodotto. Ciò permette di aggiungere alla responsabilità del proprietario e del conducente di un veicolo autonomo, per il danno cagionato dalla circolazione dello stesso ai sensi dell'art. 2054 c.c. (così anche COPPINI), anche la responsabilità

del produttore (si rinvia *funditus* a quanto riferito *supra*, nella Sez. I di questa Parte, Cap. 8, §3).

Sempre secondo questa linea di pensiero, **la responsabilità disciplinata dall'art. 2050 c.c.** sarebbe inoltre, **preferibile rispetto a quella connessa al difetto del prodotto**, come attualmente vigente (artt. 114 e ss. cod. cons.). Infatti, la stessa nozione di difetto, che fa leva sulla “sicurezza che ci si può legittimamente attendere” ai sensi dell'art. 117 cod. cons., è stata ideata assumendo come parametro di riferimento prodotti seriali e statici, sicché appare inadeguata a descrivere la condotta di una res che agisce in autonomia, il comportamento della quale, d'altronde, potrebbe discendere da una causa che non consiste in un difetto. Peraltro, anche quando l'elaboratore rientrasse nel novero dei beni seriali, lo stesso potrebbe non permanervi per l'intera durata della sua vita, potendo la cosa evolversi e, quindi, cambiare la propria condotta in base all'esperienza maturata nel contesto in cui viene immessa.

Assumendo come modello un **dispositivo medico dotato di intelligenza artificiale**, è stato osservato (CIONI) che **questo potrebbe commettere errori**, anche potenzialmente fatali, in virtù del proprio meccanismo di apprendimento. Pertanto, **la normativa sul danno da prodotto difettoso**, che, da un lato, è **incentrata sulla nozione di difetto**, e, dall'altro, **si disinteressa della condotta post-market del produttore**, sarebbe **incapace di offrire una tutela effettiva ai soggetti danneggiati**; e ciò anche perché l'elaboratore sarebbe in grado di produrre danni sconosciuti e imprevedibili al tempo della commercializzazione, per cui il produttore sarebbe esente da responsabilità perché giustificato dal rischio di sviluppo. Infatti, l'art. 118, comma 1, lett. e), cod. cons. esclude la responsabilità del produttore se lo stato delle conoscenze, sia tecniche sia scientifiche, al momento in cui il fabbricante ha messo in circolazione il prodotto, impediva di concepirlo come difettoso.

In merito al ricorso all'art. 2050 c.c., si è **obiettato** (MONATERI) che **l'AI non possa essere qualificata come attività pericolosa perché non presenta un rischio intrinseco e autonomo**: il **pericolo** nasce solo dal **modo in cui** viene **utilizzata**, dal contesto operativo e dai dati immessi dall'utente. Di conseguenza, applicare il 2050 c.c. significherebbe **forzare l'istituto**, spostando la responsabilità su soggetti – produttori e progettisti – che non hanno il controllo del rischio concreto.

4. La abortita proposta di direttiva sulla responsabilità extracontrattuale dell'IA lascia un vuoto normativo a livello comunitario

Il testé denunciato vuoto normativo sembrava prossimo a essere colmato da una **proposta di direttiva del Parlamento europeo e del Consiglio** sull'adeguamento delle norme in materia di **responsabilità civile extracontrattuale all'IA**, che riguardava le azioni giudiziarie promosse a livello nazionale, che implicano l'accertamento della colpa di un soggetto. In questo modo veniva garantito il risarcimento di qualsiasi tipo di danno a beneficio di qualunque danneggiato.

La normativa **non disciplinava gli aspetti generali della responsabilità civile**, i quali restavano affidati alle legislazioni nazionali, come le definizioni di colpa e di causalità, ma **indicava le diverse tipologie di pregiudizio** idonee a **giustificare il ricorso all'autorità giudiziaria**, la **distribuzione della responsabilità tra più**

danneggianti, il **concorso di colpa**, la **quantificazione del danno** e i termini di prescrizione, che consentiva l'integrazione delle misure previste dalla direttiva nei sistemi di responsabilità civile esistenti "senza attriti" (SCOTTO DI CARLO).

La **ratio legis** della proposta si fondava sulla necessità di garantire una **tutela efficace** alle vittime e di stabilire **criteri univoci per la ripartizione del rischio** tra operatori, sviluppatori e utilizzatori.

L'impianto della direttiva prevedeva **tre elementi fondamentali** per colmare le lacune della normativa vigente.

Il primo concerneva l'introduzione di una **presunzione di colpa** nei confronti degli operatori responsabili dei sistemi di intelligenza artificiale classificati come ad alto rischio. Questa previsione si giustificava alla luce della difficoltà di dimostrare l'esistenza di una condotta colposa in relazione a un sistema che opera in autonomia, manlevando la vittima di un gravoso onere probatorio.

Il secondo elemento riguardava **l'accesso alle prove**, mediante misure che avrebbero agevolato chi subisce un danno nell'acquisizione delle informazioni necessarie a dimostrare il nesso tra il malfunzionamento dell'intelligenza artificiale e l'evento dannoso.

L'ultimo elemento tendeva all'**armonizzazione della disciplina negli Stati membri**, con l'obiettivo di evitare una regolazione frammentata che potesse determinare disuguaglianze nelle condizioni di tutela giuridica e asimmetrie competitive nel mercato unico.

La Commissione ha posto a fondamento della propria decisione un'argomentazione sistemica per la quale l'assetto delineato dalla proposta di direttiva si poneva in rapporto con altre iniziative normative europee in materia di intelligenza artificiale, ma si distingueva per la sua specifica finalità di regolare la responsabilità *ex post*.

La Commissione europea ha, tuttavia, nel 2025, **ritirato la proposta, motivando** tale decisione con l'obiettivo di **evitare un eccesso di regolamentazione** che avrebbe potuto incidere negativamente sulla competitività delle imprese europee.

L'adozione di criteri più stringenti per l'imputazione della responsabilità in caso di danni derivanti dall'impiego di sistemi autonomi avrebbe comportato un incremento degli oneri regolatori, con **ripercussioni sullo sviluppo tecnologico e sull'attrattività del mercato** europeo per gli investitori.

L'intelligenza artificiale rappresenta un **settore strategico**, e la Commissione ha ritenuto che un **quadro normativo più rigido** avrebbe potuto **rallentare l'adozione** di nuove soluzioni basate su **sistemi autonomi**, con un **impatto** diretto sulla **capacità** dell'industria europea di **competere** con le grandi piattaforme globali.

5. La nuova disciplina per la responsabilità per i danni prodotti dall'intelligenza artificiale introdotta dal D.Lgs. 9 settembre 2026, n. 160

Come visto nei paragrafi precedenti, a livello interno si è, in un primo momento, cercato di colmare il vuoto normativo in materia, ricorrendo alle norme dettate dal codice civile in tema di responsabilità che sostanzialmente alleggeriscono l'onere probatorio per il danneggiato. Orbene, venendo incontro a queste esigenze, il **Gover-**

no ha adottato il **D.Lgs. 9 settembre 2026, n. 160**, adeguando l'ordinamento interno all'AI Act europeo (Regolamento UE 2024/1689).

In particolare, il Capo II del Titolo II dà attuazione alla delega contenuta nell'art. 24, commi 3 e 5, lett. d), L.132/2025, dettando **disposizioni** processuali civili in materia di **risarcimento** dei danni cagionati dall'utilizzo dei medesimi sistemi. **Tre** sono le **finalità** fondamentali perseguite dalle norme ivi contenute: a) **garantire** al **danneggiato** strumenti effettivi di **accesso** alla **prova**, in considerazione dell'opacità tecnologica che caratterizza i sistemi di intelligenza artificiale e della tipica asimmetria informativa tra utilizzatore e fornitore (art. 17); b) **alleggerire l'onere probatorio** sul **nesso** di **causalità**, attraverso una **presunzione** modulata in funzione della classificazione di rischio del sistema operata dal regolamento (UE) 2024/1689 (art. 18); c) **rafforzare** la **tutela** del **danneggiato** persona fisica, mediante la previsione di un **foro** speciale del **consumatore** e di un'**azione diretta** nei **confronti** dell'impresa di **assicurazione** che presta la copertura della responsabilità civile (artt. 16 e 20).

L'intervento è costruito secondo una **logica di complementarità**, e non di **sostituzione**, rispetto agli statuti speciali di responsabilità già operanti nell'ordinamento. Sono fatte **espressamente** salve le **disposizioni** dell'art. 82 del **regolamento** (UE) 2016/679, sulla protezione dei dati e la normativa nazionale di recepimento della direttiva (UE) 2024/2853 (di cui si è riferito *supra*, in questa Parte e Sezione, Cap. 4, §3.1.1), in tal modo **evitando** ogni effetto di **restrizione** delle **tutele esistenti** negli ambiti specificamente regolati dal diritto eurounitario (art. 16, comma 3).

5.1 L'accesso alle prove

In merito alla prima delle tre finalità evidenziate nel paragrafo precedente, non diversamente da quanto previsto dall'art. 9 della direttiva UE 2024/2853 e dal nuovo art. 119 cod.cons., come riscritto nello schema di decreto legislativo approvato dal Governo il 4 agosto 2026, in tema di responsabilità per danni da prodotti difettosi (di cui si è riferito *supra*, in questa Parte e Sezione, Cap. 4, §3.3.2), l'art. 17 D.Lgs. 160/2026 ha introdotto un **meccanismo di esibizione** delle **prove** volto a **superare l'asimmetria informativa** nei contenziosi relativi ai sistemi di IA, consentendo al **giudice** di **ordinare l'accesso** a documenti e dati tecnici pertinenti quando il **danneggiato dimostri la verosimiglianza** della propria **domanda**. La norma individua le principali categorie di documentazione rilevante (*logging*, gestione dei rischi, documentazione tecnica e sorveglianza umana), imponendo il rispetto dei principi di necessità, proporzionalità e tutela dei segreti commerciali. Sono inoltre previste conseguenze processuali e sanzioni in caso di mancata ottemperanza all'ordine di esibizione, al fine di garantire l'effettività della tutela risarcitoria.

In particolare, la disposizione attribuisce al giudice il potere di ordinare, su istanza della parte che allega di aver subito il pregiudizio, l'**esibizione di elementi di prova specificamente pertinenti** relativi al funzionamento del sistema di intelligenza artificiale. Tale ordine può essere **diretto** non solo alla controparte processuale, ma **anche** a un **terzo che ne disponga**, assicurando così una *disclosure* probatoria ampia e funzionale alla ricostruzione dei fatti. L'attivazione di questo meccanismo istruttorio è subordinata a un preciso onere di allegazione in capo all'istante, il quale deve