

L'errore nella decisione nell'era dell'intelligenza artificiale (raccolta di contributi)

Ermanno CALZOLAIO*

Sommario: 1. Premessa 2. Attuali tendenze del diritto europeo (UE) 3. Raccolta di contributi

1. Premessa

Vengono raccolti in questo fascicolo alcuni contributi presentati in occasione di un incontro di studio, intitolato "L'errore nella decisione nell'era dell'intelligenza artificiale"¹; svoltosi nel contesto di una più ampia riflessione sulle implicazioni degli impieghi delle nuove tecnologie sull'autonomia del processo decisionale².

È noto che il ricorso all'uso di algoritmi di intelligenza artificiale è diffuso in una estesa varietà di ambiti, incluso quello giuridico. Uno dei profili più significativi della profonda trasformazione in atto risiede nella progressiva automazione dei processi decisionali e nella conseguente erosione dell'autonomia dei singoli soggetti, privati o pubblici: sia quando la decisione viene assunta sulla base dei dati risultanti dall'applicazione dei sistemi di intelligenza artificiale; sia quando l'intervento dell'uomo è escluso in radice, perché è la decisione stessa a prodursi in modo automatizzato. In questi casi venendo meno la possibilità di imputare gli effetti eventualmente dannosi di una condotta all'autonoma volontà del soggetto responsabile della condotta medesima.

Tra i tanti aspetti che sollecitano la riflessione, vi è quello connesso ad uno dei motivi ricorrenti che hanno accompagnato la diffusione dei sistemi di intelligenza artificiale, il cui impiego contribuirebbe in modo determinante a scongiurare la possibilità di errori. Senonché, l'esperienza mostra che, per un verso, il rischio di errore può essere forse ridotto ma certamente non eliminato del tutto e che, per altro verso, nuove possibilità di errori discendono proprio dall'uso di questi sistemi. Si può trattare di 'vizi' di programmazione o di funzionamento dell'algoritmo, ma anche di 'vizi' nell'interpretazione dei dati o delle regole rilevanti, con la conseguente esigenza di accertare se, nell'uno o nell'altro caso, essi siano imputabili (anche) all'autore dell'algoritmo.

Per avvedersi della portata del problema è utile evocare qualche esempio concreto. Nel mese di settembre 2021, il New York Times³ ha riportato la notizia che alcuni utenti di *Facebook* avevano visualizzato un video realizzato da un *tabloid* britannico che mostrava uomini di colore, ricevendo nei giorni successivi un messaggio automatico che chiedeva se l'utente fosse ancora interessato a visionare video di "primati". *Facebook* è stata costretta a disabilitare il messaggio inviato automaticamente e il portavoce del noto *social network* si è affrettato a dichiarare che, nonostante i miglioramenti apportati, il sistema non è perfetto, scusandosi dell'errore. Ciò mostra che ogni sistema di IA è fallibile, con effetti che possono rivelarsi molto gravi anche sul piano della

* Professore ordinario di Diritto privato comparato, Università di Macerata.

¹ "L'errore nella decisione nell'era dell'intelligenza artificiale", Seminario internazionale a conclusione del progetto di ricerca "Dant", Macerata, 24 settembre 2021.

² Cfr. E. Calzolaio (cur.), *La decisione nel prisma dell'intelligenza artificiale*, Padova, 2020.

³ "NYT", 3 settembre 2021, *Facebook Apologizes After A.I. Puts 'Primates' Label on Video of Black Men*.

discriminazione e della tutela della persona. A questo riguardo, è significativo quanto accaduto in occasione delle paraolimpiadi di Tokio, sempre nel mese di settembre 2021. Un veicolo senza conducente, condotto mediante il ricorso ad un sistema di intelligenza artificiale, ha investito un pedone all'interno del villaggio olimpico e il direttore generale della casa produttrice si è affrettato a scusarsi, non esitando a dare atto che l'incidente ha messo in luce l'errore del sistema, che non è in grado di riconoscere persone affette da disabilità (nella specie visive).

Anche le professioni giuridiche fanno ricorso a sistemi di IA. È noto che alcuni studi legali specializzati in operazioni di *Merger & Acquisition* (M&A) si sono dotati di sistemi per facilitare la complessa attività di *due diligence*, che sovente coinvolge soggetti provenienti da diversi Paesi e richiede di esaminare una mole enorme di documenti, catalogandoli per legge applicabile e lingua di redazione. Il sistema "Luminance", ad esempio, consente di svolgere questo tipo di analisi in pochi minuti e con un livello di precisione elevato. È pur sempre indispensabile una successiva fase di controllo, ma i cui tempi sono molto ridotti⁴. Fino a che punto il professionista può fare affidamento sulle prestazioni del sistema? È ragionevole ipotizzare che sia il professionista a rispondere delle errate valutazioni a cui sia stato indotto dal sistema? Il programmatore può essere ritenuto responsabile? Analoghe questioni si pongono rispetto alla prefigurata diffusione di sistemi che consentono di agevolare addirittura l'interpretazione dei contratti⁵.

Sistemi di intelligenza artificiale sono poi diffusi nel settore degli investimenti finanziari. Un recente caso giunto all'esame della *High Court* inglese⁶ ha ad oggetto la domanda per il risarcimento dei danni conseguenti alla perdita di investimenti, a seguito dell'impiego di processi decisionali automatizzati tramite algoritmi. Il sistema in questione, denominato K1, opera attraverso un algoritmo di *machine learning*, alimentato da una molteplicità eterogenea di dati (quali notizie in tempo reale e *post* sui *social media*), in grado di individuare correlazioni e schemi ricorrenti tra i dati da cui inferire previsioni sul "sentiment" degli investitori e sull'andamento delle borse. Sulla base delle correlazioni ottenute, il sistema invia ai *broker* istruzioni sulle operazioni da compiere, aggiustando in tempo reale la propria strategia di investimento grazie all'apprendimento continuo dell'algoritmo. Nel caso di specie, l'investitore ha agito in giudizio nei confronti della società di investimenti a cui aveva affidato la gestione di una parte del proprio capitale, chiedendo un risarcimento del danno per le perdite subite, lamentando una rappresentazione errata ed iperbolica delle capacità del sistema K1 da parte della società stessa. L'investitore ha invocato ciò che nel nostro ordinamento potrebbe essere assimilabile a un vizio nella formazione della volontà negoziale (dolo), o in alternativa una forma di responsabilità precontrattuale, intesa come reticenza nelle informazioni fornite in sede di trattative. Un sistema come quello K1 è annoverato tra i mezzi di IA "debole", basato su algoritmi di *machine learning* evoluti, ma privi di capacità di autodeterminazione e di comprensione delle informazioni processate, che operano a supporto dell'uomo. Tuttavia, la capacità di apprendimento automatico conferisce a questi sistemi un grado di autonomia crescente, mettendoli in condizione di assumere decisioni che producono effetti sul mondo esterno.

Sono pure utilizzati sistemi di prevenzione del crimine mediante mezzi di sorveglianza elettronica. A questo proposito, è utile ricordare quanto avvenuto nel gennaio 2020 a Detroit, quando un uomo di colore è stato arrestato per un crimine che non aveva commesso, a causa di un errore di un sistema di riconoscimento facciale. Secondo il *New York Times*, si è trattato del primo caso conosciuto in cui un cittadino americano è stato arrestato per un errore dell'algoritmo⁷. Questa vicenda illustra bene la storia dei nostri tempi, in cui vengono utilizzati sistemi di intelligenza artificiale per monitorare, tracciare, profilare e quindi determinare la possibile innocenza o

⁴ Il "Sole 24ore", 1 marzo 2018 (https://www.ilsole24ore.com/art/perche-abbiamo-assunto-software-nostro-studio-legale-italiano-AEPmq58D?refresh_ce=1), intervista a Jan Pecoraro.

⁵ Cfr. ampiamente R. Catterwell, *Automation in contract interpretation*, in *Law, Innovation & Technology*, 2020, p. 81 ss.

⁶ *Tyndaris v MMWWW Ltd* [2020] EWHC 778 (Comm), 22 April 2020 (<http://www.bailii.org/ew/cases/EWHC/Comm/2020/778.html>).

⁷ "NYT", 3 agosto 2020, *Wrongfully Accused by an Algorithm*.

colpevolezza. Ciò non avviene solo negli Stati Uniti, ma anche in Europa, in cui la logica algoritmica si è ormai diffusa nei contesti di giustizia penale o di prevenzione del crimine⁸.

Si potrebbe continuare a lungo, ma ciò che interessa porre in luce è che, a fronte della crescente richiesta di progettazione di sistemi di intelligenza artificiale di “machine learning”, la stessa letteratura scientifica riconosce ormai apertamente che si tratta di sistemi per definizione fallibili, che possono produrre risultati non attendibili, con effetti rilevanti sul piano di possibili ingiustizie e di errori di vario genere⁹.

2. Attuali tendenze del diritto europeo (UE)

Alla luce del quadro sin qui tratteggiato, non stupisce che il legislatore europeo è da tempo impegnato nella ricerca di soluzioni normative, come testimonia, da ultimo, la proposta di Regolamento Europeo sull'intelligenza artificiale del 21 aprile 2021¹⁰. Essa poggia su tre pilastri: la tutela della persona, la responsabilità per i danni derivanti dall'uso delle nuove tecnologie e la tutela del diritto d'autore. L'approccio seguito è basato sulla identificazione dei rischi, secondo una graduazione da alto a basso rischio, fino a vietare alcune pratiche di intelligenza artificiale particolarmente dannose e comunque a restringere determinati usi dei sistemi di identificazione biometrica remota. Vengono previsti analitici meccanismi di controllo dell'attività, con riferimento in particolare ai processi di intelligenza artificiale realizzati e distribuiti come “prodotto” o inseriti in altri prodotti o utilizzati dagli operatori del mercato. Per contro, la proposta di Regolamento non si occupa direttamente dei macchinari, né dei robot, rispetto ai quali manca tuttora una definizione in ambito europeo.

Il tentativo del legislatore europeo va verso una regolamentazione in chiave ‘antropocentrica’ dell'intelligenza artificiale, perseguito attraverso la prevenzione di pratiche lesive per la persona (come le tecniche subliminali), il controllo preventivo per l'immissione di tecniche ad alto rischio, la cooperazione istituzionale tra Stati e tra le autorità preposte¹¹. Questa visione antropocentrica è stata di recente confermata dalla proposta della Commissione di “Dichiarazione europea sui diritti e principi digitali” del 26 gennaio 2022¹².

La proposta di Regolamento fa parte di un pacchetto più ampio di misure destinate ad affrontare i problemi posti dallo sviluppo e dall'utilizzo dell'IA, tra le quali le iniziative in corso volte a rivedere la normativa settoriale sui prodotti e quelle più direttamente orientate alla disciplina delle questioni connesse alla responsabilità in relazione alle nuove tecnologie. “Tali iniziative – si legge nella relazione che accompagna il testo - si baseranno sulla presente proposta

⁸ Cfr. A. Završnik, *Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings*, in *European Journal of Criminology*, 2019 (<https://journals.sagepub.com/doi/full/10.1177/1477370819876762>).

⁹ Cfr. da ultimo J. Skripchuk-Y. Shi-T. Price, *Identifying Common Errors in Open-Ended Machine Learning Projects* (<https://dl.acm.org/doi/abs/10.1145/3478431.3499397>). Nel settore medico, cfr. R. Challen-J. Denny-M. Pitt, L. Gompels-T. Edwards-K. Tsaneva-Atanasova, *Artificial Intelligence, Bias and Clinical Safety* (<https://qualitysafety.bmj.com/content/28/3/231>): cfr. altresì A. E. Evstratov-I. Y. Guchenkov, *The Limitations of Artificial Intelligence (Legal Problems)*, in *Law Enforcement Review*, 2020, p. 13 ss. Sul tema dei cd. dispositivi medici intelligenti, cfr. E. Calzolaio, *I dispositivi medici “intelligenti”: spunti di comparazione giuridica*, in *Foro it.*, 2022, fasc. 2, p. 75 ss.

¹⁰ La proposta di Regolamento è reperibile all'indirizzo <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52021PC0206&from=IT>.

¹¹ Per una accurata disamina, cfr. G. Alpa, *Quale modello normativo europeo per l'intelligenza artificiale*, in *Contratto e Impresa*, 2021, pp. 1003 ss. V. altresì G. Resta, *Cosa c'è di 'europeo' nella proposta di Regolamento UE sull'intelligenza artificiale*, in *Dir. Inf.*, 2022, p. 323 ss.

¹² La European Declaration on Digital Rights and Principles for the Digital Decade, 26/1/2022, COM(2022) 28 final, insieme al relativo documento di lavoro (Staff working document), sono reperibili all'indirizzo <https://digital-stra.tegv.ec.europa.eu/en/library/declaration-european-digital-rights-and-principles>. Sull'orientamento antropocentrico della proposta di Regolamento, v. già il documento della Commissione, dal titolo “*Creare fiducia nell'intelligenza artificiale antropocentrica*”, Bruxelles, 8.4.2019, COM(2019) 168 final.

e la integreranno al fine di apportare chiarezza giuridica e favorire lo sviluppo di un ecosistema di fiducia nei confronti dell'IA in Europa" (punto 1.3). Il riferimento riguarda le iniziative assunte sin dal 2017 dalla Commissione e dal Parlamento circa l'opportunità di modificare la Direttiva 85/374/EEC in materia di responsabilità del produttore, per adeguarla ai profili connessi all'impiego dell'intelligenza artificiale nei 'prodotti'.

Nello specifico, i problemi degli effetti dannosi dell'uso delle nuove tecnologie sono stati affrontati da un Gruppo di esperti nel 2019¹³, le cui conclusioni hanno ispirato la Relazione della Commissione del 19 febbraio 2020 "sulle implicazioni dell'intelligenza artificiale, dell'Internet delle cose e della robotica in materia di sicurezza e di responsabilità"¹⁴. Nella relazione la Commissione europea rileva che alcune caratteristiche delle tecnologie digitali emergenti rendono difficile risalire dal danno subito al comportamento umano che giustificherebbe la domanda di risarcimento sulla base della responsabilità per colpa in base alle normative nazionali. Inoltre, la Commissione pone in evidenza che la disciplina dettata dalla direttiva sulla responsabilità del produttore non è adeguata per proteggere i danneggiati da prodotti dotati di intelligenza artificiale.

Dal canto suo, il Parlamento europeo si è spinto fino a proporre l'adozione di un Regolamento sulla responsabilità per il funzionamento dei sistemi di intelligenza artificiale, il cui testo è pubblicato in calce alla Risoluzione del 20 ottobre 2020 recante raccomandazioni alla Commissione¹⁵. Anche questa proposta si basa su un sistema di responsabilità del produttore calibrato in base al rischio creato dagli operatori, a loro volta distinti in varie categorie: quelli ad alto rischio rispondono a titolo di responsabilità oggettiva, mentre quelli a rischio più moderato rispondono per colpa. In particolare, per gli operatori ad alto rischio la responsabilità non è esclusa anche in caso di dimostrazione di aver agito con diligenza, fatta salva la forza maggiore (art. 4). La proposta interviene anche sui termini di prescrizione, che per gli operatori ad alto rischio sono elevati addirittura a trenta anni.

3. Raccolta di contributi

Da quanto sin qui tratteggiato si trae ampia conferma che anche i tentativi di regolamentazione in ambito europeo testimoniano fino a che punto il tema dell'errore negli impieghi dell'intelligenza artificiale costituisca uno dei nodi più delicati nell'ambito del rapporto tra diritto e nuove tecnologie. I saggi raccolti nel presente fascicolo costituiscono un tentativo di indagare alcuni profili su cui il giurista è chiamato ad interrogarsi.

Il primo contributo, di Monica Stronati, colloca la problematica nella prospettiva storico-giuridica, ponendo in luce che non sempre gli ordinamenti giuridici hanno accolto l'ipotesi di errori nelle decisioni e la necessità di una loro riparazione. Nonostante l'art. 24 Cost. riconosca il diritto soggettivo alla riparazione degli errori giudiziari, l'impianto complessivo del nostro ordinamento continua ad essere ispirato all'ideale di una giustizia quale "regno del razionale, possibilmente calcolabile". Gli impieghi dell'intelligenza artificiale rischiano di rafforzare la falsa idea che il superamento dell'elemento umano, emotivo, possa rendere la giustizia certa e infallibile.

La riflessione di Frédéric Dourneaux, soffermandosi anche sulle implicazioni delle recenti iniziative normative in ambito europeo, si concentra sull'analisi delle varie tipologie di errori che l'intelligenza artificiale può determinare nel processo decisionale, attirando l'attenzione sulla necessità di identificarli con esattezza, al fine di prospettare soluzioni giuridiche per prevenirli e per porvi rimedio.

Dal canto suo, il saggio di Garance Cattalano si concentra sulle conseguenze degli errori prodotti dall'intelligenza artificiale, tanto nella sfera contrattuale, quanto in quella della responsabilità civile.

¹³ La relazione è reperibile all'indirizzo https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=63199.

¹⁴ [COM \(2020\) 64 \(final\)](#).

¹⁵ [2020/2014 INL](#).

In entrambi i casi, la ricerca del soggetto a cui l'errore è imputabile può costituire un limite allo sviluppo delle nuove tecnologie, sicché l'Autrice prospetta l'opportunità di immaginare meccanismi di socializzazione del rischio.

Su questa scia si muove il contributo di Arianna Alpini, che mostra fino a che punto si rivela insufficiente concentrare l'attenzione esclusivamente sull'accertamento della consapevolezza del programmatore circa la fallibilità del sistema, dovendosi considerare invece anche il livello di conoscenza tecnologica dei soggetti coinvolti.

Il saggio di Laura Vagni prende in considerazione un aspetto critico che emerge di pari passo con la diffusione della *blockchain*. Come è noto, l'immutabilità e l'incorruttibilità che caratterizzano questa tecnologia consentirebbero di superare il problema della fiducia tra le parti di una transazione (sistema *trustless*). A ben vedere, tuttavia, la *blockchain* comunica con il mondo esterno al *network* tramite cd. oracoli. L'articolo focalizza l'attenzione sulla rilevanza giuridica dell'errore commesso dall'oracolo, mettendo in evidenza che la *blockchain* non è un sistema che prescinde dalla fiducia, ma promuove una diversa idea di fiducia, basata sul consenso dei nodi, con rilevanti implicazioni sul piano giuridico.

Da ultimo, il contributo di Sirio Zolea affronta il tema dell'errore (e della responsabilità che ne può derivare) nel particolare contesto del diritto dello spazio, ove è larghissimo l'impiego di sistemi di intelligenza artificiale. In particolare, vengono esaminate le regole adottate dall'Agenzia Spaziale Europea, per poi guardare al quadro giuridico in Italia e in Francia, mostrando come la disciplina della materia si presenti molto stratificata e tale da richiedere un approccio aperto e interdisciplinare.

Abstract

Il contributo introduce una raccolta di testi presentati in occasione di un incontro di studio sul tema dell'errore nella decisione nell'era digitale, nel più ampio contesto del rapporto tra diritto e nuove tecnologie, delineandone alcuni tratti e tendenze in ambito europeo, per poi descrivere brevemente i contenuti dei singoli testi raccolti.

Parole chiave: intelligenza artificiale, errore nella decisione, diritto europeo (UE)

*

The contribution introduces a collection of texts presented on the occasion of a study meeting on the topic of error in decision making in the digital age, in the broader context of the relationship between law and new technologies, outlining some traits and trends in the European context, moving then to briefly describe the content of the individual texts collected.

Key words: artificial intelligence, decision error, European (EU) law