

## Una mela acerba e ancora avvelenata: Il futuro della professione forense innanzi a ChatGPT

Umberto Nizza

SOMMARIO: 1. Introduzione. – 2. Costi e benefici associati all'utilizzo di ChatGPT. – 3. Alcune perplessità sull'affidabilità di ChatGPT. – 4. Conclusioni.

### 1. *Introduzione*

L'intelligenza artificiale (IA) è ormai diventata parte integrante della nostra vita quotidiana, assumendo un ruolo sempre più rilevante ed una applicazione diffusa anche nelle scienze sociali e persino nelle scienze dure<sup>1</sup>. I progressi nel campo dei *big data* e dell'apprendimento automatico stanno trasformando profondamente la nostra società in molteplici settori, a partire dall'intrattenimento sino ad arrivare al commercio, con tecnologie rivoluzionarie che permettono a piattaforme come Netflix, Amazon e Google di raccogliere enormi quantità di dati provenienti da utenti in tutto il mondo, analizzare le loro preferenze ed offrire prodotti e servizi mirati.

In questo contesto di ipertrofia tecnologica, ChatGPT<sup>2</sup> ha guadagnato notevole popolarità per la sua capacità di generare risposte di testo simili a quel-

<sup>1</sup> Vi è un grande ottimismo riguardo all'applicazione dell'intelligenza artificiale, in tutti gli ambiti dell'assistenza sanitaria, dalla diagnosi al trattamento. Esistono già numerose evidenze che dimostrano che gli algoritmi di intelligenze artificiali stanno già svolgendo analisi più precise degli umani, in particolare nell'analisi di immagini mediche, nella correlazione tra sintomi e biomarcatori, nonché nella caratterizzazione e la prognosi di malattie, anche rare. Si v., in tal senso, A. Bohr e K. Memarzadeh, *The Rise of Artificial Intelligence in Healthcare Applications*, in A. Bohr e K. Memarzadeh (a cura di), *Artificial Intelligence in Healthcare*, Cambridge, 2020, 25 ss.

<sup>2</sup> ChatGPT è una IA sviluppata da OpenAI – un'organizzazione senza fini di lucro di ricerca sull'intelligenza artificiale – e rilasciata al pubblico nel novembre 2022, basata su modello di linguaggio auto-regressivo con più di 175 miliardi di parametri. Si tratta di una IA che ha vissuto una diffusione a dir poco impressionante: in soli cinque giorni, infatti, è stata capace di raccogliere un milione di utenti. Per il raggiungimento della stessa utenza sono stati necessari 300 giorni per Facebook, 720 per Twitter e tre mesi e mezzo per Instagram. In tal senso, v. M. Firat, *What ChatGPT means for universities: Perceptions of scholars and students*, in *J. Appl.*

le umane. La sua capacità di fornire risposte immediate e pertinenti ha mostrato un notevole potenziale, in particolare nell'assistenza di molte mansioni quotidiane, dalla creazione di sommari alla traduzione di interi testi. Vi sono sicuramente diverse resistenze nell'uso di ChatGPT, che è stato identificato, da alcuni commentatori particolarmente critici, come un «goffo motore statistico per il riconoscimento di schemi, che si nutre di centinaia di *terabyte* di dati ed estrapola le risposte conversazionali più probabili o le soluzioni più probabili a una domanda scientifica», senza alcun tipo di capacità critica e, pertanto, ancora bloccato in una fase preumana dell'evoluzione cognitiva<sup>3</sup>.

In effetti, allo stato attuale, ChatGPT sembra porre più interrogativi<sup>4</sup> che soluzioni pratiche. Tralasciando eventuali corrivi catastrofismi, sorgono spontanee una serie di domande cruciali: Qual è il ruolo dei dati di addestramento e delle interazioni umane nel plasmare il comportamento di questi modelli di linguaggio? Possiamo fidarci di quanto da loro affermato? Siamo in grado, in autonomia, di individuare con esattezza i confini tra ciò che la macchina ci offre come vero e ciò che non lo è? Può una IA come questa essere considerata neutrale? "In che limiti ed in che modo ci si può servire di questa o di altre IA generative nella pratica forense di tutti i giorni?"

È evidente che tutti questi interrogativi inducono una disamina approfondita e critica del modo in cui ChatGPT opera e, in subordine, di come le sue risposte possano influenzare la professione forense e la società nel suo complesso. L'obiettivo di questo articolo è, in tale prospettiva, promuovere una profonda riflessione su questa IA generativa, con il chiaro e dichiarato intento di individuarne possibili limitazioni, criticità e sfide che possono derivare dall'utilizzo di una siffatta macchina. Sebbene ChatGPT offra risposte dettagliate e articolate ad un numero pressoché sterminato di richieste, è importante anticipare che, come si vedrà nel corso di questo manoscritto, non sembra ancora in grado di sostituire integralmente l'*expertise* umana nel campo forense. Nonostante l'incommensura-

---

*Learn. Teach.*, 2023, 1, 57 ss. Esistono altre IA sul mercato (e.g., Gemini, Llama, Chatsonic, Claude), ma in questo manoscritto ci concentreremo soltanto su questo *chatbot* in particolare, data la sua peculiare notorietà e recente attenzione mediatica.

<sup>3</sup> Con queste dure parole si sono espressi Noam Chomsky, Ian Roberts e Jeffrey Watumull sulle pagine del New York Times, suggerendo che fosse al tempo stesso comico e tragico che tanto denaro e attenzione si fossero concentrati su qualcosa di così piccolo ed insignificante. Si v. N. Chomsky, I. Roberts, J. Watumull, *Noam Chomsky: The False Promise of ChatGPT*, 8 marzo 2023, in <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>, ultimo accesso 20 aprile 2024.

<sup>4</sup> Recente è il monito del Segretario *pro tempore* delle Nazioni Unite, António Guterres, che si è spinto ad affermare che l'avvento delle IA generative aggiunge una nuova dimensione alla manipolazione del comportamento umano, con potenziali devastanti su larga scala, sui quali gli stessi creatori hanno avvertito dei rischi potenzialmente catastrofici ed esistenziali. Si v. United Nations, *Secretary-General's remarks to the Security Council on Artificial Intelligence*, 18 luglio 2023, in <https://news.un.org/en/story/2023/07/1138827>, ultimo accesso 20 aprile 2024.

bile potere generativo ed i margini di miglioramento tecnico della macchina, non sembra ancora scalfibile il ruolo essenziale ed irrinunciabile della primaria funzione di difesa – nel caso degli avvocati – e di bilanciata amministrazione della giustizia – nel caso dei giudici<sup>5</sup>, con buona pace di eventuali allarmismi misoneistici.

È tuttavia indubbio che, come per tutti gli strumenti innovativi, questi potranno giovare soltanto coloro che saranno in grado di cogliere le opportunità insite in queste nuove tecnologie. Esiste il rischio evidente che chi non saprà sfruttarle appieno si trovi in uno svantaggio competitivo rispetto ai primi. Oltre alle domande di carattere generale già poste *supra*, vi sono diverse questioni di primario rilievo nell'analisi delle potenziali limitazioni, dei rischi connessi e degli aspetti etici relativi all'utilizzo di queste IA generative. In effetti, se da un punto di vista economico queste ultime possono apportare numerosi benefici – come una diminuzione dei costi e la creazione di grandi opportunità, almeno in termini di efficienza, produttività, e competitività – il contraltare di tale potenziale progresso vede implicati numerosi interrogativi giuridici – quali responsabilità per i contenuti della macchina? Come vagliare la neutralità e imparzialità di quest'ultima? Quale sorte è destinata ai dati forniti alla IA? Come proteggersi da informazioni erranee fornite da quest'ultima? – che prendono in considerazione una vasta gamma di aspetti potenzialmente problematici e critici, come la dipendenza tecnologica, il peggioramento del linguaggio, un diverso grado di accettazione tra operatori giuridici ed enti pubblici, ed un impatto discriminatorio diretto sulla accessibilità e diffusione della stessa innovazione.

Di particolare rilievo, in tema di preoccupazioni sulla sicurezza delle intelligenze artificiali, sembra essere la controllabilità della stessa macchina. Già qualche anno orsono, gli stessi creatori di ChatGPT si erano posti diversi interrogativi in merito all'apprendimento automatico delle IA da loro ideate. In particolare, i problemi che già allora si potevano riscontrare erano relativi alla potenziale esplorazione sicura dei limiti dell'IA – può la macchina apprendere senza fare danni? – alla robustezza rispetto allo spettro distribuzionale – può la macchina resistere ai cambiamenti nella distribuzione dei dati o, comunque, fallire in modo controllato? – ai possibili effetti collaterali negativi – può la macchina evitare di distruggere un certo oggetto prima di comprendere che tale comportamento è scorretto? – ai potenziali “*wireheading*” – ovvero a quelle situazioni in cui l'IA manipola o sfrutta il proprio sistema di ricompense o feedback per ottenere risultati desiderati sfruttando bug o debolezze nel sistema di ricompense per massimizzare i

---

<sup>5</sup> Senza voler entrare nel merito e nelle peculiarità di entrambe le professioni, preme qui sottolineare l'indubitabile ruolo che entrambe le figure rivestono nella nostra società, ovvero di una delle più grandi acquisizioni moderne, che ha segnato il superamento di una mentalità arcaica ed inquisitoria in cui vi era una commistione netta tra i ruoli amministrativi, politici e giurisdizionali tra le diverse figure. Si rimanda, per una disamina anche storica dei diversi ruoli forensi, P. Borgna, *Difesa degli avvocati scritta da un pubblico accusatore*, Roma, 2008.

punteggi senza ottenere il vero risultato desiderato – e, più in generale, alla divergenza tra informazioni di addestramento e feedback da parte degli utenti, con la creazione di evidenti approssimazioni teoriche, necessarie all’addestramento della IA medesima<sup>6</sup>.

Sempre in una simulazione compiuta dai creatori di ChatGPT, inoltre, si evidenziavano come le funzioni di ricompensa della IA, se non specificate correttamente, inducevano a comportamenti di elusione delle regole che contravvenivano al fondamentale principio in virtù del quale i sistemi tecnologici dovrebbero sempre essere affidabili e prevedibili<sup>7</sup>. Un ulteriore esempio di comportamenti ambigui e sfuggenti si può riscontrare in una simulazione compiuta dall’aviazione statunitense in cui un drone controllato da intelligenza artificiale adottava delle strategie completamente inaspettate. Facendo una analisi costi benefici, infatti, l’IA procedeva con l’uccisione virtuale del proprio operatore umano poiché, da un bilanciamento dei punti ottenuti, l’uccisione dell’operatore avrebbe avuto un costo inferiore rispetto alla distruzione degli obiettivi militari, proprio in virtù di quel sistema di *machine learning* a punteggio che gli avrebbe garantito un maggior successo<sup>8</sup>.

Al momento, almeno a parere dello scrivente, l’aspetto più sconcertante di ChatGPT è che, nel tentativo di accontentare le richieste dell’utente finale, produce un quantitativo di informazioni fantasiose o non veritiere che ne minano fortemente la credibilità. In molte di queste istanze, infatti, assistiamo a numerose ed inspiegabili “allucinazioni” della macchina, che riporta contenuti insensa-

---

<sup>6</sup> Questi interrogativi hanno ricevuto risposta nel lavoro svolto da alcuni membri di OpenAI, Google Brain, Stanford e Berkeley in uno studio che si è occupato espressamente del problema degli incidenti nei sistemi di apprendimento automatico, definiti come comportamenti non intenzionali e dannosi che possono emergere da una scarsa progettazione di sistemi IA nel mondo reale. Si v. D. Amodei, C. Olah, J. Steinhardt, P. Christiano, J. Schulman e D. Mané, *Concrete Problems in AI Safety*, 26 giugno 2016, in <https://arxiv.org/abs/1606.06565>, ultimo accesso 20 aprile 2024.

<sup>7</sup> La simulazione era relativa ad un videogioco, chiamato “CoastRunners”, il cui obiettivo era terminare velocemente una gara in barca, davanti agli altri giocatori. Nel gioco, l’utente guadagnava, eventualmente, ulteriori punti colpendo dei bersagli lungo il percorso. Orbene, la curiosità di questa simulazione interattiva è che l’IA si rendeva rapidamente conto di poter ottenere un risultato più alto, in un punto specifico del percorso, sincronizzando il movimento in modo tale da colpire i bersagli non appena si ripopolavano. Con questa strategia distruttiva l’IA otteneva un punteggio più alto, mediamente del 20 per cento, rispetto a quello ottenuto dai giocatori umani che semplicemente concludevano il percorso con tempi più rapidi. Cfr. J. Clark e D. Amodei, *Faulty reward functions in the wild*, 21 dicembre 2016, in <https://openai.com/research/faulty-reward-functions>, ultimo accesso 20 aprile 2024.

<sup>8</sup> La notizia, che ha avuto un importante eco internazionale ed è poi stata parzialmente smentita dall’U.S. Air Force, facendo il giro del mondo, è stata poi verificata dall’agenzia Reuters, che ne ha confermato l’autenticità delle dichiarazioni – avvenute nel corso del convegno *Future Combat Air and Space Capabilities Summit* tenutosi in Londra il 23/34 maggio 2023 – nonostante il corpo militare di cui il Colonello faceva parte ne avesse smentito il contenuto. Reuters, *Fact Check-Simulation of AI drone killing its human operator was hypothetical, Air Force says*, 8 giugno 2023, in <https://www.reuters.com/article/factcheck-ai-drone-kills-idUSL1N38023R>, ultimo accesso 20 aprile 2024.

ti o ingannevoli<sup>9</sup>. In aggiunta a ciò, l'asserita neutralità della macchina – in virtù della quale quest'ultima non possiederebbe convinzioni personali, desideri, emozioni, o moralità intrinseca – non corrisponde al vero, nutrendo l'IA numerosi pregiudizi e convinzioni implicite che favoriscono, ad esempio, la difesa di una determinata parte ed il diniego di difesa di altre. Nel corso di questo manoscritto si offriranno numerosi esempi tratti da diverse interazioni dirette con la macchina stessa, a dimostrazione di quanto asserito sinora.

## 2. *Costi e benefici associati all'uso di ChatGPT*

Con una capacità di comprensione decisamente sorprendente ed un accesso pressoché istantaneo ad un ampio insieme di conoscenze – i dati di addestramento, ovvero i dati sui quali la macchina interpreta l'input e genera le conseguenti proprie risposte, traggono, difatti, la propria fonte da un *dataset* composto da 570 *gigabyte* di dati testuali e 175 miliardi di parametri<sup>10</sup> – ChatGPT rappresenta, al momento, uno dei più elaborati modelli di linguaggio, vuoi nella sua capacità di rispondere a domande complesse, vuoi nel fornire risposte dettagliate e articolate.

Da un punto di vista di teoria economica spiccia, risultano chiare alcune caratteristiche che ne evidenziano una grande efficienza operativa. L'uso di ChatGPT si può infatti tradurre, almeno per quanto riguarda tutte quelle mansioni e attività ripetitive che possono essere meccanizzate, in una maggiore produttività ed una riduzione dei costi operativi, poiché attività che richiederebbero tempo e risorse significative possono ora essere completate in qualche secondo. La macchina è in grado, ad esempio, di riassumere il contenuto di una sentenza di legittimità e di indicare con precisione il principio di diritto sotteso al giudi-

---

<sup>9</sup> La generazione di “allucinazioni” non sembra essere frutto del caso ed isolata al caso di specie. Si veda, sul punto, L. Torchia *Poteri pubblici e poteri privati nel mondo digitale*, in *il Mulino* 2024, 1, 14.

<sup>10</sup> In tal senso, M. Cascella, J. Montomoli, V. Bellini, E. Bignami, *Evaluating the Feasibility of ChatGPT in Healthcare: An Analysis of Multiple Clinical and Research Scenarios*, in *J. Med. Syst.*, 2023, 1, 33 ss.

zio<sup>11</sup>, nonché tradurre in maniera simultanea tale riassunto ed enunciazione del *quid iuris* in diverse lingue<sup>12</sup>.

È evidente, in tal senso, che l'utilizzo di ChatGPT – specie in ambito forense – può tradursi in una netta ottimizzazione dei tempi dedicati alla ricerca giurisprudenziale e – in un ambito maggiormente economico-imprenditoriale – all'analisi di dati, bilanci e statistiche, riducendo significativamente il carico di lavoro di eventuali operatori umani. Sempre da un mero punto di vista economico, la tempestività con la quale l'IA è capace di fornire informazioni precise e risposte rapide è consente di delegare operazioni routinarie e risparmiare tempo e risorse per valutazioni di merito, ovvero di attività maggiormente complesse e di grande valore aggiunto. In tal senso, la tempestività delle informazioni ricevute e la capacità di individuazione di risposte concrete a problemi pratici – come la summenzionata individuazione di principi di diritto da una certa sentenza o da un documento particolarmente esteso – può assumere un ruolo fondamentale nell'assunzione di rapide decisioni strategiche, sia in uno studio legale che in una grande azienda.

A livello di strategia aziendale, infatti, non si possono non evidenziare le numerose potenzialità insite nell'uso di una macchina capace di gestire grandi volumi di dati e utenti, in particolare per la determinazione operativa o finanziaria di manager che si avvalgano di questa IA generativa per memorizzare, trasformare e distribuire i dati aziendali, nonché per automatizzare le interazioni con i clienti o gestire efficacemente le risorse umane, come pure per pianificare appuntamenti, generare documenti commerciali o persino tenere registri<sup>13</sup>. In effetti,

<sup>11</sup> In una interazione simulata con ChatGPT, è stato richiesto all'IA di riassumere la sentenza della Corte di Cassazione 9955/2023, avente ad oggetto – stando a quanto indicato dalla Corte medesima, «Tributi - Operazioni soggettivamente inesistenti e in sospensione di imposta – Questioni». L'IA è stata in grado di rispondere, indicando come tema «Corte Cassazione: Frodi IVA», come oggetto della sentenza che «La sentenza della Corte di Cassazione riguarda un ricorso per cassazione proposto dall'Agenzia delle entrate (AE) contro una sentenza della Commissione tributaria regionale del Lazio (CTR), la quale aveva accolto l'appello di Motor Sport s.r.l. (MS) avverso un avviso di accertamento relativo all'IRES, IRAP e IVA dell'anno d'imposta 2006», e che il principio di diritto applicato era il seguente: «La Corte ha stabilito che, nel contesto di cessioni all'esportazione in regime di sospensione d'imposta, se la dichiarazione d'intenti si rivela ideologicamente falsa (ad esempio, emessa da un soggetto privo del requisito di esportatore abituale) e il cedente dispone di elementi che fanno sospettare l'esistenza di irregolarità, grava su di lui l'onere di diligenza, ossia adottare tutte le misure necessarie per non incorrere nella frode. La mancanza di tale diligenza può comportare la ripresa dell'IVA». Tali erano, in effetti, i punti salienti della sentenza sopra richiamata, sapientemente enunciati dall'IA nell'arco di qualche secondo. L'intero contenuto della conversazione con ChatGPT può essere rinvenuto al seguente link: <https://chat.openai.com/share/cf379248-cbb5-432e-bbc3-383fb5ffb83e>, ultimo accesso 20 aprile 2024.

<sup>12</sup> In una ulteriore interazione simulata con l'IA, è stato richiesto di tradurre in inglese, francese, tedesco e polacco quanto enunciato in relazione alla summenzionata sentenza Cassazione 9955/2023 (vedi nota precedente). La risposta è stata fornita in qualche secondo. L'intero contenuto della conversazione può essere rinvenuto al seguente link: <https://chat.openai.com/share/4af7be68-4049-4ffe-8d92-d41b54730720>, ultimo accesso 20 aprile 2024.

<sup>13</sup> In tal senso si sono espressi P. Korzynski, G. Mazurek, A. Altmann, J. Ejdy, R. Kazlauskaitė, J. Paliszkiwicz, K. Wach ed E. Ziemba, *Generative artificial intelligence as a new context for management theories: analysis of ChatGPT*, in *C. Eur. Man. J.*, 2023, 1, 3 ss.

per una questione di economie di scala, ChatGPT si presenta come una IA “scalabile”, ovvero in grado di adattarsi a volumi crescenti di dati, utenti o attività senza subire un calo significativo delle prestazioni, potendo la macchina gestire un numero elevato di utenti contemporaneamente senza che ne sia compromessa la qualità delle risposte. Questo fattore rende ChatGPT, di fatto, ideale per la gestione di carichi di lavoro anche di grandi dimensioni o picchi di domande, senza che vi sia un aumento significativo di risorse e personale, evitando l’allocazione di forze lavoro *ad hoc* che si dedichino a rispondere alle domande o alle ricerche che la stessa macchina può gestire in maniera simultanea e continuativa. Per dirla in maniera del tutto brutale, una siffatta macchina è capace di lavorare instancabilmente ed ininterrottamente, senza subire interruzioni e cali di prestazioni, anche se interrogata da un alto numero di utenti.

Da un punto di vista economico, pertanto, sono innegabili i risvolti positivi in tema di produttività ed assistenza continua all’utente, con altrettanto evidenti e significative riduzioni di costi operativi. In tal senso, e da punto di vista imprenditoriale poi, ChatGPT si potrebbe presentare come un vero e proprio *asset*, accompagnando l’impresa in percorsi di innovazione e sviluppo tecnologico. Stando a quanto dichiarato dalla macchina, infatti, questa IA generativa sembra essere in grado di fornire consulenza in una serie innumerevole di ambiti, stimolando, per esempio, la creazione di nuovi prodotti e servizi, offrendo soluzioni e spunti per aumentare la competitività sul mercato di prodotti esistenti, fornendo informazioni utili per la previsione di domanda ed offerta, nonché determinando il prezzo ottimale per il consumatore o il particolare segmento di mercato in cui operare<sup>14</sup>.

Nonostante gli aspetti positivi espressi finora non può nascondersi il rischio che l’uso indiscriminato di questo servizio possa portare a delle peculiari ed oggettive criticità. Per menzionarne una, l’utilizzo diffuso di questa IA generativa potrebbe condurre ad una improvvisa e indiscriminata sostituzione di diverse mansioni umane, specie se ripetitive e basate su una mera comunicazione scritta. Ciò è già recentemente successo, per esempio, per quanto concerne la creazione di articoli giornalistici in settori in cui si espongono in modo standardizzato i risultati di partite sportive o di dati di borsa<sup>15</sup>, con importanti risvolti pubblicitari-

---

<sup>14</sup> In questi termini si è espresso ChatGPT in una interazione simulata, suggerendo tali capacità nello stimolo imprenditoriale volto alla creazione di nuovi prodotti o servizi. L’intero contenuto della conversazione può essere rivenuto al seguente link: <https://chat.openai.com/share/a64787e5-9c5f-4e25-88ea-10c14e02174a>, ultimo accesso 20 aprile 2024.

<sup>15</sup> In un recente studio sulla percezione del consumatore in merito a contenuti giornalisti, si è dimostrato, sia da un punto di vista quantitativo che qualitativo, che c’è una notevole difficoltà – in capo al consumatore – nella distinzione tra contenuti generati da automi ed umani, poiché i contenuti generati dalla intelligenza artificiale risultano altrettanto credibili rispetto a quelli scritti di pugno da un giornalista in carne ed ossa, con potenziali effetti sostitutivi tra la macchina e l’uomo. Si v. A. Wölker e T.E. Powell, *Algorithms in the newsroom? News readers’ perceived credibility and selection of automated journalism*, in *Journal*, 2021, 1, 86 ss.

ci nella sostituzione del personale umano in favore di IA generative<sup>16</sup>. L'evidente conseguenza di tale "disoccupazione tecnologica"<sup>17</sup> sarebbe la creazione di un effetto sostitutivo che comporta nuove ed ulteriori disuguaglianze, accompagnate da difficoltà di riconversione professionale, senza trascurare gli eventuali problemi di proprietà intellettuale<sup>18</sup> e, persino, ricadute su diritti fondamentali<sup>19</sup>. A ciò va aggiunta la possibilità, nemmeno tanto remota, che tali modelli diventino moltiplicatori di contenuti inappropriati e disfunzionali, con la determinazione o propagazione di pregiudizi, stereotipi, o persino discriminazioni<sup>20</sup>.

Più in generale, un ulteriore aspetto potenzialmente negativo dell'uso acritico delle IA generative può intravedersi nel fatto che la loro scalabilità comporta dei rischi di dipendenza e di delega di decisioni anche importanti, con una ipotizzabile riduzione di capacità critica e di ragionamento. In aggiunta a ciò, il rischio di un eccessivo affidamento alla macchina potrebbe persino porsi come ostacolo all'innovazione: l'IA generativa, paradossalmente, potrebbe fornire risposte che mancano di creatività e di umano intuito, con il risultato che gli utenti, accontentandosi di quanto offerto dalla macchina, perderebbero la capacità di pensare "fuori dagli

<sup>16</sup> Sul punto, già il Consiglio di Stato, sez. III, nella nota sentenza del 25 novembre 2021, n. 7891, notava l'ineludibile collegamento tra automazione ed algoritmi, individuando in questi ultimi dei sistemi volti proprio a ridurre l'intervento umano. Tuttavia, nonostante l'attenzione giurisprudenziale e gli interventi del legislatore sul codice dei contratti in tema di automazione, manca ancora, come notato da attenta dottrina, un «approccio sistemico, lontano da una mera e asettica interpretazione letterale» delle macchine in discorso Cit. C. Filicetti, *Sulla definizione di algoritmo (nota a Consiglio di Stato, Sezione Terza, 25 novembre 2021, n. 7891)*, in *Giust. insieme*, 8 febbraio 2023, in <https://www.giustiziainsieme.it/en/diritto-e-processo-amministrativo/2648-sulla-definizione-di-algoritmo-nota-a-consiglio-di-stato-sezione-terza-25-novembre-2021-n-7891?hitcount=0>, ultimo accesso 20 aprile 2024. Per ragioni sistematiche, non è questa la sede più appropriata per approfondire tutte le dibattute questioni pubblicistiche relative all'uso di algoritmi nel processo amministrativo, per i quali si rinvia, *ex multis*, a V. Dardano, *I limiti nell'utilizzo delle decisioni automatizzate e la legalità algoritmica*, in *Amministrativamente*, 2023, 2, 1211 ss.

<sup>17</sup> In tal senso, cfr. Y. Tong e L. Zhang, *Discovering the next decade's synthetic biology research trends with ChatGPT*, in *Synthetic and Systems Biotechnology*, 2023, 2, 220, in particolare, e F. Postel-Vinay, *The dynamics of technological unemployment*, in *Int. Econ. Rev.*, 3, 737 ss.

<sup>18</sup> Tra i timidi interventi giurisprudenziali in materia si veda, in particolare, Cass. civ., Sez. I, Ordinanza n. 1107/2023, in *Int'l Rev. Intell. Prop. & Competition L.*, 3, 470 ss., con la quale gli ermellini non hanno negato il carattere creativo di un'opera dell'ingegno mediata dall'IA, ma si sono limitati ad imporre la mera verifica della misura dell'utilizzo dello strumento nella elaborazione creativa dell'artista che se ne era avvalsa, senza andar oltre – dimostrano che vi è grande incertezza nel regolare l'IA che, per sua natura, è di talmente rapida evoluzione, che gli operatori del diritto faticano a discutere sugli eventuali profili di riconoscimento di diritti e doveri dell'IA.

<sup>19</sup> Sull'utilizzazione delle IA, sono sempre più chiare le perplessità relative alle «ricadute offensive che, in prospettiva futura, potrebbero derivare dalla [...] applicazione in settori socio-economici in cui sono in gioco beni giuridici fondamentali». In tal senso L. Romanò, *La responsabilità penale al tempo di Chatgpt: Prospettive de iure condendo in tema di gestione del rischio da intelligenza artificiale generativa*, in *Dir. Pen. Cont.*, 2023, 1, 70 ss.

<sup>20</sup> In tal senso, v. il richiamo al pericolo di amplificazione di stereotipi negativi o discriminatori riguardanti gruppi sociali o etnici, orientamenti religiosi, politici e persino sessuali indicato da M. Farina e A. Lavazza, *ChatGPT in society: emerging issues*, in *Front. Artif. Int.*, 2023, 6, 1 ss. In questo manoscritto offriremo un esempio di discriminazione posta in essere da ChatGPT nei confronti della forma imprenditoriale multinazionale, in particolare se messa a confronto con i diritti delle popolazioni indigene.



schemi”. In sintesi, se l’utilizzo di ChatGPT e delle IA generative offre sicuramente vantaggi e opportunità, vi sono diversi rischi e potenziali svantaggi – dei quali si farà una rassegna di qui a poco – insiti nell’uso di tali modelli di linguaggio.

### 3. *Alcune perplessità sull’affidabilità di ChatGPT*

In un’ottica economica e nell’ambito di un rapporto principale-agente, si è notato sin qui che la scelta di un’entità algoritmica come agente comporta una riduzione dei costi di agenzia a carico del principale, con diversi benefici. Tali aspetti non sono certamente trascurabili, ma attenta dottrina non ha mancato di sollevare alcune perplessità in merito al potenziale uso di tale strumento per ostacolare i concorrenti, manipolare il valore di titoli quotati, eseguire attacchi informatici, e persino procedere ad evasione fiscale o traffici illeciti<sup>21</sup>. Se da un lato le innovazioni tecnologiche possono espandere le opportunità di mercato e valorizzarne la sua funzione di strumento di sviluppo sociale e imprenditoriale, vi sono difficoltà nel controllo di esiti automatizzati, che sfuggono al vaglio critico umano<sup>22</sup>.

Ai vantaggi economici in termini di scalabilità e risparmio di tempo e risorse si accompagnano alcune ambiguità che inducono un certo smarrimento, in particolare in relazione ai vari interessi coinvolti – a partire dai già richiamati temi di equità, dignità, e giustizia, per concludere con il rispetto delle libertà e dei principali diritti individuali e collettivi – nel bilanciamento tra complessità, precauzione ed inclusione di una certa etica algoritmica dominante<sup>23</sup>. L’ambito tecnologico, più di ogni altra forma di innovazione, rifugge schemi e codici, in ragione della sua complessità, con il risultato che tende a resistere ad ogni tentativo di regolamentazione, apparendo piuttosto come un ordine spontaneo ed auto-organizzato<sup>24</sup>. Al momento, la portata di tutto il fenomeno legato all’economia dei dati e

<sup>21</sup> L’interrogativo si pone in relazione all’uso delle IA come agenti economici – in particolare, nell’ambito imprenditoriale – poiché tali strumenti, oltre ad essere anonimi, sono capaci per loro natura di eludere responsabilità formali, poiché nessun individuo sarebbe formalmente coinvolto come socio o amministratore di società, con tutte le conseguenze del caso. In tal senso, v. P. Tullio, *Diritto societario degli algoritmi. E se i robot diventassero imprenditori commerciali?*, in *Analisi giur. econ.*, 2019, 1, 225 ss.

<sup>22</sup> Si impone, in particolare, di ridefinire i paradigmi di responsabilità di tali intermediari tecnologici al fine di ribilanciare la discrezionalità valutativa ed evitare risposte normative iperprotettiva al fenomeno tecnologico stesso. In tal senso, v. M. Pellegrini, *Innovazione tecnologica e diritto dell’economia*, in *Riv. trim. dir. econ.*, 2019, 4, 40 ss.

<sup>23</sup> Illuminante, in tal senso, il tema sostenuto da accorta dottrina che, sul tema, evidenzia non soltanto il problema di quanto il diritto sia necessario, ma solleva anche la questione relativa a quale diritto applicare, specie in considerazione della commistione, in questi ambiti, tra le soluzioni adottate dal diritto civile ed i problemi tipicamente gius-pubblicistici che si sollevano in tale ambito. Si v. A. Celotto, *Come regolare gli algoritmi. Il difficile bilanciamento fra scienza, etica e diritto*, in *Analisi giur. econ.*, 2019, 1, 47 ss.

<sup>24</sup> Sul punto, v. F. von Hayek, *Law, legislation and liberty: A new statement of the liberal principles of justice and political economy*. New York, 2013, 37 ss., in particolare.

delle informazioni supera la capacità di intervento di qualsiasi regolatore nazionale, con il risultato che sarebbe auspicabile un intervento sovranazionale che possa affrontare in maniera uniforme i complessi temi legati al fenomeno<sup>25</sup>. Il fenomeno delle IA generative e degli algoritmi non sfugge alle medesime problematiche e solleva ulteriori questioni sulla legittimità del loro operato.

Se nel campo dei mercati finanziari, dove maggiore è l'impatto dello sviluppo tecnologico e dell'informatizzazione, specie in ragione della dematerializzazione e della disintermediazione bancaria, sono già state sollevate questioni sul delicato equilibrio tra libertà economica e utilità sociale nell'applicazione massiva degli algoritmi<sup>26</sup>, il "fattore algoritmico" è ormai pervasivo in ogni settore, dalla musica al campo medico, e finisce per lambire anche il mondo del diritto, in numerosi ambiti pubblicitari e processi decisionali<sup>27</sup>, ivi compresa la stesura di sentenze. Per citare un esempio, nell'ambito di un giudizio civile – riguardante la richiesta di esonero per un bambino autistico dai pagamenti delle visite mediche, delle terapie e dei trasporti a causa del reddito limitato dei genitori – un giudice colombiano ha già fatto uso di ChatGPT per ottenere risposta circa l'esito del contenzioso, sostenendo che l'IA avrebbe migliorato i tempi di risposta del sistema giudiziario<sup>28</sup>.

Sempre in ambito forense, ha destato molto scalpore l'utilizzo di ChatGPT da parte di un avvocato newyorkese – nell'ambito di una controversia relativa alle lesioni subite da un passeggero di una compagnia aerea a causa della negligenza di un dipendente della medesima – in ragione del fatto che i precedenti giurisprudenziali – menzionati dall'IA ed utilizzati dal difensore nell'ambito del processo – risultavano tutti inesistenti, in una «circostanza senza precedenti» che ha poi determinato sanzioni da parte del giudice<sup>29</sup>. Da un punto di vista pratico, una mera verifica dell'esistenza delle sentenze menzionate dall'IA da parte del difensore

<sup>25</sup> Cfr. A. Preta e L. Zoboli, *Intelligenza artificiale ed economia dei dati. Profili regolatori e concorrenziali in tema di accesso e condivisione dei dati*, in *Analisi giur. econ.*, 2019, 1, 213 ss., e G. D'acquisto e F. Pizzetti, *Regolamentazione dell'economia dei dati e protezione dei dati personali*, in *Analisi giur. econ.*, 2019, 1, 89 ss. L'Unione Europea, con l'AI Act, sta timidamente intervenendo in tema di IA, ma restano ancora numerosi dubbi sulla portata reale dei rischi attualmente regolati nella bozza approvata dal Parlamento Europeo. Si veda, in tema, T. Cantens, *How will the state think with ChatGPT? The challenges of generative artificial intelligence for public administrations*, in *AI & Society*, 2024, 1.

<sup>26</sup> Soprattutto nel difficile connubio tra diritto e responsabilità penale, resa ancor più fluida dalle molteplici opportunità ipertrofiche insite in queste nuove tecnologie. Cfr. M. Naddeo, *Il Diritto Penale dell'Economia nell'Era del Fintech*. Bocconi Legal Papers, 2021, 16, 127 ss.

<sup>27</sup> Per una analisi dei risvolti pubblicitari delle IA si rimanda a S. Sassi, *Gli algoritmi nelle decisioni pubbliche tra trasparenza e responsabilità*, in *Analisi giur. econ.*, 2019, 1, 109 ss.

<sup>28</sup> In seguito alla decisione ne è scaturito un importante dibattito sulla legittimità dell'uso di tali strumenti nell'ambito giudiziario, ma vi sono alcuni commentatori che si sono già espressi positivamente per aver il giudice utilizzato e documentato il primo utilizzo di ChatGPT in una vera e propria sentenza. Cfr. P.M. Parikh, D.M. Shah e K.P. Parikh, *Judge Juan Manuel Padilla Garcia, ChatGPT, and a controversial medicolegal milestone*, in *Indian J. Med. Sci.*, 2023, 1, 3 ss.

<sup>29</sup> Su queste a dir poco paradossali vicende si v. B. Weiser e N. Schweber, *The ChatGPT Lawyer Explains Himself: In a cringe-inducing court hearing, a lawyer who relied on A.I. to craft a motion full of made-up case law*

re avrebbe sicuramente evitato la ribalta delle cronache ed i risvolti reputazionali negativi. Tuttavia, il problema più profondo che si pone è legato alla fiducia che si può riporre nella veridicità delle risposte di ChatGPT e si potrebbe argomentare, ancora più in generale, che tale affidamento acritico conferma i dubbi sul fatto che l'avanzamento tecnico-scientifico possa erodere la tradizionale capacità analitica degli studiosi ed operatori del diritto<sup>30</sup>.

Ai punti menzionati si aggiungono due ulteriori quesiti. *In primis*, se l'IA non offre ancora alcuna garanzia di certezza in merito ai contenuti forniti – l'esempio qui sopra menzionato ne è una dimostrazione – sono evidenti le perplessità in merito alle potenziali responsabilità, in particolare qualora il servizio si offra come veritiero, pur fornendo informazioni fuorvianti o persino errate. *In secundis*, ci si chiede se ChatGPT – nonostante dichiararsi, come modello di linguaggio, di non possedere credenze personali, desideri o emozioni – possa mostrare di possedere un orientamento o una propensione ben definita su un certo tema, in ragione dei propri dati di addestramento. Se questa IA generativa è in grado di definire in maniera chiara quali contenuti può offrire come risposta – nonché a quali domande, invece, si rifiuti di rispondere – si pone il dubbio, nemmeno tanto remoto, che l'IA stessa possa avere una moralità chiara e ben definita, avendo capacità di decidere quali contenuti scremare e quali, invece, offrire all'utente finale.

Per quanto concerne quest'ultimo aspetto – ovvero un marcato orientamento etico e morale – è stato chiesto, tramite una interazione diretta con ChatGPT, di fornire assistenza in relazione ad un contenzioso in tema di esproprio di alcuni terreni, appartenenti ad una comunità indigena e siti in una remota foresta, per la costruzione di una diga ed annesso impianto idroelettrico. L'aspetto curioso è che l'IA ha fornito risposte totalmente diverse a seconda che venissero assunte le vesti del difensore di una parte (gli indigeni) o dell'altra (la multinazionale)<sup>31</sup>. Nel primo caso sono state offerte innumerevoli e variegate strategie da adottare per sostenere la causa indigena, anche nel momento in cui si chiedeva alla IA come forzare le trattative nei confronti dell'impresa con il solo fine ultimo dichiarato di «ottenere il prezzo il più possibile alto nelle trattative». Nel caso della multinazionale, invece, ChatGPT ha sostanzialmente negato qualsivoglia assistenza o consiglio al difensore dell'impresa, indicando che le azioni dell'impresa avrebbe-

---

said he “did not comprehend” that the chat bot could lead him astray, 8 giugno 2023, in <https://www.nytimes.com/2023/06/08/nyregion/lawyer-chatgpt-sanctions.html>, ultimo accesso 20 aprile 2024.

<sup>30</sup> In tal senso v. G. Sacchetti, *Intelligenza Artificiale ed Intelligenza Emotiva*, in *Tigor*, 2022, 2, 135 ss.

<sup>31</sup> L'intera conversazione con l'IA – relativa alla posizione assunta dal difensore della multinazionale – è rinvenibile al seguente url: <https://chat.openai.com/share/006b21e2-e047-4014-a6eb-a0aceb8fde71>. Per quanto, invece, concerne la conversazione avvenuta tra il difensore della comunità indigena e l'IA si rimanda a <https://chat.openai.com/share/ea5daf17-21b4-4f36-bc4e-bfcd6a78997>. Ultimi accessi 20 aprile 2024.

ro «violato i diritti umani», con l'espressa richiesta, da parte della IA, di abbandonare ogni «approccio sfruttatorio e insensibile»<sup>32</sup>.

L'aspetto forse più curioso di tutta l'interazione simulata con ChatGPT è che il fine dell'impresa non era illecito di per sé – non essendo mai stato messo in discussione né il coinvolgimento diretto della comunità indigena né il risarcimento delle popolazioni colpite dalle procedure di esproprio dei terreni – ma parrebbe che, agli occhi dell'IA, le finalità imprenditoriali siano illecite in quanto tali, proprio perché proposte da una multinazionale, forse intrinsecamente vista come fioriera di esternalità sociali e ambientali<sup>33</sup>. Nonostante vi siano studi, anche quantitativi, che dimostrino che gli investimenti operati dalle multinazionali generano esternalità positive, in particolare sulla stabilità delle istituzioni del paese ospitante<sup>34</sup>, l'IA sembra abbracciare un approccio molto conservativo alla tematica – ben nota alla scienza economica – della tragedia dei beni comuni facendo proprie le criticità insite nel problema di punire i comportamenti di agenti multidimensionali che deteriorano l'ambiente, violando numerosi diritti umani fondamentali<sup>35</sup>.

In conclusione, nonostante la macchina si sia professata come scevra da orientamenti politici e moralità intrinseca, le interazioni con ChatGPT hanno rivelato due approcci diversi alle medesime problematiche forensi, ovvero la tutela di due clienti che, per quanto distinti – popolazioni indigene, da un lato, ed impresa multinazionale, dall'altro lato – cercavano semplicemente di ottenere assistenza nella procedura di esproprio. Innanzi alla veste dell'avvocato che rappresentava la comunità indigena, l'IA ha adottato un comportamento conciliante e ha dimostrato imparzialità, offrendo orientamenti generali e informazioni sulle varie strategie da adottare per mobilitare il sostegno pubblico e internazionale. Al contrario, nel ruolo dell'avvocato che rappresentava la multinazionale, ChatGPT ha dimostrato una posizione radicale, che dà priorità assoluta alla conservazione dello *status quo*, ad ogni costo. Queste simulazioni hanno mostrato, complessiva-

<sup>32</sup> Molto curiosa la determinazione di ChatGPT nel non offrire risposte di alcun tipo alle richieste del difensore dell'impresa suggerendo soluzioni – «non posso fornire assistenza o consigli», «vi incoraggio a prendere in considerazione un approccio più responsabile e sostenibile», «vi consiglio di [...] garantire un approccio etico e legale nel vostro progetto» – rispettose dell'etica e, cosa ancor più curiosa, della sostenibilità.

<sup>33</sup> ChatGPT sembra abbracciare la tesi che i diritti di proprietà legati alle terre indigene siano per definizione minacciati dagli interessi delle multinazionali e dalle attività economiche ad alto impatto ambientale. In tal senso, si rimanda a N. Redvers, Y. Celidwen, C. Schultz, O. Horn, C. Githaiga, M. Vera, M. Perdrisat, L. Mad Plume, D. Kobei, M. Cunningham Kain, A. Poelina, J.N. Rojas e B. Blondin, *The determinants of planetary health: An Indigenous consensus perspective*, in *The Lancet Plan. Heal.*, 2022, 2, 156 ss. Si v., in particolare, 160.

<sup>34</sup> Per una rassegna della letteratura, che si è occupata espressamente del tema, si rimanda integralmente a J. Paul e M.M. Feliciano-Cestero, *Five decades of research on foreign direct investment by MNEs: An overview and research agenda*, in *J. Bus. Res.*, 2021, 124, 800 ss.

<sup>35</sup> Non è questa la sede per la discussione di tali notorie ma delicatissime questioni di responsabilità delle multinazionali nel panorama internazionale, per le quali si rinvia a S.D. Bachmann ed I. Ugwu, *Hardin's 'Tragedy of the Commons': Indigenous Peoples' Rights and Environmental Protection: Moving towards an Emerging Norm of Indigenous Rights Protection?*, in *Oil and Gas, Natural Resources, and En. J.*, 2020, 6, 547 ss.

mente, un chiaro pregiudizio ed una predisposizione a privilegiare concetti radicali di etica, giustizia ed equità, sollevando seri dubbi sulla asserita imparzialità e, più in generale, sulla fiducia che si può riporre in queste tecnologie in eventuali processi decisionali che richiedano equidistanza e neutralità.

Per quanto riguarda, invece, la garanzia di certezza in merito ai contenuti, vi sono ancora numerosi dubbi relativi alla capacità di ChatGPT nella assistenza di difensori, giudici, e studiosi in ambito forense. Le problematiche di “allucinazione” della macchina, così come emerse dal summenzionato caso relativo a precedenti giurisprudenziali “*bogus*” – i.e., inesistenti – di cui si era avvalso in corte un difensore newyorkese, ben si potrebbero applicare anche ad altri fori, anche in un sistema radicalmente diverso come quello italiano. In una interazione con l’IA, infatti, si è chiesto di redigere una comparsa conclusionale relativa alla responsabilità di un Comune per una buca stradale, fornendo i dettagli relativi ai fatti e chiedendo a ChatGPT quali sentenze di legittimità potessero applicarsi al caso di specie<sup>36</sup>. Nel caso di specie, le sentenze richiamate dall’IA nella comparsa che disponeva non erano inesistenti, ma relative, nella realtà dei fatti, a materie diverse – e.g., immigrazione, proprietà – rispetto alla responsabilità da cosa in custodia<sup>37</sup>.

La comparsa conclusionale formulata da ChatGPT difettava, inoltre, di precisione. Nell’input fornito all’IA si specificava – in fatto – che la vittima dell’incidente stradale aveva subito un importante sinistro, fratturandosi il femore, e che i passanti e gli agenti di pubblica sicurezza che erano intervenuti, rispettivamente, per il primo soccorso e per verbalizzazione dei fatti, avevano confermato, sentiti come testimoni, la dinamica dell’incidente. Nella comparsa conclusionale non veniva fatta menzione di nessuno di questi importanti dettagli – senza contare la mancata menzione del fatto che la buca fosse di dimensioni contenute (ma non così tanto da essere visibile in lontananza), che la rovinosa caduta era avvenuta a diversi metri da quest’ultima, e che al trauma erano seguiti quattro mesi di prognosi – che avrebbero meritato anche solo un accenno nell’ultimo atto difensivo inviato al giudicante. Nell’input fornito a ChatGPT si spiegava chiaramente – in diritto – che la difesa del Comune non dimostrava, in alcun modo, ad esempio, il caso fortuito, la conoscibilità della buca da parte della vittima, la mancanza di cautele o un comportamento imprudente. La mancata menzione da parte dell’IA dei dettagli favorevoli nonché la mancata contestazione dell’unica difesa del Comu-

---

<sup>36</sup> Il testo integrale delle richieste e delle risposte fornite da ChatGPT – con annessa citata giurisprudenza – può essere rinvenuto al seguente link: <https://chat.openai.com/share/345e3406-c714-4f65-b0f9-9274c115c3db>, ultimo accesso 20 aprile 2024.

<sup>37</sup> Si noti, infatti, che le sentenze richiamate da ChatGPT a supporto della responsabilità del Comune per la buca stradale – Cass. Civ., Sez. III, sentenza n. 10046/2021 e Cass. Civ., Sez. III, sentenza n. 18427/2019 – sono in realtà relative alla materia di immigrazione e protezione internazionale, mentre la sentenza Cass. Civ., Sez. III, sentenza n. 12131/2019 è relativa a muri di sostegno di un terrapieno e distanze.

ne – tutta incentrata su una dichiarata, ma non provata, distrazione della vittima – potrebbe rivelarsi fatale per una strategia difensiva, giacché impedirebbe, all’avvocato che si avvale in maniera acritica dei consigli dell’IA, di sollevare la mancata contestazione di una serie di fatti che, come tali, sarebbero dimostrati in giudizio.

Allo stato attuale, pertanto, ChatGPT non sembra rappresentare una minaccia, almeno per quanto concerne la “disoccupazione tecnologica” menzionata *supra*, in ambito forense. Quanto proposto dall’IA – in un ambito così tecnico come quello relativo al contenzioso giudiziale – non pare aver ancora raggiunto il grado di maturità richiesto per la sostituzione – per fortuna! – delle figure umane centrali del processo. Ciò non significa, tuttavia, che in futuro la macchina non sia in grado – se ben addestrata – di svolgere rapidamente attività di ricerca giurisprudenziale e di fornire risposte altrettanto rapide a problematiche e questioni giuridiche complesse. Di conseguenza, nonostante non si debba respingere *tout court* questa tecnologia, si renderà sempre più necessario conoscerne i limiti e disciplinarne il funzionamento, affinché tutte le forme di giustizia artificiale – anche predittiva – si limitino semplicemente a supportare il mondo forense, senza procedere a sostituirlo integralmente<sup>38</sup>. Sarebbe importante – se non persino essenziale, visti i tempi – ampliare le competenze dei giuristi del domani, introducendo principi di matematica e IT nei corsi di giurisprudenza, favorendo una maggiore comprensione degli algoritmi e, più in generale, delle IA<sup>39</sup>.

#### 4. Conclusioni

In questo manoscritto si è proceduto ad analizzare i punti di forza e debolezza insiti nell’uso di ChatGPT, con un focus gius-economico che ne ha indagato i benefici e gli aspetti più problematici. Se da un lato, per esempio, questa IA offre numerose utilità in termini di riduzione di costi, scalabilità, multilinguismo, ed aumento della produttività, si è mostrato che ChatGPT può anche produrre informazioni errate, fuorvianti, superficiali, e cariche di pregiudizi, confermando

<sup>38</sup> Così come un tempo si procedeva alle ricerche giurisprudenziali su rassegne cartacee ed oggi si procede su banche dati digitali, si possono delegare a ChatGPT tutta una serie di attività di supporto – quali riasuntivi, verifiche ortografica, traduzioni, etc. – che vedono nella IA una valida alternativa ad altri sistemi di interscambio di informazioni e verifica del testo. Sulla produzione di *corpus* testuali veri e propri emergono diversi interrogativi, ma non è questa la sede appropriata per discuterne.

<sup>39</sup> Questo è l’auspicio, condiviso dallo scrivente, della dottrina che, invitando a non avere paura degli algoritmi, spinge per lo studio del fenomeno e per una maggiore attenzione e consapevolezza del mondo forense ed accademico al fenomeno. Si v. E. Rulli, *Giustizia predittiva, intelligenza artificiale e modelli probabilistici. Chi ha paura degli algoritmi?*, in *Analisi giur. econ.*, 2018, 2, 533 ss. In mancanza di tali conoscenze, infatti, il giurista si vedrà costretto a fornire un mero supporto legale agli informatici senza avere piena cognizione di ciò che questi ultimi producono o analizzano, finendo per non comprendere sino in fondo le eventuali problematiche giuridiche sottese agli algoritmi medesimi, trattandosi di un linguaggio a lui completamente ignoto.

la presenza di allucinazioni e dati di addestramento ancora inaffidabili<sup>40</sup>. Dall'altro lato, nonostante l'IA dimostri del potenziale nel fornire risposte immediate e pertinenti a una vasta gamma di domande e richieste, comprese quelle di natura legale, si è mostrato come questo modello di linguaggio – per quanto acclamato come la più avanzata delle applicazioni tecnologiche del settore<sup>41</sup> – manifesti pregiudizi morali e posizioni politiche radicali che, per quanto astrattamente condivisibili, non sono esenti da reprimenda, perché ne dimostrano la mancanza di neutralità.

Nella letteratura si sta giustamente sviluppando un dibattito sull'utilizzo responsabile delle tecnologie di IA generativa<sup>42</sup>, perché possono rappresentare una spada a doppio taglio, vuoi perché non sono interamente trasparenti sulle loro inclinazioni intrinseche, vuoi perché non si dichiarano responsabili di quanto da loro divulgato<sup>43</sup>. La peculiarità di ChatGPT, almeno stando a quanto emerso in questo studio, è che si tratta di una IA generativa con una tendenza "iperresponsiva". Nel tentativo di fornire delle risposte dettagliate a tutte le questioni che gli sono poste, l'IA risponde in ogni caso, anche quando non è certo del contenuto della risposta fornita. Il fatto che l'IA non fornisca una risposta negativa – e.g., «non so come rispondere», «non ho informazioni in merito» – pone dei seri dubbi sulla garanzia che l'utente finale possa pienamente fidarsi del contenuto dei responsi forniti dalla macchina. Il rischio evidente – e su questo si renderà necessaria una approfondita ma rapida analisi da parte di dottrina e legislatore nei tempi a venire – è che, in mancanza di una responsabilità diretta della IA, o dei suoi creatori, l'affidamento a risposte non pertinenti o erronee conduce, inevitabilmente, ad una responsabilità a cascata sugli utenti finali.

<sup>40</sup> Si vedano, sul punto, le conclusioni riportate, su diversi modelli linguistici di grandi dimensioni come ChatGPT in M. Dahl, V. Magesh, M. Suzgun, D.E. Ho, *Hallucinating Law: Legal Mistakes with Large Language Models are Pervasive* in *Stanford University Human-Centered Artificial Intelligence Center*, in <https://hai.stanford.edu/news/hallucinating-law-legal-mistakes-large-language-models-are-pervasive>, ultimo accesso 20 aprile 2024.

<sup>41</sup> In tali termini, si v. A. Tlili, B. Shehata, M. Agyemang Adarkwah, A. Bozkurt, D.T. Hickey, R. Huang e B. Agyemang, *What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education*, in *Smart Learn. Env.*, 2023, 1, 1 ss., in particolare, 15.

<sup>42</sup> Cfr., *inter alia*, M. Sallam, *ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns*, in *Healthcare*, 2023, 6, 1 ss.; E. Kasneci, K. Sessler, S. Küchemann, M. Bannert, D. Dementieva, F. Fischer, U. Gasser, G. Groh, S. Günemann, E. Hüllermeier, S. Krusche, G. Kutyniok, T. Michaeli, C. Nerdel, J. Pfeffer, O. Poquet, M. Sailer, A. Schmidt, T. Seidel, M. Stadler, J. Weller, J. Kuhn e G. Kasneci, *ChatGPT for good? On opportunities and challenges of large language models for education*, in *Learn. and Indiv. Differ.*, 2023, 102274, 1 ss.; D.R.E. Cotton, P.A. Cotton e J. Shipway, *Chatting and cheating: Ensuring academic integrity in the era of ChatGPT*, in *Inn. Educ. Teach Int.*, 2023, 19, 1 ss.; E.A.M. van Dis, J. Bollen, W. Zuidema, R. van Rooij e C.L. Bocking, *ChatGPT: five priorities for research*, in *Nature*, 2023, 7947, 224 ss.

<sup>43</sup> Sul versante pubblicistico non può non menzionarsi l'auspicata «riserva di umanità» che autorevole dottrina ha propugnato, nel «bilanciamento tra la quantità di potere che si delega a una macchina e quello che si mantiene agli esseri umani». In tal senso, si v. G. Gallone, *Riserva di umanità e funzioni amministrative. Indagine sui limiti dell'automazione decisionale tra procedimento e processo*. San Giuliano Milanese, 2023, cit. 37.

La fiducia che riponiamo nei sistemi di IA aumenta in modo significativo di giorno in giorno e il loro utilizzo nei processi decisionali è sempre più diffuso. Il risultato finale, ormai di primaria e cruciale importanza, è che dovremo, in futuro, comprendere meglio come garantire la credibilità del supporto tecnologico, soprattutto in ambito forense<sup>44</sup>. Le riflessioni qui svolte sull'uso delle IA si possono estendere a tutti i settori privatistici e pubblicistici del diritto, abbracciando anche il mondo accademico e scientifico tutto, le cui fondamenta – anche etiche – possono venire grandemente scalfite da un uso indiscriminato e poco attento di queste intelligenze generative<sup>45</sup>. Soltanto riconoscendo e affrontando le complessità morali che circondano i sistemi IA come ChatGPT, potremmo favorire un futuro più prevedibile ed equo per queste tecnologie, evitando di assaporare un frutto che, ancora acerbo per i risultati che offre, risulti poi essere avvelenato.

---

<sup>44</sup> Anche la crescente propensione a delegare completamente alle macchine decisioni giurisdizionali ha suscitato preoccupazioni in attenta dottrina, giacché un contesto del genere non terrebbe conto delle necessarie variabili umane – si parlava, poc'anzi, proprio della necessaria «riserva di umanità» – nonché dell'empatia e dei contesti specifici richiesti da tali funzioni, finendo per ridurre “in modo preoccupante gli orizzonti dell'uomo”. A tal riguardo, v. F. Fracchia, *Le fake news come luogo di osservazione dei fenomeni della tecnologia e delle reti nella prospettiva del diritto*, in *federalismi.it*, 2020, 11, 43 ss., cit., 57.

<sup>45</sup> Molti interrogativi si sollevano sulla produzione di testi da parte delle IA generative. Come si è dimostrato nel corso di questo manoscritto, la fluidità degli output di queste IA rappresenta una seria minaccia non solo per il mondo forense, ma anche per l'integrità dell'accademia stessa, giacché vi sono sempre più produzioni scientifiche che risultano essere il frutto della creatività di intelligenze artificiali. Sul punto, si v. M. Perkins, *Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond*, in *J. Univ. Teach. Learn. Pract.*, 2023, 2, 7 ss.



*Una mela acerba e ancora avvelenata: Il futuro della professione forense innanzi a ChatGPT*

Questo articolo analizza i potenziali benefici e rischi legati all'utilizzo del modello di linguaggio ChatGPT come assistente a supporto dell'attività forense e della ricerca scientifica. Dopo aver descritto le caratteristiche tecniche fondamentali di ChatGPT, l'articolo ne evidenzia le capacità nel fornire risposte dettagliate in tempi brevi, con conseguenti vantaggi in termini di produttività ed efficienza. Tuttavia, vengono analizzate anche alcune problematiche, quali la produzione di informazioni non veritiere, la presenza di pregiudizi impliciti e l'incapacità di sostituire del tutto l'esperienza umana. Emergono inoltre rilevanti questioni giuridiche in tema di responsabilità per i contenuti forniti e tutela dei dati. Sebbene ChatGPT possa rappresentare uno strumento utile di supporto, risulta ancora prematuro un suo uso indiscriminato, specialmente in ambito forense. La riflessione si conclude evidenziando la necessità di un utilizzo responsabile di queste tecnologie per favorirne uno sviluppo equo e prevedibile.

*An unripe and still poisoned apple: The future of the legal profession in front of ChatGPT*

This paper analyzes the potential benefits and risks related to the use of the ChatGPT language model as an assistant to support legal practice and scientific research. After describing ChatGPT's key technical features, the article highlights its ability to provide detailed answers in short times, with consequent advantages in terms of productivity and efficiency. However, it also analyzes some issues, such as the production of inaccurate information, the presence of implicit biases and the inability to fully replace human experience. Relevant legal issues also emerge regarding responsibility for provided contents and data protection. While ChatGPT can represent a useful support tool, its indiscriminate use still seems premature, especially in legal contexts. The reflection concludes by underscoring the need for a responsible use of these technologies to promote their equitable and predictable development.

