
Estratto

Estratto da un prodotto in vendita su **ShopWKI**, il negozio online di Wolters Kluwer Italia

Vai alla scheda →

Wolters Kluwer opera nel mercato dell'editoria professionale, del software, della formazione e dei servizi con i marchi: IPSOA, CEDAM, Altalex, UTET Giuridica, il fisco.



Wolters Kluwer

CAPITOLO XVI

DIFFAMAZIONE E INTELLIGENZA ARTIFICIALE

Temi

- *L'intelligenza artificiale generativa e la possibilità di ledere l'altrui reputazione attraverso l'utilizzo di un chatbot.*
- *Le prime vittime delle allucinazioni di AI: gli avvocati e i precedenti giurisprudenziali inesistenti citati negli atti giudiziari.*
- *Le fonti di normazione sovranazionali: l'AI ACT.*
- *Anche l'Italia corre ai ripari: la nuova legge n. 132 del 23 settembre 2025 sull'intelligenza artificiale, in vigore dal 10 ottobre 2025.*
- *Un nuovo reato (il deepfake, art. 612-quater c.p.), una nuova aggravante generale (art. 611, 1° co. undecies, c.p.) e nuove aggravanti speciali: ma manca la diffamazione.*
- *L'impossibilità di costruire una responsabilità penale diretta dell'AI: dietro la macchina c'è sempre un utilizzatore umano.*
- *I primi casi giudiziari nel mondo: NYTimes c. Microsoft Corporation e OpenAI Inc. su copyright e AI generativa.*
- *AI e il falso articolo del New York Times sugli effetti negativi del succo d'arancia.*
- *Il caso di Mark Walters, noto commentatore radiofonico statunitense, e delle false accuse di appropriazione indebita generate da AI.*
- *Altri casi in Australia, negli Stati Uniti e in Cina: il sindaco e il docente universitario accusati di reati mai commessi e la notizia dell'inesistente disastro ferroviario creato con ChatGPT che ha portato al primo arresto in Cina.*
- *I casi di deepfake di personaggi famosi: Joe Biden, Taylor Swift e Tom Hanks*
- *Gli Artificial Intelligence Crimes ovvero i reati commessi con coscienza e volontà di diffondere un contenuto falso.*
- *La diffusione inconsapevole di contenuti diffamatori rivelatisi falsi: il problema delle fake news e delle "allucinazioni" di AI e l'obbligo di verifica della fonte.*

- *L'utilizzatore professionale e i limiti per i giornalisti: l'art. 19 del nuovo codice deontologico del 2025.*
- *I pericoli per l'utente non professionale: chi usa AI si assume il rischio dei suoi errori.*
- *La possibile responsabilità del fornitore di AI e, in particolare, il possibile concorso nel reato (e nell'obbligo risarcitorio) con l'utilizzatore.*
- *La tutela civilistica offerta in via generale dall'art. 2043 c.c. e dalla specifica disciplina del danno da prodotto (la EU Product Liability Directive del 2024).*
- *La possibilità di creare un nuovo reato colposo di mera condotta.*
- *Forse sarebbe meglio smettere di chiamarla "intelligenza artificiale": la macchina elabora, non pensa.*

Sommario: **16.1.** L'alba di un nuovo mondo: avanza l'intelligenza artificiale. – **16.1.1.** L'AI generativa e i *chatbot* possibili propalatori di contenuti diffamatori. – **16.1.2.** Gli avvocati e le sentenze inesistenti citate negli atti frutto delle allucinazioni di AI. – **16.2.** Sì all'uso dell'AI, ma con giudizio: una premessa metodologica. – **16.3.** Le fonti normative sovranazionali e nazionali. – **16.3.1.** L'AI ACT del 2024. – **16.3.2.** La legge n. 132 del 23 settembre 2025: un approccio antropocentrico e settoriale. – **16.3.3.** Un nuovo reato (il *deepfake*, art. 612-*quater* c.p.), una nuova aggravante generale (art. 611, 1° co. *undecies*, c.p.) e nuove aggravanti speciali: ma manca la diffamazione. – **16.4.** AI e diffamazione: l'elemento psicologico del reato e il (falso) problema del dolo di diffamazione. – **16.5.** I primi casi giudiziari nel mondo. – **16.5.1.** Il *leading case* su *copyright* e AI generativa: *NYTimes c. Microsoft Corporation e OpenAI Inc.* – **16.5.2.** Un punto a favore di AI: il caso di Mark Walters, noto commentatore radiofonico statunitense. – **16.5.3.** Altri casi in Australia, negli Stati Uniti e in Cina: il sindaco e il docente universitario accusati di reati mai commessi e l'inesistente disastro ferroviario. – **16.5.4.** I numerosi casi di *deepfake* di personaggi famosi: Joe Biden, Taylor Swift, Tom Hanks. – **16.6.** Nessun problema per la configurabilità del delitto quando c'è coscienza e volontà di diffondere un contenuto falso (gli *Artificial Intelligence Crimes*). – **16.7.** La diffusione inconsapevole di contenuti diffamatori rivelatisi falsi: il problema delle *fake news* e delle "allucinazioni" di AI. – **16.7.1.** L'utilizzatore professionale i limiti per i giornalisti: l'art. 19 del nuovo codice deontologico del 2025. – **16.7.2.** I pericoli per l'utente non professionale: chi usa AI si assume il rischio dei suoi errori. – **16.8.** La possibile responsabilità del fornitore di AI e, in particolare, il possibile concorso nel reato (e nell'obbligo risarcitorio) con l'utilizzatore. – **16.8.1.** La tutela civilistica: la via generale dell'art. 2043 c.c. e quella specifica del danno da prodotto (la *EU Product Liability Directive* del 2024). – **16.8.2.** *De iure condendo*: la possibilità di creare un nuovo reato colposo di mera condotta. – **16.9.** Ma forse sarebbe meglio smettere di chiamarla "intelligenza artificiale": la macchina elabora, non pensa.

16.1. L'alba di un nuovo mondo: avanza l'intelligenza artificiale (AI)

Non è questa la sede per descrivere quali e quanti siano i sistemi di intelligenza artificiale esistenti oggi. Sistemi di *machine learning* capaci di emulare in maniera sempre più autonoma e indipendente il processo decisionale dell'uomo.

L'intelligenza artificiale (a seguire, indicata più semplicemente con l'acronimo anglosassone **AI**) si accinge a sostituire l'uomo in numerose mansioni. Non è difficile immaginare, ad esempio, un futuro in cui anche nel nostro Paese, come accade già in molti Stati, dove già si pensa a chi dovrà essere chiamato a rispondere di eventuali incidenti¹, sarà del tutto normale che vi siano auto che guidino da sole. Oppure – fenomeno già in voga – al *marketing* finanziario e agli investimenti “suggeriti” da AI.

16.1.1. L'AI generativa e i *chatbot* possibili propalatori di contenuti diffamatori

L'intelligenza artificiale che interessa in questa sede, al fine di verificarne le possibili interazioni con il reato di diffamazione, è quella c.d. **generativa (GenAI)**. Semplificando il concetto, si tratta di un tipo di intelligenza artificiale che è in grado di generare testi, immagini, video, musica o altri contenuti in risposta a delle richieste o dei comandi specifici da parte di un utente, detti *prompt*. Questi sistemi utilizzano modelli generativi, che producono dati a partire da un *dataset* di addestramento utilizzato per crearli. In commercio – fruibili in ormai rari casi gratis o, per lo più, a pagamento- vi sono numerose applicazioni, installabili su computer o su dispositivi mobili, quali *smartphone* e *tablet*, di AI generativa di tipo “conversazionale”, costituite da programmi informatici – denominati **chatbot** – integrati in vari canali di comunicazione, come siti *web* e motori di ricerca, *app* di messaggistica o piattaforme di *social media* utilizzati per intrattenere conversazioni in tempo reale con capacità di comprensione e risposta coerenti con quelle attese da un agente umano. Si tratta di assistenti virtuali basati su AI progettati per interagire con gli utenti simulando conversazioni simili a quelle umane.

¹ Si veda in proposito Lanzi, *Self-driving cars e responsabilità penale. La gestione del “rischio stradale” nell'era dell'intelligenza artificiale*, Giappichelli, 2023.

Il più noto di tutti è *Chat Generative Pre-Trained Transformer*, meglio noto come **ChatGPT**, che da subito si è caratterizzato per essere un *chatbot* basato su un nuovo approccio probabilistico al linguaggio naturale che, elaborando *database* prestabiliti e imparando da esempi secondo complessi algoritmi di apprendimento, arriva a generare nuovi contenuti. È il *chatbot* più diffuso, soprattutto per la scrittura creativa e si caratterizza per un'estrema versatilità soprattutto nella generazione di testi su domanda.

Accanto a quello si sono diffusi gli Assistenti AI, per cui, ormai, utilizzando un qualunque motore di ricerca, lo stesso, oltre a “dragare” il *web* alla ricerca di contenuti corrispondenti all'*input* introdotto, offre un iniziale riquadro in cui l'Assistente AI fornisce la sua risposta.

Tra quelli più noti vi sono **Gemini** (ex Bard) che è il *chatbot* di Google, **Copilot**, che è quello di Microsoft, **Grok**, che è il *chatbot* sviluppato da xAI, la startup dedicata all'intelligenza artificiale fondata da Elon Musk.

Vi sono poi altri programmi e *app* di AI che servono a creare immagini e suoni. Le più recenti indagini di mercato hanno messo in luce che, al di là dell'utilizzo professionale, di cui pure si dirà di qui a poco, il cittadino medio utilizza AI alla stregua di un motore di ricerca tipo “Google” ma più sofisticato. E, quindi, nel momento in cui è alla ricerca di informazioni su Mario Rossi, invece di digitarne il nome sulla stringa di ricerca di Google o di un altro motore di ricerca tradizionale, le chiede a *Chat GPT*. Con una differenza basilare e foriera per lui di grandi problemi e anche di responsabilità. Perché il motore di ricerca andava a reperire le informazioni su Mario Rossi dal *web*. E se non le trovava ti restituiva un messaggio “risultato non trovato”. Il *chatbot*, invece, pure utilizza quanto c'è sul *web*, ma vi associa le informazioni provenienti dal proprio algoritmo, addestrato a fornire comunque una risposta, **giusta o sbagliata che sia**.

Certo, anche la ricerca su Google o su un altro motore tradizionale poteva condurre ad un contenuto pubblicato *online* che non fosse veritiero. Ma il *chatbot* non fa questo. Per banalizzare: se non sa una risposta, la inventa.

Il problema, dunque – e questo sarà il filo conduttore della nostra analisi – è che la risposta del *chatbot* non sempre è esatta ed affidabile. Si pensi al *chatbot* che, interrogato su temi completamente differenti, rispondeva ai suoi utenti parlando di “genocidio dei bianchi in Sudafrica”.

È quanto hanno scoperto, ben presto, tanti avvocati che, piuttosto che andare ad impelagarsi nella classica e costosa ricerca sulle banche dati giuridiche dei precedenti utili al caso che stavano studiando, hanno ben pensato di richiederli più velocemente al *chatbot*, con domande del tipo: “quali sono le ultime pronunce in materia di...?”. E si sono visti fornire tutta una serie di dati relativi a sentenze, con tanto di date, nomi e contenuti. Peccato che si

trattava di **riferimenti inesistenti**. Avrebbero potuto accorgersene facilmente andando a verificarlo in una banca dati ufficiale. Ma non l'hanno fatto, perché evidentemente, nella loro ottica, lo scopo per cui si erano rivolti ad AI era proprio quello di evitare di “perdere tempo”. O, magari, una ricerca su una banca dati non erano nemmeno in grado di farla. Ed allora hanno preso, pari pari, quelle sentenze che il *chatbot* aveva indicato loro e le hanno inserite nei loro atti. Con il risultato di rendere una prestazione professionale inadeguata. Incontrando, per adesso, la benevolenza dei giudici.

16.1.2. Gli avvocati e le sentenze inesistenti citate negli atti frutto delle allucinazioni di AI

Il primo caso che è assurto alla notorietà delle cronache è avvenuto a **Firenze nel marzo 2025**, trasformando una causa civile su plagio e riproduzione di marchio industriale, in un caso-scuola su giustizia e intelligenza artificiale. Tutto era iniziato durante la discussione in aula, quando l'avvocato dell'attore aveva depositato una memoria difensiva in cui venivano richiamate delle sentenze di legittimità in tema di acquisto di merce contraffatta. Il problema, però, è che era risultato che le sentenze in questione erano inesistenti. Il legale non l'aveva negato, ma si era difeso spiegando che i riferimenti giurisprudenziali citati nell'atto erano il frutto della ricerca effettuata da una collaboratrice di studio mediante lo strumento dell'intelligenza artificiale *ChatGpt*, del cui utilizzo lui non era a conoscenza.

I giudici del Tribunale di Firenze, sezione imprese, nella motivazione della sentenza parlano di “allucinazioni giurisprudenziali” provocate dall'intelligenza artificiale. Ovvero quando “l'IA genera risultati errati che, anche a seguito di una seconda interrogazione, vengono confermati come veritieri”. Oltre alla figuraccia in aula, l'avvocato era stato accusato durante il procedimento dalle controparti di “abusivo utilizzo dello strumento processuale” per avere cercato di influenzare il giudizio del tribunale con dati fasulli, provocando così dei danni alla controparte. Il difensore del convenuto aveva anche chiesto la condanna per “responsabilità aggravata” e cioè per aver “agito o resistito in giudizio con mala fede o colpa grave”. L'avvocato dell'attore, tuttavia, pur riconosciuto l'omesso controllo, aveva evidenziato come le sue buone ragioni fossero state già provate e prescindessero dai (fasulli) riferimenti giurisprudenziali (quella che i giudici di legittimità chiamano “prova di resistenza”). E i giudici fiorentini, pur sottolineando “il disvalore relativo all'omessa verifica dell'effettiva esistenza delle sentenze risultanti dall'interrogazione dell'IA”, gli hanno dato ragione, ritenendo che

“l’indicazione di estremi di legittimità” fosse stata fatta “ad ulteriore conferma della linea difensiva già esposta” e dunque unicamente “diretta a rafforzare un apparato difensivo già noto e non invece finalizzata a resistere in giudizio in malafede”. È stata così rigettata la richiesta di condanna al pagamento di ulteriori somme, presentata dalla parte che si è poi rivelata vittoriosa.

Nel caso specifico i giudici hanno ritenuto che le loro valutazioni non fossero state influenzate da queste sentenze inesistenti perché la causa si ‘giocava’ su altri piani. Ma, come si diceva, non è detto che vada sempre così.

Ed infatti, nei mesi successivi non sono mancate le pronunce di altri giudici di merito che, contrariamente a quello fiorentino, in ragione di un utilizzo di AI senza un adeguato controllo nella redazione degli atti di causa, hanno **condannato la parte per lite temeraria ai sensi dell’art. 96, 3° co., c.p.c.**

È stato così per la sentenza del 23 settembre 2025 del Tribunale di Latina, sezione lavoro, in cui si legge: “Il ricorso giudiziario – così come tutti gli altri centinaia di giudizi patrocinati dal medesimo difensore, tutti redatti a stampone – risulta evidentemente redatto con strumenti di intelligenza artificiale; tanto è evidente non solo dalla gestione del procedimento (deposito di note *ex art. 127 ter c.p.c.* il giorno successivo al deposito del decreto di fissazione di udienza), ma soprattutto dalla scarsa qualità degli scritti difensivi e dalla totale mancanza di pertinenza o rilevanza degli argomenti utilizzati; l’atto è infatti composto da un coacervo di citazioni normative e giurisprudenziali astratte, prive di ordine logico ed in gran parte inconferenti rispetto al *thema decidendum* ed, in ogni caso, tutte manifestamente infondate. In questa situazione, si giustifica una condanna *ex art. 96 comma 3 c.p.c.*”².

E qualche giorno prima, il 16 settembre 2025 il Tribunale di Torino sez. Lavoro aveva a sua volta condannato una parte per lite temeraria ai sensi dell’art. 96, 3° co., c.p.c. sul rilievo che: “La ricorrente ha (...) agito in giudizio con malafede o, quantomeno con colpa grave, dal momento che ha proposto opposizione nei confronti di avvisi di addebito che le erano stati tutti notificati in precedenza, già oggetto di plurimi atti di esecuzione anch’essi tutti regolarmente notificati ed ha svolto – tramite un ricorso redatto ‘col supporto dell’intelligenza artificiale’, costituito da un coacervo di citazioni normative e giurisprudenziali astratte, prive di ordine logico e in larga parte inconferenti, senza allegazioni concretamente riferibili alla situazione oggetto del giudizio – eccezioni tutte manifestamente infondate”³.

² Trib. di Latina, sez. lavoro, sentenza del 23.9.2025, in www.lanuovaproceduracivile.com.

³ Trib. di Torino, sez. lavoro, sentenza del 16.9.2025, su www.lavorosi.it.

Estratto

Estratto da un prodotto in vendita su **ShopWKI**, il negozio online di Wolters Kluwer Italia

Vai alla scheda →

Wolters Kluwer opera nel mercato dell'editoria professionale, del software, della formazione e dei servizi con i marchi: IPSOA, CEDAM, Altalex, UTET Giuridica, il fisco.



Wolters Kluwer