

# OpenAI è responsabile per i danni che ChatGPT ha prodotto

Intelligenza artificiale/3

Luca De Biase, Giusella Finocchiaro e Oreste Pollicino

**L**e allucinazioni di ChatGPT possono fare del male. Un professore di diritto della Georgetown University accusato di indebite avances sessuali contro una studentessa in base a un'invenzione dell'intelligenza artificiale è diventato famoso perché il «Washington Post» ha ricostruito correttamente la vicenda e restituito al professore il suo onore. Un sindaco australiano ha dovuto difendersi da accuse infondate di corruzione che erano generate dal programma realizzato da OpenAI. E recentemente un conduttore radiofonico in Georgia, Mark Walters, è stato accusato di essersi indebitamente appropriato di denaro della Second Amendment Foundation: non era vero per niente, era una pura invenzione di ChatGPT. E Walters ha fatto causa per diffamazione contro OpenAI. Questo caso apre la strada a una nuova idea della diffamazione? Si tratta di una forma di diffamazione paragonabile alla diffamazione a mezzo stampa? È una diffamazione semplice? C'è un diritto di rettifica immediata? La tentazione di nascondersi dietro un "disclaimer" che dichiara la possibilità che ChatGPT possa avere "allucinazioni" può essere forse comprensibile, ma sarebbe anche giuridicamente accettabile? In realtà, se un'intelligenza artificiale provoca dei danni a una persona, chi l'ha progettata e messa sul mercato ne dovrebbe essere responsabile. Pesantemente. La quantità di casi simili che si possono verificare è tale che OpenAI dovrebbe essere obbligata a mettere da parte un bel gruzzolo per pagare gli avvocati e poi i danni. Ma sarà davvero così? Certamente, dal punto di vista giuridico, di diffamazione si tratta. La questione dell'analogia con la stampa si ripropone in tutta la sua evidenza, come già è stato per la responsabilità dei provider e delle piattaforme, anche a causa dell'enorme propagabilità della notizia, che può diffondersi velocissima in tutto il mondo. Oggi in Italia il reato di diffamazione è aggravato se l'offesa è recata col mezzo della stampa o con qualsiasi altro mezzo di pubblicità e sono molte le sentenze che hanno ritenuto diffamazione aggravata la diffusione di informazioni lesive attraverso i social network. Allo stesso modo si potrebbe procedere per ChatGPT, trattandosi comunque di uno strumento che comporta la pubblicità della notizia.

Non v'è dubbio che un danno sia stato cagionato a chi è stato infondatamente accusato di molestie, di corruzione, di appropriazione indebita e di reati che non avevano alcun collegamento con la realtà, magari anche con l'allegazione di prove costruite ad hoc dall'applicazione di intelligenza artificiale. OpenAI è dunque

responsabile, civilmente e penalmente, per i danni che ChatGPT ha prodotto. Poco rileva la questione della soggettività giuridica. Come già è accaduto per le auto a guida autonoma, il produttore risponde. Certo, nel caso di diffusione delle informazioni, occorre che il rimedio giunga rapidamente e che la tutela, anche cautelare, sia effettiva.

L'informazione confezionata da ChatGPT appare credibile e verosimile, anche se non sempre vera, e induce ad affidarsi. La rappresentazione della conoscenza con la pagina scritta, anche se digitale, crea affidamento. Anche se tutti sanno che Wikipedia contiene molti errori, così come blog e social. Tuttavia, il fatto di leggere le informazioni spesso attenua il senso critico. L'esigenza di rapidità fa il resto.

C'è un contrasto, fra la percezione dell'informazione e la sua qualità. Essa è presentata come affidabile, ma il disclaimer nega che lo sia.

Il codice deontologico dei giornalisti, e i principi enunciati nella famosa sentenza decalogo della Corte di Cassazione, non sono direttamente applicabili alle applicazioni di intelligenza artificiale. Perché? Anche se l'effetto può apparire il medesimo, quello di informare e produrre una rappresentazione della conoscenza, tuttavia la struttura e l'organizzazione imprenditoriale sono assai diversi. Nel caso del giornale, c'è una struttura imprenditoriale che ha lo scopo di produrre e vendere informazione. C'è un direttore responsabile che ha la funzione di vagliare le notizie. Nel caso, invece, delle applicazioni di intelligenza artificiale, lo scopo imprenditoriale è un altro e la struttura organizzativa è differente. E un principio di civiltà del nostro ordinamento giuridico, come è noto, è quello del divieto dell'analogia *in malam partem*. Ciò non significa, come già si è detto, che non ci sia responsabilità giuridica: è

comunque configurabile il reato di diffamazione aggravata e il danno deve essere risarcito. I disclaimer non consentono di escludere ogni responsabilità: sia la disciplina dei contratti che quella sulla responsabilità del produttore configurano dei limiti precisi. In sintesi: non si può, da un lato, generare un affidamento e un'apparenza e, dall'altro, escluderne le conseguenze giuridiche. Forse se non chiamassimo questi gravi errori "allucinazioni", ma invece più prosaicamente "difetti" o "malfunzionamenti", sarebbe più

facile vederne chiaramente le conseguenze giuridiche. Qui il linguaggio, che è lo strumento di ChatGPT, gioca un ruolo fondamentale: la retorica e il fascino delle parole inducono essi stessi a una diversa rappresentazione della realtà. Personalizzano, mitizzano e conducono verso una dimensione fascinosa e apparentemente lontana. Ma se le "allucinazioni" sono "errori", allora le regole ci sono già.

© RIPRODUZIONE RISERVATA

**MA SE LE  
«ALLUCINAZIONI»  
DI CHATGPT  
SONO «ERRORI»,  
ALLORA LE REGOLE  
SULLA DIFFAMAZIONE  
CI SONO GIÀ**

