
Il copyright al tempo dell'IA generativa

Francesco Posteraro

L'intelligenza artificiale può recare grandissimi benefici in termini di efficienza a numerosi comparti, dalla sanità alla tutela dell'ambiente, dai trasporti all'agricoltura, dalla sicurezza alla gestione delle infrastrutture e via dicendo. Nel contempo, come e forse più di quanto è accaduto per altre conquiste della tecnologia nell'era digitale, anche l'intelligenza artificiale può comportare gravi rischi per i diritti delle persone, tra i quali la *privacy* e la proprietà intellettuale, e per il sistema dell'informazione. Da qui la necessità che le istituzioni pubbliche intervengano per prevenire e mitigare i suddetti rischi. Purtroppo, di fronte a una tecnologia che evolve in maniera sempre più rapida, il compito del diritto assomiglia sovente alla fatica di Sisifo: fatica tanto più ardua se la si intraprende con ritardo, come si è verificato a proposito della rete internet, a lungo non sottoposta ad alcuna regolamentazione allo scopo di favorirne lo sviluppo.

L'IA generativa – ossia quella capace di produrre opere simili alle creazioni dall'ingegno umano – dà luogo a problemi che si riflettono sulla tutela del diritto d'autore e dei diritti connessi da un duplice punto di vista. Da un lato, si pone la questione della tutela dei diritti autoriali afferenti ai contenuti utilizzati per addestrare gli algoritmi. D'altro lato, ci si chiede se le opere prodotte dall'IA generativa possano a loro volta essere protette da diritto d'autore.

La possibilità di riconoscere la tutela del *copyright* a opere generate da sistemi di IA è destinata verosimilmente a trovare risposta, più e prima che nei testi legislativi, nelle aule dei tribunali. Alla luce di alcune recenti vicende processuali è già possibile tracciare una linea interpretativa. Mi riferisco in primo luogo alla decisione con la quale un tribunale federale di Washington ha respinto una domanda giudiziale volta a ottenere la protezione del *copyright* per un disegno generato da un sistema di IA denominato *Creativity Machine*, statuendo che solo alle opere create da autori umani possono riconoscersi i diritti d'autore. Il Copyright Office aveva motivato il diniego da cui ha tratto origine il giudizio con l'argomentazione che l'immagine generata da *Creativity Machine* non include alcun elemento di paternità umana. Dunque, nessuna tutela se l'opera è ascrivibile esclusivamente all'IA; se, invece, vi è un contributo dell'essere umano, allora la tutela può essere riconosciuta, ma all'uomo, non alla macchina.

Alle stesse conclusioni conduce un'ordinanza della I sezione civile della Corte di cassazione, la n. 1107 del 16 gennaio 2023, concernente l'utilizzazione da parte della RAI, quale scenografia per il Festival di Sanremo 2016, di un'opera grafica generata attraverso un *software*. Nel respingere il motivo di ricorso della RAI fondato su tale profilo – inammissibile perché non prospettato nel giudizio di merito – la Corte ha sostenuto, *incidenter tantum*, che l'uso di un *software* per generare un'immagine «è pur sempre compatibile con l'elaborazione di un'opera dell'ingegno con un tasso di creatività che andrebbe [...] scrutinato». Per cui si rende necessario, argomenta la Corte,

«un accertamento di fatto per verificare se e in qual misura l'utilizzo dello strumento» abbia «assorbito l'elaborazione creativa dell'artista». Anche per la Corte di cassazione, dunque, la tutela del *copyright* spetta esclusivamente all'essere umano, e a condizione che la tecnologia digitale sia utilizzata – come si legge nell'ordinanza – soltanto come parte del processo creativo.

Così individuato il criterio al quale attenersi, la sua applicazione concreta rimane nondimeno tutt'altro che agevole, dovendosi stabilire, nei singoli casi, se l'apporto umano abbia o meno contribuito alla creazione dell'opera in misura sufficiente a determinare il riconoscimento della tutela. Di certo, non può bastare a questo fine l'aver progettato e messo in funzione l'algoritmo: la questione, altrimenti, non si porrebbe affatto, in quanto ci sarebbe comunque, in tutti i casi, un essere umano al quale far risalire la paternità dell'opera. Si consideri, al riguardo, che il 5 settembre scorso il Copyright Office ha negato la protezione a un'immagine creata da una IA denominata Midjourney nonostante essa fosse il risultato dell'inserimento nel sistema di un numero elevatissimo di *prompt*.

Problemi più complessi e soprattutto più delicati si pongono in ordine alla tutela dei diritti d'autore e dei diritti connessi concernenti i contenuti impiegati per addestrare gli algoritmi. Negli ultimi tempi il tema si è imposto all'attenzione di numerose istituzioni pubbliche, non soltanto europee. Pochi mesi or sono il Copyright Office USA ha dato avvio a un'iniziativa volta a esaminare i problemi creati dall'IA alla disciplina del *copyright*. La World Intellectual Property Organisation si è prefissa l'obiettivo di contribuire ad affrontare le sfide poste al diritto d'autore dall'IA generativa. Ancora negli USA, su richiesta dell'Amministrazione e in vista dell'emanazione di un *Executive Order* da parte del Presidente Biden, le aziende leader del settore dell'IA hanno sottoscritto nel luglio scorso una serie di impegni, nessuno dei quali però contempla la tutela del *copyright*.

Nel contempo molti titolari hanno adottato iniziative intese a tutelare i loro diritti. Negli USA scrittori, sceneggiatori e autori teatrali hanno intentato *class action* contro Open AI e in seguito contro Google. Nel Regno Unito Getty Images ha convenuto in giudizio la piattaforma Stability AI per l'utilizzo indebito di fotografie contenute nel suo database. Il New York Times e successivamente altri media – tra cui il Washington Post, la CNN, la Reuters, Espn, Bloomberg, Disney, nonché Radio France e TF1 – hanno vietato l'accesso ai propri siti al *crawler* di Open AI, ossia al *software* che copia i dati sul web e li analizza. Lo stesso New York Times – non avendo finora prodotto risultati il negoziato con Open AI volto a ottenere una remunerazione per l'accesso alla sua banca dati – sembra inoltre intenzionato ad agire in giudizio per i materiali che il sistema di IA ha utilizzato in passato.

Nel settore musicale appare invece avviata verso un accordo la trattativa tra Universal Music e Google volta alla realizzazione di una piattaforma che permetterebbe agli utenti di creare musica basandosi su brani già presenti sul mercato e ai titolari dei diritti di ottenere un ricavo per l'addestramento dell'algoritmo.

Se negli USA e nel Regno Unito la soluzione del problema sarà con ogni probabilità influenzata dagli esiti dei procedimenti giudiziari, nell'UE è stato avviato fin dall'aprile 2021 l'iter di un regolamento, l'AI Act, volto ad assicurare che lo sviluppo della nuova tecnologia non rechi pregiudizio ai diritti e ai valori protetti dall'ordinamento europeo.

Quando è stata presentata la proposta di regolamento l'IA generativa non aveva ancora fatto la sua comparsa. Per disciplinarne l'uso, il PE ha approvato nel giugno scorso un emendamento – destinato probabilmente a essere accolto nel testo definitivo dell'AI Act – che reca l'obbligo per gli sviluppatori dei sistemi di GenAI di rendere noti i dati protetti da diritto d'autore utilizzati per addestrare gli algoritmi. Imponendo ai fornitori di sistemi di IA generativa soltanto un obbligo di trasparenza, l'emendamento non fornisce dunque alcuna indicazione in ordine alla liceità o meno dell'uso di contenuti protetti da *copyright*, ma si limita a fare salva la vigente normativa europea in materia di diritto d'autore. Alla quale occorre pertanto riferirsi per individuare le disposizioni applicabili nella fattispecie.

Deve innanzi tutto escludersi che nei confronti dell'utilizzazione dei contenuti da parte di un sistema di IA generativa i diritti d'autore e i diritti connessi possano trovare protezione nell'art. 17 della Direttiva *Copyright*. Tale utilizzazione non può infatti configurarsi in alcun modo come un atto di comunicazione al pubblico, lecito, ai sensi del citato art. 17, solo a condizione di essere stato previamente autorizzato dai titolari dei diritti.

La norma cui fare riferimento è invece quella dell'art. 4 della stessa direttiva, che prevede l'eccezione di *text and data mining*. Prevede, cioè, che sia consentita l'estrazione di testo e di dati da opere o da altri materiali cui si abbia legalmente accesso (compresi dunque i contenuti disponibili online), a meno che il loro utilizzo non sia stato espressamente riservato dai soggetti cui spettano i diritti esclusivi di riproduzione. I titolari dei diritti possono quindi, esercitando un *opt-out*, sottrarre le proprie opere all'operatività dell'eccezione ed evitarne in tal modo l'uso ai fini dell'addestramento dei sistemi di IA. In questo quadro, l'adempimento dell'obbligo di trasparenza di cui all'emendamento del PE può servire a verificare se siano stati utilizzati contenuti per i quali i titolari abbiano esercitato la riserva.

Non occorre sottolineare che un meccanismo di *opt-in* sarebbe stato assai più efficace dell'*opt-out* ai fini della tutela del *copyright*. E se si deve supporre che all'epoca i legislatori europei nulla sapessero del futuro avvento dell'IA generativa, l'AI Act avrebbe fornito l'occasione, a quanto pare purtroppo non colta, per cambiare rotta in direzione di una più energica difesa dei diritti autoriali.

La direttiva precisa che i titolari dei diritti devono esercitare la riserva «in modo appropriato, ad esempio attraverso strumenti che consentano una lettura automatizzata in caso di contenuti resi pubblicamente disponibili sul web». L'art. 70-*quater* della legge sul diritto d'autore, con il quale è stato recepito nell'ordinamento italiano l'art. 4 della direttiva, non reca invece alcuna prescrizione in proposito. Con riferimento ai contenuti disponibili online l'impiego di strumenti tali da consentire una lettura automatizzata sembra tuttavia indispensabile per rendere possibile il riconoscimento della riserva da parte dei *crawler*.

Nei singoli Paesi membri dell'UE l'eccezione di *text and data mining* opera ovviamente a decorrere dal recepimento della direttiva nei rispettivi ordinamenti nazionali. Ciò significa che l'utilizzo di contenuti protetti da *copyright* avvenuto prima del suddetto recepimento deve considerarsi illecito; e può invece ritenersi lecito, a partire dal recepimento, a condizione che – e fino a quando – i titolari dei diritti non abbiano esercitato l'*opt-out*.

Relativamente all'accertamento delle eventuali violazioni dei diritti commesse prima dell'entrata in vigore dell'eccezione di *text and data mining* si pone però un problema di prova pressoché insolubile. Anche nel caso di *web scraping* eseguiti lecitamente sussistono peraltro difficoltà probatorie con riferimento al rispetto dell'obbligo – sancito dall'art. 4, par. 2, della Direttiva *Copyright* – di conservare le riproduzioni solo per il tempo necessario ai fini dell'estrazione di testo e di dati.

Quanto alle eventuali violazioni commesse dopo l'esercizio dell'*opt-out*, spetta alla tecnologia fare in modo di impedirle o almeno di individuarle. Nei confronti delle piattaforme e dei motori di ricerca designati come molto grandi a norma del DSA si potrebbe anche utilizzare a questo scopo la previsione dell'art. 37 di quel regolamento, che prevede la possibilità di sottoporre tali operatori a un *audit* esterno e indipendente. Un'ultima notazione prima di concludere: assicurare la tutela dei diritti d'autore e dei diritti connessi può non equivalere, di fronte all'IA generativa, a tutelare anche i titolari. Non sembri un gioco di parole. Se pure dovesse continuare a svilupparsi rispettando il *copyright*, e dunque acquisendo i contenuti protetti solo in accordo con i titolari e dietro compenso, l'IA potrebbe comunque fare concorrenza all'industria culturale e quindi rischiare nel tempo di impoverirla. E il gioco non sarebbe in questo caso a somma zero, perché trovando nei settori creativi minori risorse cui attingere i sistemi di IA finirebbero nel lungo termine per produrre a loro volta contenuti di più modesta qualità. L'auspicio è che il diritto e la tecnologia sappiano scongiurare l'avverarsi di questo scenario distopico.