

Dipartimento di Scienze Giuridiche
European School of Economics
Direttore Prof. Giuseppe Cassano

INTELLIGENZA ARTIFICIALE E DIRITTO D'AUTORE

*Guida operativa per giuristi alla ricerca di soluzioni dal legislatore
e da tutta la giurisprudenza nota di diritto interno e comparato*

GIUSEPPE CASSANO - VIRGILIO D'ANTONIO
BRUNO TASSONE - VIRGINIA ZAMBRANO

Abstract

La tutela del diritto d'autore nell'era di ChatGPT richiede una riflessione approfondita su alcuni temi specifici, che l'evoluzione normativa più recente ha in parte ridefinito. Il lavoro muove dalla constatazione che ad ogni evoluzione tecnica di rilievo ha sempre fatto da contraltare la necessità di una rimodulazione della tutela autoriale, come dimostrano le vicende delle opere fotografiche, delle banche dati e del software; e ricostruisce anzitutto la nozione di sistema di IA accolta dall'art. 3 dell'AI Act, richiamato dalla stessa Legge n. 132/2025.

Si passa poi al quesito con cui ci si chiede se le opere generate dall'intelligenza artificiale possano essere considerate o no «opere dell'ingegno» e quindi protette dal diritto d'autore: tema particolarmente rilevante se si considera che tali opere sono spesso il risultato di un processo di apprendimento automatico che utilizza dati e informazioni già esistenti. L'indagine verifica allora la configurabilità di un autore «artificiale» attraverso una ricognizione comparata - dal caso Naruto alle linee guida dello U.S. Copyright Office, dalla legge cinese alla giurisprudenza della Suprema Corte messicana, sino alle decisioni dell'EPO - sostenendo che il riferimento normativo alla persona fisica, interno ed euro-unitario, è storicamente volto a delimitare l'acquisto in capo alla persona giuridica e non opera ad excludendum verso l'IA; da cui l'ipotesi ricostruttiva di una «soggettività dimezzata», oggi però scartata dalle scelte del legislatore del 2025.

La novella recata dalla Legge 23 settembre 2025, n. 132, infatti, restringe la tutela alle opere dell'ingegno umano, ammettendo quelle realizzate con l'ausilio di strumenti di IA purché costituenti risultato del lavoro intellettuale dell'autore. Ne deriva una matrice categoriale a quattro quadranti, le cui zone intermedie vengono perimetrare alla luce della giurisprudenza della Corte di Giustizia sulle «scelte libere e creative», articolando il processo creativo nelle fasi di concezione, esecuzione e redazione. In tale prospettiva il prompt - input

fornito all'algoritmo e momento fondamentale nella creazione dell'opera – può rilevare non già quale autonomo oggetto di protezione, bensì quale sede elettiva delle scelte creative dell'autore e criterio di misura del suo apporto intellettuale, segnatamente nella zona grigia del prompt crafting. Si esamina quindi la sorte delle opere che, per difetto di apporto umano percepibile, restano prive di tutela.

Passando ad altro tema centrale, il lavoro si sofferma poi sulla gestione dei diritti relativi ai dati utilizzati per addestrare gli algoritmi: venendo ad analizzare il nuovo art. 70-septies l.d.a., introdotto dall'art. 25 della Legge n. 132/2025, i suoi rapporti con gli artt. 70-ter e 70-quater, con l'art. 4 della direttiva (UE) 2019/790 e con l'art. 53 dell'AI Act, il regime dell'opt-out e le sue modalità di espressione, nonché l'inidoneità dell'eccezione a coprire la fase di output. Viene così in rilievo la possibilità che il modello generi opere lesive dei diritti di terzi, con i conseguenti interrogativi sulla responsabilità di fornitori e deployer.

Seguono poi l'esame del Codice di buone pratiche e dei codici di condotta ex art. 95 dell'AI Act, dei quali si propone una qualificazione in termini di espressione della regolazione «mista» e se ne argomenta il rilievo tanto ai fini delle sanzioni ex artt. 99-101 quanto della responsabilità civile (artt. 2043 e 2050 c.c.); nonché l'analisi del ruolo dell'AGCOM e dell'ingiunzione dinamica amministrativa. Le riflessioni conclusive assumono infine il bilanciamento di interessi quale filo conduttore, proponendo criteri operativi per l'accertamento dell'apporto creativo umano e segnalandone le criticità sistemiche di medio periodo.

Il lavoro è poi corredato dalla pubblicazione delle sentenze e, ove il procedimento non sia concluso, dei principali provvedimenti resi nei contenziosi ad oggi più rilevanti in materia su entrambe le sponde dell'Atlantico - Kneschke v. LAION, GEMA v. OpenAI, Getty Images v. Stability AI, Bartz v. Anthropic, Concord Music v. Anthropic e Disney Enterprises v. Midjourney - ciascuno accompagnato da apposita nota di commento.

SOMMARIO:

SEZIONE 1. 1. INTRODUZIONE. – 2. EVOLUZIONE DELLA TECNICA E IA. – 3. OPERE GENERATE DALLA IA E DIRITTO D'AUTORE. – 4. IA E INFORMAZIONE VISIVA. – 5. L'AUTORE ARTIFICIALE: FU VERO AUTORE? – 6. (SEGUE:) LE OPINIONI FAVOREVOLI. – 7. (SEGUE:) LA SOGGETTIVITÀ DIMEZZATA. – 8. L'IA, LE MODIFICHE INTRODOTTE DAL LEGISLATORE CON LA LEGGE 132/2025 E LA GIURISPRUDENZA DELLA UE. – 9. (SEGUE:) ORIGINALITÀ PROGRESSIVA, FASI DEL PROCESSO CREATIVO E OPERE IN PUBBLICO DOMINIO. – 10. ADDESTRAMENTO DEI SISTEMI DI IA E TUTELA AUTORIALE. – 11. CODICE DI BUONE PRATICHE E CODICI DI CONDOTTA PER L'IA. – 12. CO-REGOLAZIONE, AUTOREGOLAZIONE E ALTRE RIFLESSIONI SUL RUOLO DELLA AUTONOMIA PRIVATA. – 13. IL RUOLO DELL'AGCOM. – 14. RIFLESSIONI CONCLUSIVE ATTORNO AL BILANCIAMENTO DI INTERESSI. – 15. POSTFAZIONE (RETHINKING COPYRIGHT: LA PROPRIETÀ INTELLETTUALE ALLA PROVA DEI GRANDI MODELLI DI IA). – 16. POSTFAZIONE (LA PERSONA NELL'ECOSISTEMA DIGITALE: TUTELA DEI DATI E NUOVE FORME DELLA PROPRIETÀ INTELLETTUALE).

SEZIONE 2. RASSEGNA DI GIURISPRUDENZA STRANIERA RAGIONATA E COMMENTATA

SEZIONE 1*

1. Introduzione (IA e proprietà intellettuale: la metamorfosi dogmatica del concetto di creazione).

L'intersezione tra sistemi di Intelligenza Artificiale (da ora in poi IA) e disciplina della Proprietà Intellettuale (da ora in poi PI) si pone all'origine di una delle più complesse aporie dogmatiche dell'ordinamento giuridico contemporaneo. Ed infatti, storicamente incardinata sul presupposto immanente dell'esclusività dell'apporto creativo e inventivo umano, l'impianto dogmatico delle privative industriali e intellettuali — si trova oggi a fare i conti con una produzione di beni immateriali

(opere artistiche, software, invenzioni e modelli industriali) che si avvale di processi automatizzati di tipo autonomo o semi-autonomo.

Il che — escluso che sotto il profilo giuridico possa trovare approdo in soluzioni ermeneutiche emergenziali — solleva più di una questione. In discussione, a ben vedere, è la tenuta concettuale delle categorie fondative del diritto d'autore e del diritto brevettuale, atteso che requisiti quali quello di creatività; impronta personale dell'autore; inventive step o, ancora l'individuazione della person skilled in the art appaiono configurati muovendo da una evidente prospettiva eminentemente antropocentrica.

Si tratta di una prospettiva che non regge a fronte dell'irruzione sul mercato di modelli linguistici di grande scala, reti avversarie generative, modelli di diffusione e così via. Si assiste, infatti, ad una torsione della dimensione antropocentrica che è chiamata a fare i conti con un mutato concetto di creazione. A cadere in discussione è la stessa unitarietà di un processo creativo condizionato dalla stretta interrelazione input/output, la cui decostruzione consente di cogliere la peculiarità di siffatta metamorfosi. Ed invero, la collisione della proprietà intellettuale con la tecnologia si apprezza già a partire dal fatto che — nell'era dell'IA — l'apporto dell'uomo si trasforma in atto materiale di impulso. Un atto sprovvisto per l'algoritmo di qualsivoglia densità semantica.

Certo l'attività di addestramento dei modelli trae vantaggio dalla elaborazione ed estrazione di una grande quantità di dati alcuni dei quali, non raramente, tutelati dal diritto d'autore. Ma — come ben evidenziato da un case law che va progressivamente formandosi in tema — non si ha a che fare con una memorizzazione diretta dell'opera. In discussione infatti sono (a dispetto dell'enfasi posta sul lemma "intelligenza") modelli matematici in grado di operare correlazioni statistiche, produrre pattern sintattici e dar vita a rappresentazioni che creano l'apparenza della predittività.

* Sebbene il saggio sia dovuto all'opera congiunta degli Autori, l'Introduzione va attribuita a **Virgilio D'Antonio** e **Virginia Zambrano**; a **Giuseppe Cassano** i Par. 1-4, a

Bruno Tassone i par. 5-14; a **Virginia Zambrano** il par. 15 e a **Virgilio D'Antonio** il par. 16.

Certo, piaccia o no, questa fase di training implica pur sempre che sul server dello sviluppatore si origini una attività di “riproduzione” più o meno temporanea o strumentale dei dati.

Certo (almeno in Europa) l’art. 2 Dir. 2001/29/CE sembra accorgersi di tutto ciò concedendo agli Stati il diritto esclusivo di autorizzare o vietare la riproduzione delle opere. Eppure non v’è chi non veda come si tratti di previsione “tormentata” dal dilemma della difficoltà di conciliare opposti interessi. Complessità di una opzione che (abbandonando la prospettiva eurounitaria) la comparazione giuridica disvela pienamente, evidenziando come – a fronte di un medesimo dilemma – differenti siano le strategie di bilanciamento degli interessi utilizzate al di qua e al di là dell’oceano.

Non a caso, in Europa, la Direttiva UE 2019/790 – DSM nell’introdurre una disciplina speciale per il Text and Data Mining ha individuato, per un verso, un regime di eccezione destinato a valere – sul presupposto dell’interesse alla conoscenza – per enti di ricerca e istituti di tutela del patrimonio culturale (art. 3) e, per l’altro, ha subordinato la liceità dell’uso a fini commerciali di contenuti pubblicamente accessibili online, al mancato esercizio della facoltà di un opt out (art. 4) la cui effettività si è scontrata con l’assenza di standard tecnici univoci e con la frammentazione dei metadati.

Tanto senza considerare che l’assenza di rigorosi obblighi (su cui è opportunamente intervenuto l’AIA ristrutturando la dimensione probatoria e articolandola su due pilastri trasversali quali 1) l’obbligo dei fornitori di modelli di IA di redigere e rendere pubblica una sintesi dettagliata dei dataset impiegati per l’addestramento; 2) l’adozione di tecniche di marcatura digitale (watermarking) per garantire la tracciabilità degli output) e l’opacità dei dataset rendeva praticamente impossibile assolvere all’onere probatorio circa l’avvenuta utilizzazione non autorizzata della propria opera. In questo senso, il vero fronte delle contese giudiziarie è rappresentato proprio dalla fase di addestramento e acquisizione dei dati (input).

Negli Stati Uniti, vigilata dalla fair use doctrine (17 U.S.C. § 107) la liceità dell’addestramento è invece – come noto – costantemente chiamata a muoversi entro i confini di un transformative use la cui interpretazione è all’origine di un vigoroso contenzioso (tra cui le cause Andersen v. Stability AI, Tremblay v. OpenAI e Getty Images v. Stability AI) che sta mettendo a dura prova la tenuta di quella dottrina. L’ubi consistam di queste azioni collettive promosse da autori, illustratori ed editori contro i principali sviluppatori di modelli generativi non solo ha a che vedere con la interpretazione di quel “famoso” quarto componente della fair use doctrine rappresentato dalla valutazione del pregiudizio al mercato dell’opera originale. Nelle riflessioni giurisprudenziali più mature l’algoritmo non “legge per ispirarsi”, ma compie atti materiali di riproduzione tecnica e vettorizzazione industriale su scala massiva di cui non può non tenersi conto.

Né, spostando la lente di osservazione sul piano della qualificazione giuridica degli output prodotti dall’IA le cose sembrano andare meglio. Qui l’attrito con i principi generali del diritto d’autore arriva ad un punto (se possibile) ancora più critico, e che né l’approccio degli ordinamenti di Civil Law né di Common Law riesce effettivamente a chiarire. In entrambi i casi la tutelabilità del bene immateriale presuppone che esso costituisca indefettibilmente il “risultato dell’attività creativa della persona umana”, riflettendone la personalità e le scelte libere e creative. In questa direzione, nelle prassi applicative dell’US Copyright Office e, di recente, in Thaler v. Perlmutter, No. 23-5233 (D.C. Cir. 2025) i giudici non esitano a precisare che, sebbene il Copyright Act non definisca la parola autore “traditional tools of statutory interpretation show that, within the meaning of the Copyright Act ‘author’ refers only to human beings” (pag. 10).

Sotto la ruga antropocentrica si nasconde tuttavia l’esigenza di approdare ad una regolazione dell’IA generativa che investa in modo coordinato le tre fasi fisiologiche del fenomeno: l’addestramento (input), la

tracciabilità (processo) e la valorizzazione dei risultati (output).

Sul piano comparato, costituisce allora un'eccezione la previsione del Copyright, Designs and Patents Act inglese 1988, sec. 9(3). Si tratta di una previsione che promuove una logica utilitarista, pragmatica ed economicamente incentivante e che lascia trasparire il timore che beni di valore commerciale cadano nel pubblico dominio per il solo difetto di un contributo creativo umano diretto. Un legislatore inglese in preda all'horror vacui, nel disciplinare le opere computer-generated, ha così assegnato la titolarità dei diritti patrimoniali a chi ha posto in essere le condizioni per la genesi dell'opera.

Non sfugge, tuttavia, come la norma – originariamente pensata per software rudimentali – sia del tutto inadeguata a gestire le problematiche proprietarie legate all'IA generativa. La ratio incentivante della privativa rischia di capovolgersi nel suo esatto opposto, arrivando a giustificare la concessione di monopoli legali su una mole potenzialmente illimitata di beni generati a un costo marginale pressoché azzerato.

Preso atto della metamorfosi della creatività, su tutto sembra allungarsi l'ombra dell'irrisolto rapporto uomo-macchina. Che si tratti di snodo cruciale è evidente. In discussione è la necessità di “misurare” la rilevanza dell'apporto umano nel processo generativo, stante il confine sempre più labile che separa la mera formulazione concettuale di un'idea dalla sua concreta esecuzione espressiva. Laddove è chiaro che se il contributo umano si risolve in mere indicazioni generiche viene meno la soglia minima di creatività richiesta dalla legge, come non manca di precisare la più recente giurisprudenza (Zarya of the Dawn e Théâtre D'opéra Spatial). Mentre, per contro, la dimostrazione della centralità dell'apporto della persona fisica sul processo di generazione non può non riflettersi sul regime di tutela, sia pur limitatamente alle componenti espressive direttamente imputabili all'intervento dell'uomo.

L'utente che inserisce un prompt assume le vesti del committente che affida un incarico a un

artista: fornisce l'idea, l'indirizzo concettuale o il tema di fondo, ma la trasposizione in forma espressiva viene eseguita in autonomia dal software. Perché possa configurarsi una tutela autorale, l'apporto umano deve dunque manifestarsi o in una fase post-generativa di selezione, modifica e rielaborazione sostanziale dell'elaborato, ovvero in una combinazione complessa di elementi direttamente conformati e nel rispetto dello human-in-the-loop principle. Né il carattere determinante dell'apporto umano riguarda solo le opere artistiche o letterarie. Esso si presenta in chiave speculare e ancora più rigida sul versante della proprietà industriale. La centralità della persona umana sembra qui trovare conferma anche attraverso l'itinerario tracciato da uffici brevettuali e corti di settore (tra cui l'EPO - EPO Board of Appeal, decisione J 0008/19 - , l'USPTO e la UK Supreme Court) che, a partire dal caso DABUS (Device for the Autonomous Bootstrapping of Unified Sentience), hanno manifestato consenso nell'affermazione del principio secondo cui l'inventore designato nella domanda di brevetto deve possedere indefettibilmente la qualifica di persona fisica.

Come dire che l'assenza di capacità giuridica e di agire rappresenta ad oggi un ostacolo insormontabile per il riconoscimento della soggettività giuridica del software. Tanto, a dispetto di tentativi isolati di provocazione giuridico-politica, come è accaduto nel caso del robot umanoide Sophia cui l'Arabia Saudita ha concesso simbolicamente la “cittadinanza” nel 2017.

Al netto di queste provocazioni un dato rimane. L'attribuzione della cittadinanza o di una fittizia e-personhood alla macchina non riesce a risolvere i problemi della proprietà intellettuale né della responsabilità civile rischiando, al contrario, di creare un pericoloso schermo di responsabilità per gli sviluppatori umani. Il che fa intendere come, a ben riflettere, l'approccio antropocentrico si confermi quale unico presidio di razionalità del sistema delle privative. La privativa brevettuale costituisce un monopolio temporaneo concesso dallo Stato quale incentivo economico per i soggetti che

sostengono il rischio finanziario e intellettuale dell'innovazione, non già un riconoscimento automatico attribuibile a un algoritmo operante a costi marginali azzerati.

Insomma, sia che si guardi al problema dal punto di vista dogmatico, sia che si adotti una prospettiva economico-funzionale, è difficile liberarsi della componente umana che – come Erik nel Fantasma dell'Opera – continua a innervare siffatta materia.

Ineludibilità di un fattore umano, non solo nella dimensione fisiologica dell'attribuzione dei diritti di privativa, che si avverte con pari vigore allorché si sposti l'attenzione alla dimensione patologica dell'allocazione dei rischi e delle responsabilità. Nel momento in cui l'algoritmo valica i confini della mera rielaborazione per ledere l'esclusiva altrui, il problema della centralità dell'uomo si sposta dall'individuazione dell'autore all'imputazione dell'illecito. In questo contesto presunzione di accesso, criterio della convergenza espressiva e frammentazione dell'imputazione si rivelano gli snodi di un sistema probatorio che se è vero che ha il pregio di adattare la prova dell'accesso al contesto digitale, è anche vero che mantiene una rigidità concettuale sul lato dell'output e dell'imputazione. Da questo punto di vista, l'IA generativa finisce con l'essere equiparata all'"esecutore materiale" di un plagio, del tutto trascurando che l'illecito algoritmico è un processo distribuito, probabilistico e privo di una singola volontà cosciente. Il che, sposta il discorso sulla difficoltà di prova di un nesso causale il quale si smarrisce fra gli anelli della catena dei soggetti che hanno preso parte alla "produzione" del risultato, con la conseguenza, per un verso, di rendere complessa la dinamica risarcitoria e, per l'altro, di spingere il mercato verso clausole di manleva o forme di assicurazione obbligatoria.

Se, alla luce di quanto si è andato allineando, è possibile approdare ad una conclusione, questa va nella direzione di una disciplina dell'IA che prenda le distanze dalla creazione di fittizie soggettività algoritmiche. In discussione è piuttosto la giusta allocazione di responsabilità e potere economico tra chi sviluppa la tecnologia

e chi fornisce la materia prima creativa. Fermo restando la centralità della componente umana, quale unico momento di legittimazione della tutela d'autore e brevettuale, la vera battaglia sembra destinata a spostarsi sul piano della gestione della fase di input, in un contesto destinato a segnare un graduale passaggio dal piano dello scontro giudiziale a quello delle licenze collettive e degli accordi negoziali di massa, quali unici strumenti capaci di assicurare un equo compenso alla creatività posta alla base dell'addestramento algoritmico.

In questo senso, lungi dal proporre l'adattamento di vecchie logiche a nuove realtà, si tratta di ripеримetrare la funzione economico-sociale della privativa, dinamicamente bilanciando l'ecosistema della creatività e dell'espressione dell'essere umano con la necessaria remunerazione degli investimenti nell'innovazione tecnologica.

2. Evoluzione della tecnica e IA

È sotto gli occhi di tutti che l'intelligenza artificiale (di seguito "IA") sta da qualche anno producendo grandi cambiamenti nella realtà giuridica, oltre che – fra le altre – in quella politica, sociale, economica, educativa, ambientale e lavorativa.

A tali cambiamenti non si sottrae il diritto d'autore, che anzi è, per certi versi, più inciso da essi e, per altri, può dare un particolare contributo allo sviluppo del dibattito sulle migliori forme della regolazione della IA.

In proposito, uno spartiacque storico è certamente rappresentato dalla creazione di applicativi come il ben noto – e ormai iconico – ChatGPT (*Generative Pre-trained Transformer*), definibile come una classe di modelli di IA basati sul *deep learning*, ossia progettati per comprendere, elaborare e generare linguaggio naturale in modo simile a quello umano, che ha in brevissimo tempo ridefinito il campo dell'elaborazione del Linguaggio Naturale stesso (NLP).

In effetti, proprio mentre conducevamo alcuni studi sulla soggettività il 3 novembre 2022 aveva luogo il lancio ufficiale di detta *chatbot*, tramite la quale qualsiasi utente registrato è stato subito

messo in grado di interagire come se avesse di fronte un essere umano e – soprattutto – di chiedere la redazione di un testo (fosse esso la *brochure* di presentazione di una casa vinicola che opera nell'Oltrepò Pavese o un tema sulle campagne napoleoniche), differenziandolo anche in base al destinatario (ad esempio, chiedendo di spiegare il concetto di base ad un minore di sei anni)¹.

Oggi sembrano passati non tre, ma trenta anni, considerando le successive versioni di ChatGPT e i suoi omologhi come Gemini e Claude: ma già allora i risultati erano sorprendenti perché, grazie all'auto-apprendimento, lo stesso migliorava la sua capacità relazionale man mano che dialogava con l'utente e già al "primo tentativo" di scrittura forniva un prodotto "pressoché perfetto", senza che – ecco il punto – i suoi contenuti e la sua forma espressiva fossero in alcun modo prevedibili da parte del programmatore².

Di qui, allora, il titolo di questa nostra digressione sull'impatto dell'IA sul diritto d'autore e sulla sua tutela, appunto nell'era di ChatGPT.

Al riguardo, metterà conto chiedersi, innanzi tutto, se le opere generate dall'IA possano essere considerate "opere dell'ingegno" e quindi opere protette dal diritto d'autore, per poi soffermarsi sul diverso quesito con cui ci si domanda come avvenga la gestione dei diritti di proprietà intellettuale relativi ai dati utilizzati per addestrare gli algoritmi di IA e delle stesse "creazioni" della IA: mentre un distinto quesito riguarderà da un lato l'eventuale tutela del *prompt* (ossia l'input fornito all'algoritmo per generare un'opera, che da taluni è considerato un elemento fondamentale nella creazione dell'opera stessa), dall'altro la possibilità che esso sia alla base della creazione di opere in violazione dei diritti di proprietà intellettuale di terzi, con ogni conseguente responsabilità.

Questi, dunque, i principali argomenti che verranno qui trattati mediante una rassegna dei temi e problemi che essi pongono, al fine di consentire al lettore una prima – ma non superficiale – informazione. Gli stessi, inoltre, saranno completati dalla pubblicazione delle sentenze e (in assenza di queste, laddove il procedimento non è concluso) dei principali provvedimenti dei sei giudici che attualmente stanno interessando i giudici su entrambe le sponde dell'Atlantico in materia di IA e diritto d'autore, con il corredo di apposite note di commento.

In proposito, è noto che alla fine del 2022 sono state avviate negli Stati Uniti azioni individuali e di classe di artisti che si sentono lesi per il fatto che algoritmi come quelli usati dai *software* Midjourney, Dall-E, Unstable Diffusion e altri – accessibili gratuitamente o con piani di abbonamento con funzionalità avanzate – consentono agli utenti di far realizzare immagini complesse digitando semplici frasi: ciò grazie al fatto che gli stessi algoritmi vengono "addestrati" usando (letteralmente milioni di) immagini, molte delle quali protette dal diritto d'autore, sicché alcuni primari siti che gestiscono portafogli artistici come DeviantArt o ArtStation hanno iniziato a offrire agli iscritti la possibilità di segnalare le proprie opere come non utilizzabili ai fini indicati³.

Ebbene, dopo aver notato che consimili casi si sono ancora più di recente avuti nel Vecchio Continente, alle società che offrono il servizio si contesta in sostanza il *download* massivo delle immagini presenti su *internet* – se del caso grazie alla loro estrazione automatica con tecniche di *scraping* – affinché esse siano impiegate per permettere all'IA di imparare a creare "nuove" immagini senza che vi sia alcun consenso degli aventi diritto: anche se, come dimostra sempre la raccolta giurisprudenziale di cui in appresso, l'asse del contenzioso si sta spostando dalla

¹ Ne davamo conto con B. TASSONE, *Riflessioni su intelligenza artificiale e soggettività giuridica*, in *Dir. internet*, 2023, 213 ss.

² Così C. MORELLI, *La ChatGPT che scriveva come un avvocato*, disponibile presso www.altalex.it, la quale rilevava altresì che la capacità di aggregazione delle

conoscenze avrebbe già a quel tempo permesso al programma di superare l'esame di avvocato.

³ Sul tema, fra gli altri, C. GEIGER e V. IAIA, *The forgotten creator: Towards a statutory remuneration right for machine learning of generative AI*, in *Computer Law & Security Review*, 52, 2024, n. 105925.

liceità del *training* e dall'utilizzo dei *dataset*, alla responsabilità derivante dagli *output* prodotti dai modelli generativi e dalle modalità di reperimento dei materiali destinati all'addestramento.

Peraltro, l'analisi di tale giurisprudenza fa capire perché in tale contesto è più che mai necessario che il giurista si confronti con la fattispecie concreta e la comprenda a fondo, altresì considerando che la complessità tecnica dei casi portati all'attenzione dei giudici induce a chiedersi se occorranzo nuove prerogative autoriali o se quelle esistenti offrano una sufficiente tutela⁴.

Quel che è certo è che allo stato l'apparato tradizionale del diritto d'autore continentale ha portato ad alcune recenti "vittorie" degli aventi diritto, in qualche modo preconizzate da chi riteneva che le forme di sfruttamento economico già previste dalla legge potessero servire allo scopo⁵.

La diffusione incessante, inarrestabile e quasi autoritaria, dell'IA nella quotidianità di singoli, imprese e soggetti pubblici tutti, pone nuove

frontiere al dibattito in tema di tutela del diritto d'autore. Né potrebbe essere diversamente, giacché ad ogni evoluzione tecnica di rilievo ha fatto sempre da contraltare la necessità della rimodulazione della tutela qui in esame.

L'intreccio tra evoluzione della tecnica e diritto d'autore non è dunque un tema esclusivo dei nostri giorni, né proprio della IA generativa (qui indicata anche come GenAI), sulla quale si è di recente soffermato anche il Garante per la Protezione dei Dati Personali⁶.

In proposito, è intanto del tutto ovvio che il problema del rapporto fra IA e diritto d'autore si pone solo o prevalentemente rispetto agli applicativi dotati di maggiore capacità elaborativa e, al riguardo, proprio la "autonomia" di cui essa è dotata fa sorgere non pochi interrogativi anche in altri contesti.

Del resto, quando alla stessa IA è stato chiesto alcuni anni fa di definirsi grazie all'uso di un altro linguaggio auto-regressivo basato sulla tecnica GPT-3, la stessa ha indicato proprio gli elementi di cui si tratta fra quelli che la caratterizzano aggiungendo, fra gli altri, la capacità di adattamento, la capacità di risolvere problemi e

⁴ L'interrogativo di cui nel testo trova riscontro nel numero di *causes of action* poste alla base dei primi contenziosi. Ad esempio, nel procedimento citato nel testo – il n. 3:23-cv-00201-WHO, pendente di fronte alla United States District Court del Northern District of California – nella *Opposition to Defendants Motion to Dismiss* (depositata quale documento n. 65 in data 2 giugno 2023 e reperibile presso <www.courtlistener.com>) si rileva che «the Complaint contains at least three distinct theories of direct infringement [...]: (1) direct infringement by distributing Stable Diffusion, which contains compressed copies of the training images; (2) direct infringement by creating and distributing their DreamUp, which is itself an infringing derivative work; and (3) generating and distributing output images which are infringing derivative works». Di poi, nel giudizio avviato alla fine del 2023 dal New York Times contro OpenAI e Microsoft per aver usato senza consenso milioni di suoi articoli per addestrare le rispettive chatbot, fra cui ChatGPT e Copilot (sul presupposto che le stesse sarebbero potenziali suoi *competitor* che per rispondere alle domande degli utenti impiegano materiali protetti dal diritto d'autore), si fa leva sui *trademarks* e sulla *unfair competition* oltre che sul *copyright*.

⁵ Sulla linea fatta propria dagli attori nel caso per primo citato si colloca M. BARBONE, *Valorizzazione degli asset di proprietà intellettuale attraverso la rete*, in G. CASSANO e B.

TASSONE (a cura di), *Diritto industriale e d'autore nell'era digitale*, Milano, 2022, 66 ss., il quale fa notare come le condotte telegraficamente descritte nel testo possono comportare appunto l'elaborazione dell'opera (artt. 4 e 18 l.d.a.), la riproduzione della stessa (art. 13 l.d.a.), la sua traduzione e/o trascrizione in un linguaggio binario (art. 14 l.d.a.), la riproduzione della medesima anche se soltanto temporanea (art. 64-bis, lett. b), l.d.a.), il tutto in violazione, quando tali diritti non sono autorizzati, anche dei relativi diritti connessi (artt. 72 e ss. e 80 e ss. l.d.a.).

⁶ Nel senso che «per "intelligenza artificiale generativa" si intende il campo dell'intelligenza artificiale che si concentra sulla creazione di contenuti nuovi e originali rispetto ai dati di input in risposta alle richieste (prompt) dell'utente, attraverso l'utilizzo di algoritmi prevalentemente di tipo neurale» e per "rete neurale" si intende a sua volta «un modello computazionale standard applicabile nei contesti più diversificati che permette il riconoscimento di oggetti, forme o pattern all'interno di un dato o un insieme di dati (ad esempio, un volto umano in una fotografia)», sicché «[g]li algoritmi di intelligenza artificiale generativa sono impiegati in una vasta gamma di applicazioni, tra cui il riconoscimento e la generazione di immagini, di tracce vocali o musicali, di testi e di video», Garante per la protezione dei dati personali, provvedimento del 10 aprile 2025, n. 232, disponibile presso garanteprivacy.it.

la capacità di riprodurre comportamenti umani, sempre al di fuori degli *input* originari⁷.

E tale seminale elenco è stato poi completato dalla dottrina indicando quali altre caratteristiche dei “robot cognitivi” la capacità di ragionamento, notando che essi sono in grado di modificare la conoscenza di base attraverso la manipolazione logica delle informazioni già disponibili, nonché la capacità di pianificazione, dacché possono procedere al calcolo e alla selezione delle attività da compiere, con assunti che trovano conferma nelle fonti rilevanti in materia⁸.

In effetti, nella Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni del 24 aprile 2018 dal titolo “L’intelligenza artificiale per l’Europa” vengono ricondotti alla IA «i sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere specifici obiettivi», altresì specificandosi che essi possono o no essere incorporati in dispositivi *hardware*⁹.

Proprio per tale ragione nel presente lavoro si preferisce usare il termine onnicomprensivo IA

in luogo di “agenti *software*”, da un lato, oppure “macchine”, “robot” e simili dall’altro lato, ancorché con la precisazione per cui non si vogliono escludere dal discorso le applicazioni c.d. *embodied*, antropomorfe o no che siano: le quali meritano anzi e per certi versi uno specifico interesse (e suscitano i timori maggiori) perché in esse la sfera delle capacità intellettive della IA si unisce a quella delle capacità fisiche delle macchine e, grazie alla sensoristica, queste sviluppano connotati sempre più simili a quelli umani¹⁰.

Inoltre, per parte sua l’AI Act, con un testo varie volte modificato sul punto durante il suo *iter* legislativo, stabilisce al suo art. 3 – richiamato dalla stessa legge 132/2025 quanto alle definizioni da usare per perimetrare (anche) la sua sfera di applicazione – che per «sistema di IA» si intende «un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall’input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali»: mentre la nozione di

⁷ La definizione di cui si tratta – riportata in AA.VV., *Intelligenza artificiale e sicurezza: opportunità, rischi e raccomandazioni*, marzo 2021, 22, reperibile presso <https://iasecurity.clusit.it/> – è la seguente: “[l]’intelligenza artificiale (IA) è la capacità delle macchine o dei computer di eseguire compiti che vanno oltre la loro programmazione, come l’apprendimento e la soluzione di problemi. Il termine intelligenza artificiale è stato applicato in molti modi diversi, ma il suo uso più comune si riferisce a programmi per computer che hanno imparato comportamenti di tipo umano attraverso una vasta esperienza. Ciò include elementi come il riconoscimento vocale, il pattern matching, l’analisi delle immagini e molto altro ancora. Affinché una macchina diventi intelligente deve imparare dalle esperienze passate e avere la capacità di ragionare su nuove situazioni. Queste due abilità sono strettamente correlate e sono necessarie l’una all’altra. La macchina deve inoltre dimostrare capacità di adattamento in risposta a circostanze mutevoli. Ciò è particolarmente importante se il suo scopo è quello di interagire con l’uomo o con la società in generale: un robot che non possa cambiare il suo comportamento quando necessario non sarebbe

considerato intelligente da nessuna definizione del termine”.

⁸ A. SANTOSUOSSO e M. TOMMASI, *Diritto, scienza e nuove tecnologie*, 3ª ed., Milano, 2021, 324.

⁹ COM(2018) 237 final, ove pure si precisa – come esposto nel testo – che «i sistemi basati sull’IA possono consistere solo in *software* che agiscono nel mondo virtuale (per esempio assistenti vocali, *software* per l’analisi delle immagini, motori di ricerca, sistemi di riconoscimento vocale e facciale); oppure incorporare l’IA in dispositivi *hardware* (per esempio in robot avanzati, auto a guida autonoma, droni o applicazioni dell’Internet delle cose)».

¹⁰ Sul rilievo della componente fisica, anche in ragione dei progressi della sensoristica che consentono di costruire sistemi di visione e di percezione molto avanzati nonché sensori tattili e sistemi uditivi integrati in robot con capacità biomeccaniche sempre migliori, R. CINGOLANI e D. ANDRESCIANI, *Robot, macchine intelligenti e sistemi autonomi: analisi della situazione e delle prospettive*, in G. ALPA (a cura di), *Diritto e intelligenza artificiale*, Pisa, 2020, 25 ss. Per l’uso della locuzione “agenti autonomi” al fine di ricomprendere “droni, robot e agenti *software* puramente digitali”, SANTOSUOSSO e TOMMASI, *Diritto, scienza e nuove tecnologie*, cit., 321.

autoapprendimento è contenuta nel Considerando 12 e il Considerando 97 precisa che i modelli di IA per finalità generali «sono solitamente addestrati su grandi quantità di dati con diversi metodi, come l'apprendimento autosupervisionato, non supervisionato o per rinforzo»¹¹.

Orbene, anche se qui non ci si può soffermare sulle critiche rivolte a tale definizione e sul modo in cui essa richiama i fondamenti dell'IA, è chiaro che il citato tema dell'apprendimento è fondamentale¹²: venendo soprattutto in rilievo il c.d. *deep learning*, cioè un modello che emula quello della mente umana e tramite il quale l'IA impara a svolgere una determinata azione anche se la stessa non era stata programmata, come detto, tra quelle possibili¹³.

¹¹ Secondo il Considerando 12 «[I]e tecniche che consentono l'inferenza nella costruzione di un sistema di IA comprendono approcci di apprendimento automatico che imparano dai dati come conseguire determinati obiettivi e approcci basati sulla logica e sulla conoscenza che traggono inferenze dalla conoscenza codificata o dalla rappresentazione simbolica del compito da risolvere. La capacità inferenziale di un sistema di IA trascende l'elaborazione di base dei dati consentendo l'apprendimento, il ragionamento o la modellizzazione. [...] I sistemi di IA sono progettati per funzionare con livelli di autonomia variabili, il che significa che dispongono di un certo grado di autonomia di azione rispetto al coinvolgimento umano e di capacità di funzionare senza l'intervento umano. L'adattabilità che un sistema di IA potrebbe presentare dopo la diffusione si riferisce alle capacità di autoapprendimento, che consentono al sistema di cambiare durante l'uso. I sistemi di IA possono essere utilizzati come elementi indipendenti (stand-alone) o come componenti di un prodotto, a prescindere dal fatto che il sistema sia fisicamente incorporato nel prodotto (integrato) o assista la funzionalità del prodotto senza esservi incorporato (non integrato)». Per una panoramica sui documenti ufficiali che a livello unionale prendono in considerazione la IA, F. RODI, *Gli interventi dell'Unione Europea in materia di intelligenza artificiale e robotica: problemi e prospettive*, in ALPA, *Diritto e intelligenza artificiale*, cit., mentre per un recente commento all'AI ACT si vedano i vari contributi editi in G. CASSANO e E.M. TRIPODI (a cura di), *Il regolamento europeo sull'intelligenza artificiale - Commento al Reg. UE n. 1689/2024*, Santarcangelo di Romagna, 2024.

¹² Nel senso che gli algoritmi che si basano sulla necessaria formalizzazione matematica delle tre aree fondamentali dagli studiosi esplorate negli anni, ossia logica, computazione e probabilità, S. RUSSEL e P. NORVIG, *Intelligenza Artificiale - Un approccio Moderno*, I, Torino,

3. Opere generate dalla IA e diritto d'autore

L'evoluzione tecnologica è sempre andata di pari passo con il dibattito circa l'estensione della tutela offerta dal diritto d'autore, come dimostrano ad esempio le vicende inerenti alle opere fotografiche, alle banche dati, ai *software* e ai progetti tecnici¹⁴.

Limitandoci, per brevità e a titolo evidentemente esemplificativo, a considerare le sole opere fotografiche, già quando ancora non si parlava di IA era chiaro in dottrina quanto la loro trattazione richiedesse la considerazione delle dinamiche evolutive che investivano allora e che oggi investono ancor di più il diritto d'autore nel suo complesso¹⁵.

2005, 9 ss. Per il resto, la definizione contenuta nel Reg. (UE) 2024/1689 è ben diversa dalla definizione contenuta nella prima bozza di regolamento perché tra quest'ultima e la versione finale i sistemi di IA si sono evoluti al punto che il legislatore si è trovato a doverne aggiornare appunto la definizione. In effetti, nella Proposta di Regolamento del 2021 la definizione di sistema di IA fornita era di «un software sviluppato con una o più delle tecniche e degli approcci elencati nell'allegato I, che può, per una determinata serie di obiettivi definiti dall'uomo, generare output quali contenuti, previsioni, raccomandazioni o decisioni che influenzano gli ambienti con cui interagiscono»: laddove, come si è visto nel testo, ora la medesima norma definisce un sistema di IA come «un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali».

¹³ Così V. VIETRI, *FinTech: la rivoluzione nel mondo della finanza*, in ALPA, *Diritto e intelligenza artificiale*, cit., 511.

¹⁴ Sulle banche dati e sui progetti tecnici, rispettivamente, CGUE, sez. III, 18 ottobre 2012, n. 173/11, in *eur-lex.europa.eu*, nonché da ultimo T.a.r. Emilia-Romagna, Bologna, sez. I, 22 ottobre 2025, n. 1180, in *OneLegale*.

¹⁵ Sul tema B. TASSONE, *Dei diritti sulle fotografie - Commento agli artt. 87-92 della legge 22 aprile 1941, n. 633*, in A. Gambino e C. Galli (a cura di), *Codice ipertestuale commentato della proprietà industriale e intellettuale*, Torino, 2011, 3282 ss. e, più di recente, A. PEDRIALI-KINDLER *Ad artt. 87-92 L. 22 aprile 1941, n. 633*, in L. C. UBERTAZZI (a cura di), *Commentario breve alle leggi su proprietà intellettuale e concorrenza*, Padova, 2019, 2053 ss.

In effetti, considerando anche solo la “storia” della disciplina dettata in materia ci si rende agevolmente conto del fatto che – di là da ciò che *illo tempore* prevedeva il R.D.L. 7 novembre 1925, n. 1950 – le opere fotografiche hanno ricevuto espresso riconoscimento all’interno di quelle creative solo con il d.p.r. 8 gennaio 1979, n. 19, peraltro sulla scia delle modifiche apportate alla Convenzione d’Unione di Berna per la protezione delle opere letterarie e artistiche (cui lo Stato italiano ha dato ratifica ed esecuzione con la L. 20 giugno 1978, n. 399) nel testo di Parigi-Stoccolma¹⁶.

Da tale scelta normativa è infatti derivato il problema di distinguere le opere fotografiche di cui all’art. 2, le fotografie semplici di cui alla prima frase dell’art. 87 e le fotografie di mera documentazione di cui alla seconda frase della disposizione, le quali non sono in alcun modo tutelate dalla legge (fatte salve, ovviamente, le regole civilistiche e penalistiche che presiedono alla protezione dei beni mobili e del patrimonio).

¹⁶ In proposito, soltanto detto atto normativo ha infatti aggiunto all’art. 2 della legge 22 aprile 1941, n. 633 (di seguito anche “l.d.a.”) il n. 7, a tenore del quale sono oggetto di tutela «le opere fotografiche e quelle espresse con procedimento analogo a quello della fotografia sempre che non si tratti di semplice fotografia protetta ai sensi delle norme del capo V del titolo II»: sicché si tratta di opere che hanno ricevuto protezione, almeno in parte, in ragione di quella dinamica economico-sociale che obbliga il legislatore a una costante attività di adeguamento delle norme poste a tutela del diritto d’autore per via dell’evoluzione tecnologica, più di recente messa prepotentemente in evidenza dall’avvento di *internet* e, poi, dalla stessa IA.

¹⁷ Su tale profilo generale della tutela autoriale, G. CASSANO e B. TASSONE, *La disciplina del plagio evolutivo tra effettività della tutela del diritto di autore ed evoluzioni in tema di diritto morale: il caso del pupazzo o personaggio di fantasia*, in *Riv. dir. ind.*, 2021, 214 ss.

¹⁸ Non a caso una pronuncia ha posto già molti anni addietro in luce la necessità che il giudice si addentri «in un campo non propriamente suo, quello dell’estetica, in condizioni di carenza di oggettivi ed assoluti criteri di valutazione» (Pret. Saluzzo 13 ottobre 1993, in *Dir. aut.*, 1994, 484). E ciò per non considerare che, dato tassonomico a parte, le difficoltà sono accresciute da ragioni per così dire “strutturali”, poiché le fotografie migliori – che sono insieme «documento e creazione» (secondo la felice formula di S. DELL’ARTE, *Opere fotografiche e plagio trasversale - Tutela cautelare*, in *Dir. aut.*, 2007, 268) – vengono spesso (anche se non sempre)

E anche in tale contesto si sono subito manifestate tutte le difficoltà più in generale legate all’elaborazione di una nozione onnicomprensiva di «opere creative» e di parametri validi per tutte le tipologie di opere elencate dall’art. 2, oltre che di modelli decisionali idonei ad assicurarne la uniforme e costante applicazione invece che trasformarli in formule linguistiche *bonnes à tout faire*, secondo un profilo anche di recente messo in luce dalla dottrina¹⁷: secondo difficoltà qui accresciute dal fatto che la legge non si limita a distinguere le fotografie creative e no, prevedendo come detto un piano intermedio di tutela che risente comunque delle categorie impiegate per individuare le vere e proprie opere dell’ingegno¹⁸.

Premesso che il *caveat* così formulato deve indurre ad un’attenta analisi delle singole fattispecie concrete, le oscillazioni giurisprudenziali sul punto fanno subito sorgere un interrogativo¹⁹: se è vero che la creatività del

alla luce solo grazie al «mestiere», cioè all’impiego di strumenti e tecniche altamente professionali (come fra gli altri evidenzia P. AUTERI, *Commento al Decreto del Presidente della Repubblica 8 gennaio 1979, n. 19*, in *Nuove leggi civ. comm.*, 1980, 158 ss.).

¹⁹ In proposito, secondo una parte della giurisprudenza costituisce opera fotografica quella che «per l’originalità dell’inquadratura, l’impostazione dell’immagine e la capacità stessa di evocare suggestioni che trascendono il comune aspetto della realtà raffigurata, rappresenta una realizzazione artistica e non costituisce un mero fatto riproduttivo idoneo soltanto a documentare determinate azioni o situazioni reali» (Trib. Firenze 16 febbraio 1994, in *Dir. aut.*, 1994, 480); altra parte della giurisprudenza, invece, sembra adottare un canone meno “oggettivo”, affermando che la tutela autoriale «è operante tutte le volte che il fotografo abbia inserito nell’opera la propria fantasia, il proprio gusto e la propria sensibilità, così da trasmettere le proprie emozioni a chi esamini la fotografia in tal guisa realizzata» (Trib. Milano 6 marzo 2006, in *AIDA*, 2007, 792, nonché Trib. Milano 28.6.1993, *id.*, 1993, 757); e, al riguardo, si è altresì precisato che «la creatività necessaria ai fini della qualificazione di una fotografia come opera dell’ingegno non si identifica con l’originalità assoluta, né con la novità, ma richiede che l’opera riveli un’impronta personale dell’autore, e cioè sia frutto non solo della tecnica, ma dell’ingegno, della sensibilità e della perizia del medesimo» (Trib. Catania 24 novembre 2004, in *AIDA*, 2006, 411). Per altre pronunce, poi, occorre «l’apporto di una creatività particolare, connotata da un’impronta personale e da una valenza estetica che

fotografo si esprime fra l'altro nel momento della scelta fra le centinaia di "scatti possibili", come si può traslare un tale insegnamento alle immagini create dall'IA²⁰?

Ecco dunque che la citata novella della legge sul diritto d'autore è stata solo il punto di partenza della "nuova vita" delle opere fotografiche, anche perché già l'avvento dell'era digitale aveva consentito agli autori da un lato giovare di inedite opportunità di sfruttamento e diffusione delle opere dell'ingegno (che inducevano a preconizzare una profonda trasformazione del

diritto d'autore stesso e delle strutture che lo regolano, anche solo considerando la possibilità di instaurazione di rapporti diretti con il fruitore dell'opera), dall'altro incrementava esponenzialmente i rischi di contraffazione e stimolava purtroppo la nascita di una cultura "anti-copyright", rendendo necessario ridisegnare e potenziare l'apparato delle misure di tutela²¹.

Di poi, come è facile immaginare, si tratta di problemi esponenzialmente incrementati dall'avvento dell'IA, perché se prima un

possa essere apprezzata di per sé prescindendo dal soggetto rappresentato», sicché «la rappresentazione di un soggetto comune o il ritratto di un personaggio possono assurgere alla dignità di opere d'arte ove risultino compiute con una interpretazione del tutto personale, fuori dall'ordinario, finalizzata ad offrire un prodotto unico, frutto dell'ingegno del suo autore e purché dette immagini si differenzino notevolmente da quelle che altri avrebbero potuto realizzare in situazioni analoghe» (App. Milano 10 ottobre 2003, in *Foro pad.*, 2005, I, 129 e, analogamente, Trib. Roma 28.3.2003, in *Dir. aut.*, 2004, 101). In proposito, si è anche affermato che «non può riconoscersi l'integrale applicazione della tutela del diritto d'autore prevista dall'art. 2 n. 7 l.d.a., bensì la più limitata protezione prevista per le fotografie "semplici" dagli art. 87 ss. della stessa legge, nel caso di fotografie relative ad una competizione automobilistica che, pur rivelando elevata e sicura professionalità nella cura dell'inquadratura e nella capacità di cogliere in modo efficace gli eventi ripresi, non evidenziano l'espressione di un'originale interpretazione personale dell'autore, tale da imprimere alle immagini un peculiare carattere di creatività in ragione di uno specifico contributo estetico e figurativo riconducibile alla personalità dell'autore medesimo» (Trib. Milano, 9 novembre 2000, in *AIDA*, 2002, 578). Viceversa, al cospetto di fotografie delle celebrazioni in onore di una santa patrona si è sostenuto che «la prevalenza del profilo artistico e creativo sull'aspetto prettamente tecnico, da cui discende la tutelabilità dell'opera fotografica come opera dell'ingegno, emerge quando l'attimo fotografato coglie il momento significativo di un evento, ricorrendo ad un linguaggio connotativo che crea una composizione di prospetti, luci, ombre e colori del tutto peculiari» (Trib. Catania 11 settembre 2001, in *Foro it.*, 2002, I, 1236). Infine, appare degna di nota la pronuncia secondo cui possiede i requisiti di creatività necessari ad essere tutelata come opera dell'ingegno la «fotografia che riproduce un'opera architettonica e sia caratterizzata da giochi di luci ed ombre, taglio particolare dell'immagine e fuga prospettica della facciata dell'edificio lungo il corso che lo costeggia» (Pret. Torino 27 maggio 1996, in *AIDA*

1997, 751), alla quale si può accostare quella della Suprema Corte secondo cui va considerata opera dell'ingegno «la fotografia che non si limita alla mera riproduzione di un dipinto, in quanto ne costituisce una rielaborazione» (Cass. 12 marzo 2004, n. 5089, in *Foro it.*, 2004, I, 2441). Il rilievo delle pronunce testé citate è dato dal fatto che solitamente si tende ad escludere il carattere creativo delle fotografie relative ad opere architettoniche, in considerazione della fedeltà della rappresentazione del soggetto riprodotto che usualmente le caratterizza (Pret. Saluzzo 13 ottobre 1993, in *Dir. aut.*, 1994, 484), al pari di quanto si è deciso per le fotografie di opere d'arte sul presupposto che la "creatività" vada in principio attribuita all'oggetto riprodotto (Trib. Bari 20 febbraio 1998, in *Foro it.*, Rep. 2000, voce *Diritti d'autore*, n. 103, nonché Trib. Roma 24 febbraio 1998, in *Dir. inf.*, 1998, 793), mentre maggior margine – sempre in linea di principio – può sussistere per il ritratto fotografico (come emerge da Trib. Milano 28.6.1993, in *AIDA*, 1993, 757). In ultimo, non sembra che i diversi approcci seguiti dalla giurisprudenza testé citata siano poi mutati, come emerge per tutti da A. PEDRIALI-KINDLER *Ad artt. 87-92 L. 22 aprile 1941, n. 633*, in UBERTAZZI, *Commentario breve alle leggi su proprietà intellettuale e concorrenza*, cit., 2053 ss.

²⁰ Così App. Milano 7 novembre 2000, in *AIDA*, 2001, 789, cui *adde* Trib. Milano 9 ottobre 2000, *id.*, 2002, 831, nonché Cass. 4 luglio 1992, n. 8186, in *Foro it.*, 1993, I, 127, le quali riconoscono il valore di fotografie «frutto di sicura professionalità, in quanto vi si riscontra un'accurata ricerca e composizione degli oggetti fotografati, un'equilibrata relativa reciproca collocazione, uno studiato e non casuale uso della luce», ma negano alle stesse tutela autoriale perché non ritrasmettono «alcuna reinterpretazione soggettiva della realtà e non evocando alcuna particolare suggestione».

²¹ Su questi aspetti già B. TASSONE, *DRM e rifiuto di licenza nel caso Virgin v. Apple: questione di FairPlay?*, in *AIDA*, 2005, 376 ss., nonché gli scritti editi in R. CASO (a cura di), *Digital Rights Management: problemi teorici e prospettive applicative. Atti del Convegno tenuto presso la Facoltà di Giurisprudenza di Trento il 21 ed il 22 marzo 2007*, Università degli Studi di Trento, Trento, 2008.

fotomontaggio richiedeva la sovrapposizione – fisica o almeno virtuale – di immagini, ora anche una GenAI disponibile grazie ad un abbonamento gratuito è in grado di inserire accanto a due amici seduti ad un tavolo da pranzo un terzo con un bicchiere in mano, anche se lo stesso non è mai stato lì, a seguito di un semplice prompt: che, in generale, basta a creare un'immagine fotografica – di nuovo senza alcun uso dei consueti “ferri del mestiere”, analogici o digitali che siano – secondo i meccanismi (e con i problemi affrontati dalla casistica straniera) di cui in appresso, sui quali le corti supreme o unionali non hanno ancora potuto pronunciarsi.

In proposito, la CGUE ha infatti avuto solo occasione di ribadire che un ritratto fotografico può essere protetto dal diritto d'autore alla condizione, che spetta al Giudice nazionale di verificare in ogni singolo caso, che esso costituisca una creazione intellettuale dell'autore la quale ne rifletta la personalità e si manifesti attraverso scelte libere e creative di quest'ultimo nella realizzazione di tale ritratto²²: secondo insegnamento che, peraltro, non risolve ancora il citato problema degli standard²³.

Tuttavia, proprio in Italia si è avuta una primissima pronuncia della Cassazione in materia di software e tutela del diritto d'autore per un'opera generata dalla IA, anche se solo in *obiter*.

Nella specie, tutto ha avuto inizio con la richiesta della creatrice dell'opera grafica “The Scent of the Night” dell'accertamento della violazione dei propri diritti d'autore, nonché della condanna della convenuta (RAI) al risarcimento dei danni per l'utilizzo, non autorizzato, della stessa come scenografia fissa del Festival di Sanremo 2016.

E il Giudice di prime cure ha accolto tale richiesta con pronuncia che ha poi superato indenne il vaglio del Giudice di appello²⁴: mentre la Suprema Corte, pur pronunciandosi (come detto) solo incidentalmente sul tema, con la pronuncia n. 1107/2023 ha comunque dato semaforo verde alla possibilità di attribuire carattere creativo ad una fotografia realizzata dal suo autore tramite un *software* di IA²⁵.

In specie, secondo gli Ermellini l'utilizzo di un *software* «per generare l'immagine [...] è pur sempre compatibile con l'elaborazione di un'opera dell'ingegno con un tasso di creatività che andrebbe solo scrutinato con maggior rigore»: secondo una scelta di *policy* poi smentita dalla successiva Legge n. 132/2025, come meglio si vedrà in appresso dopo aver allargato il discorso ad altri profili del “diritto delle immagini”.

4. IA e informazione visiva

Secondo parte della dottrina la GenAI di immagini rappresenta una frontiera verso cui muoversi stando attenti a ben governare i temi

²² CGUE, sez. III, 1° dicembre 2011, n. 145, in *eur-lex.europa.eu*.

²³ In effetti, la definizione degli standard appare oggetto di continue discussioni in giurisprudenza: se secondo una sentenza della Suprema Corte «il concetto di creatività *ex art.* 1 l. 633/1941 non coincide con quello di creazione, originalità e novità assoluta, riferendosi, per converso, alla personale e individuale espressione di un'oggettività appartenente alle categorie elencate, di modo che un'opera dell'ingegno riceva protezione a condizione che sia riscontrabile in essa un atto creativo, seppur minimo, suscettibile di manifestazione nel mondo esteriore» (Cass. 12 marzo 2004, n. 5089, in *Foro it.*, 2004, I, 2441), una quasi coeva decisione di merito vuole che «la protezione assicurata dall'art. 2 n. 7 della l. 633/1941 riguarda esclusivamente le fotografie artistiche, dotate di un livello di creatività particolarmente elevato» (Trib. Roma 22 settembre 2004, in *Dir. aut.*, 2005, 378 ss.), mentre poco dopo si è tornati a richiedere un «minimo gradiente di

creatività» (App. Trento 14 dicembre 2007, massimata in *Dir. aut.*, 2008, 535).

²⁴ Trib. Genova, 6 giugno 2018, n. 1640, e App. Genova, sez. spec. in materia di imprese, 11 novembre 2020, n. 1066, entrambe in *OneLegale*.

²⁵ Cass. 16 gennaio 2023, n. 1107, si legge in *Foro it.*, 2023, I, 1874, con nota di richiami, in *Dir. internet*, 2023, 3, 481 ss., con nota di A. LA ROSA, *La tutelabilità dell'opera creata col supporto di sistemi di intelligenza artificiale*, in *Arte e Diritto*, 2023, 2, II, 391 ss., con nota di A. BETTONI, *Tutela dell'opera d'arte algoritmica. Spunti comparatistici a margine di una recente pronuncia italiana*, nonché *ivi*, 417 ss., con nota di P. LEANZA, *Il diritto d'autore tra ingegno umano e intelligenza artificiale*, oltre che in *Riv. dir. ind.*, 2023, 252, con nota di V. IAIA, *Il «maggior rigore» nell'esame di originalità delle opere dell'ingegno realizzate con l'ausilio del software*, e ancora in *Giur. it.*, 3, 2024, 571, con nota di A. AMIDEI, *Diritto d'autore - La tutela autoriale dell'opera generata dall'A.I. (o da un semplice software?)*.

controversi che essa solleva ma, al tempo stesso, profittando delle numerose sue applicazioni utili per singoli e collettività²⁶.

In proposito, un'applicazione della GenAI di immagini certamente utile per la collettività può rinvenirsi nella visione artificiale applicata alla ricostruzione di artefatti visivi (quali antichi affreschi, superfici pittoriche, dipinti frammentari o deteriorati) per riconoscere pattern cromatici, texture, tracce materiali, porzioni mancanti e stratificazioni delle singole opere attenzionate.

Le tecniche di computer vision permettono infatti di ricostruire parti perdute, suggerire integrazioni plausibili, identificare stesure pittoriche successive, classificare pigmenti o individuare interventi di restauro pregressi. L'obiettivo diventa, in altre parole, la ricostruzione dell'immagine, la comprensione della sua struttura e la restituzione quanto più fedele possibile dell'aspetto originario dell'artefatto.

In proposito, è chiaro che l'IA è qui chiamata ad operare esclusivamente sull'informazione visiva, tramite un lavoro centrato sull'elaborazione dell'immagine, che richiede competenze di tipo tecnico-informatico (computer vision, modelli generativi, segmentazione visiva)²⁷.

Come è evidente, siamo qui già ai confini del diritto d'autore, che – da un lato e come è noto – tutela anche opere munite di creatività assai bassa che non sono vere e proprie opere d'arte o comunque beni culturali (che a loro volta non

sono sempre opere dell'ingegno), ma dall'altro è distinto dal diritto dei beni culturali stessi²⁸.

Ebbene, tralasciando i pure noti problemi che sorgono quando un bene culturale che sia anche un'opera dell'ingegno è caduto in pubblico dominio – quanto esclude la tutela autoriale, ma non quella di cui agli artt. 107 e 108 del Codice dei Beni Culturali e del Paesaggio (D.Lgs. 42/2004), il che ha consentito di configurare la nascita di un "diritto all'immagine" del bene culturale di natura dominicale (il quale permette allo Stato di bloccare gli usi commerciali non autorizzati o indecorosi dell'immagine dei capolavori nostrani, anche non coperti più dal copyright) –, gli usi della IA sopra cennati dimostrano che essa può essere talvolta persino funzionale alla tutela del diritto d'autore: il che ci porta ora a indagare meglio il ruolo che essa gioca nella produzione delle opere dell'ingegno.

5. L'autore artificiale: fu vero autore?

Come emerge già dai seminali accenni fatti in apertura, le numerose questioni che oggi gli operatori del diritto sono chiamati ad affrontare con riguardo al rapporto tra IA generativa e diritto d'autore impongono la ricerca di soluzioni di ampio respiro e tali da collocarsi ben oltre il solco tracciato dalle esperienze normative nazionali per come modellate, nel corso degli anni, dall'apporto della giurisprudenza: con la precisazione per cui la GenAI implica anche il ripensamento delle coordinate che descrivono i soggetti e l'oggetto del diritto.

²⁶ In specie, per Diana-Urania Galetta, *Intelligenza artificiale e pubblica amministrazione: aspetti critici* (seminario dell'Università degli Studi di Milano, 12 novembre 2025), non è seriamente contestabile che «sapere e comprendere ... è fondamentale, prima di decidere se e come servirsi dello strumento (...) sebbene di recente siano tutti intenti a parlare di Intelligenza Artificiale - dividendosi più o meno equamente fra coloro che si avvicinano all'IA con utopico ottimismo (formulando affascinanti ipotesi sulle sue innumerevoli e mirabolanti applicazioni possibili alla vita quotidiana, al lavoro, ma anche all'attività dello Stato e della Pubblica Amministrazione nelle sue molteplici attività) e coloro che, invece, vi si avvicinano con maggiore paura e diffidenza - pochi sanno di che cosa stanno davvero parlando. Anche gli studiosi di area umanistica che si avvicinano al tema IA - e i giuristi rientrano ovviamente

all'interno di questa categoria - si trovano sovente in una situazione di vero e proprio "deficit cognitivo": nel senso che discutono di un oggetto per loro largamente sconosciuto, una sorta di UFO, potremmo dire».

²⁷ T.a.r. Campania, Napoli, sez. V, 5 marzo 2026, n. 1525, in *OneLegale*.

²⁸ In proposito, è pure noto che i due plessi disciplinari sono regolati dal sistema del c.d. "doppio binario", tale per cui vi sono due tutele parallele, indipendenti e non sovrapponibili, che proteggono due profili completamente diversi dello stesso oggetto (nella misura in cui l'opera dell'ingegno sia anche bene culturale): mentre il diritto d'autore protegge l'opera dell'ingegno (l'atto creativo immateriale), il diritto dei beni culturali tutela tendenzialmente il supporto fisico (la cosa materiale) in quanto parte della memoria e dell'identità storica della Nazione.

Quanto al profilo soggettivo della tutela, l'interrogativo che pone la GenAI attiene alla possibilità, o no, di configurare una sorta di autore "artificiale".

In proposito, va subito notato che il dato normativo di riferimento – nazionale, euro-unionale e di matrice internazionale – non è dirimente al fine di risolvere l'interrogativo in quanto, a volte, fa riferimento all'autore quale persona fisica, altre volte ammette che l'autore sia una "persona giuridica", senza mai spingersi fino ad accordare tutela anche all'autore artificiale, così come a negarla espressamente (perlomeno non in ogni esperienza giuridica e fatto salvo quanto si dirà circa la legge 132/2025).

È però vero, al tempo stesso, che diversi fattori lasciano propendere per la limitazione della qualità di autore alla sola persona fisica (o entro certi limiti giuridici) così rendendo più arduo configurare l'esistenza di un autore artificiale²⁹. Ad esempio, è sicuramente con riferimento alla persona umana che il Legislatore accorda la tutela anche del diritto morale.

Inoltre, la stessa creatività non è solo definita tramite uno standard assai basso, ma almeno fino ad oggi essa è stata riferita dalla giurisprudenza alla «personale e individuale espressione», al «patrimonio intellettuale» e/o

alla «soggettività» di un autore persona fisica³⁰: senza contare che sebbene la legge n. 633/1941 non definisca espressamente e in tal senso la nozione di autore, il dato si ricava da una serie di altri indici³¹.

D'altronde, la stessa ricostruzione del diritto d'autore come diritto duale – in contrapposizione al modello continentale tedesco – e la riconduzione dei diritti morali a quelli della personalità di rilevanza costituzionale, dovrebbero supportare tale interpretazione³²: mentre di recente, proprio valorizzando i profili indicati, la Cassazione ha sancito il diritto di ogni autore di vedersi sempre riconosciuta la paternità dell'opera, superando l'orientamento che escludeva una violazione di tale diritto ove l'utilizzatore legittimo (cioè il titolare dei diritti patrimoniali) non avesse attribuito a sé o ad altri la paternità dell'opera stessa, ma avesse solo ommesso il nome dell'autore, imponendogli un «anonimato forzoso»³³.

Di poi, pure la disciplina euro-unitaria – alla luce della quale dev'essere interpretata, come è noto, quella nazionale – è sempre più orientata a rafforzare la tutela morale degli autori, come emerge dall'art. 17, comma 2, della Carta dei Diritti Fondamentali dell'Unione Europea³⁴. Inoltre, in varie occasioni la Corte di Giustizia UE

²⁹ Fra i tanti *laia*, *To Be, or Not to Be ... Original Under Copyright Law, That Is (One of) the Main Questions Concerning AI-Produced Works*, in «GRUR International», 71, 9, 2022, 793 ss.

³⁰ BARBONE, *Valorizzazione degli asset di proprietà intellettuale*, cit., 57 ss. Nel senso che la creatività tutelabile dal diritto d'autore può essere solo quella proveniente dall'uomo, V.M. DE SANCTIS, *La protezione delle opere dell'ingegno*, t. 1, *Le opere figurative, le opere audiovisive, le opere utilitarie*, Milano 2004, 74, secondo il quale «[l]a qualità di autore, nel senso dell'acquisto originario del diritto, spetta sicuramente a qualsiasi persona fisica, anche minore, inabilitato, interdetto o anche inconsapevole, che crei un'opera dell'ingegno».

³¹ BARBONE, *Valorizzazione degli asset di proprietà intellettuale*, cit., 58 ss., ove si rammenta che l'art. 8 l.d.a. considera autore dell'opera chi è in essa indicato come tale, che l'art. 25 l.d.a., nel definire la durata dei diritti, fa riferimento alla vita dell'autore e al termine di settanta anni dopo la sua morte e che l'art. 115 l.d.a. contempla espressamente la trasmissione dei diritti agli eredi dell'autore dopo la morte di quest'ultimo.

³² Sulle teorie dualistiche e sulla teoria unitaria nella dottrina italiana, di nuovo DE SANCTIS, *La protezione delle opere dell'ingegno*, cit., 75 ss. Per alcuni cenni alla storia del diritto d'autore che spiegano le differenze ricostruttive che hanno portato al modello dualista e monista in Europa, nonché a quello 'patrimoniale' ancora vigente nei sistemi di *Common Law*, B. TASSONE e M. BARBONE, *Diritti sul cibo e tutela dei piatti creativi. Una prospettiva comparata*, in L. SCAFFARDI e V. ZENO-ZENCOVICH (a cura di), *Cibo e diritto – Una prospettiva comparata. Atti del Colloquio Biennale della Associazione Italiana di Diritto Comparato sul tema, Parma, 23-25 maggio 2019*, vol. II, 2020, 587 ss.

³³ Cass. 5 luglio 2019, n. 18220, in *Foro it.*, I, 2019, 3957.

³⁴ Così delineati gli argomenti che dovrebbero legare a doppio filo la nascita di un'opera dell'ingegno ad un atto creativo umano, è da notare che con riferimento al suo movente giustificativo molto si è discusso in passato sul se esso dovesse essere il frutto di un ragionato sforzo psichico o di un lavoro mediato teso ad un risultato nuovo ovvero sul se bastasse l'atto istintivo carente di volontà proprio in ragione della distinzione che in letteratura, già

ha stabilito che «perché un oggetto possa essere considerato originale, è necessario e sufficiente che rifletta la personalità del suo autore, manifestando le scelte libere e creative di quest'ultimo»³⁵.

Del resto, una specificazione nel senso indicato si rinviene sia nella Direttiva 96/9/CE del Parlamento Europeo e del Consiglio, dell'11 marzo 1996, relativa alla tutela giuridica delle banche di dati (c.d. "Direttiva Database") per la quale «l'autore di una banca di dati è la persona fisica o il gruppo di persone fisiche che l'ha creata o, qualora la legislazione dello Stato membro interessato lo consenta, la persona giuridica individuata da tale legislazione come titolare del diritto» (art. 4, I) sia nella Direttiva 2009/24/CE del Parlamento Europeo e del Consiglio del 23 aprile 2009 relativa alla tutela giuridica dei programmi per elaboratore (c.d. "Direttiva Software") secondo cui «la tutela è riconosciuta a tutte le persone fisiche o giuridiche aventi i requisiti previsti dalla legislazione nazionale sul diritto d'autore applicata alle opere letterarie» (art. 3).

Di poi, sempre a sostegno di una risposta negativa al quesito – passando agli Stati Uniti d'America –, si colloca ancora un caso molto discusso dalla comunità scientifica nazionale e internazionale relativo allo scatto fotografico posto in essere dal macaco *Naruto*, il quale ha portato una corte nordamericana a chiarire che

il *copyright* presuppone che «to qualify as a work of 'authorship' a work must be created by a human being» e che «[t]he [U.S. Copyright] Office will not register works produced by nature, animals, or plants»³⁶.

D'altronde, anche di recente le linee guida (del 16 marzo 2023) dell'U.S. Copyright Office hanno chiarito come il copyright è uno strumento di tutela delle sole opere create dagli uomini, osservando che «public guidance is needed on the registration of works containing AI-generated content. This statement of policy describes how the Office applies copyright law's human authorship requirement to applications to register such works and provides guidance to applicants»³⁷.

Infine, anche allargando lo sguardo ad altri sistemi, il *mainstream* pare quello indicato.

Così, l'art. 11 della legge cinese sul diritto d'autore è chiaro nell'indicare come autore di un'opera intellettuale la sola persona fisica suo "creatore", così mettendo fuori gioco ogni interpretazione atta a ricondurre in tale concetto l'autore artificiale: «[t]he author is the natural person who creates the work. For works hosted by a corporate or unincorporated organization, created on behalf of the will of the corporate or unincorporated organization, and for which the corporate or unincorporated organization assumes responsibility, the corporate or

alla fine degli anni sessanta del secolo scorso, veniva delineata per distinguere tra risultati creativi dell'uomo e risultati tecnici di cervelli elettronici: sicché già allora la dottrina si era divisa tra chi aveva ritenuto che la tutelabilità potesse essere estesa o no a quelle opere realizzate da un elaboratore elettronico, purché dietro l'azione della macchina vi fosse sempre l'azione dell'uomo, come emerge da R.S. SVELSON, *La musica elettronica e il diritto di autore nel pensiero di alcuni giuristi. - Musica elettronica e legge sul diritto di autore*, in *Dir. aut.*, 1966, 18 ss.

³⁵ Corte Giust. UE, 12 settembre 2019, causa C-683/17, nonché Corte Giust. UE, 16 luglio 2009, causa C-5/08, entrambe in www.curia.europa.eu.

³⁶ *Naruto v Slater (Monkey Selfie)*, no. 16.15469 (9th Circ. 2018), deciso dalla United States District Court Northern District of California. Nel caso di specie, è stato rilevato come il fatto che la macchina fotografica fosse stata impostata dal fotografo per la combinazione delle scelte tecniche necessarie per procedere allo scatto,

attribuirebbe a questi la paternità, sicché l'opera da essa risultante non sarebbe potuta essere di colui che materialmente ha proceduto allo scatto stesso, al di là del fatto che nel caso di specie si è trattato di un animale la cui appartenenza al genere di riferimento esclude *ab origine* la possibilità che egli sia riconosciuto autore dell'opera. Inoltre, il medesimo principio del caso *Naruto* lo si trova espresso nel risalente caso *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53 (1884). Sul tema, nella dottrina italiana, G. Rossi, *L'intelligenza artificiale e la definizione di "opera dell'ingegno"*, in *AIDA*, 2018, 268 ss.

³⁷ Ciò con la precisazione che «this policy statement sets out the Office's approach to registration of works containing material generated by AI technology. The Office continues to monitor new factual and legal developments involving AI and copyright and may issue additional guidance in the future related to registration or the other copyright issues implicated by this technology».

unincorporated organization shall be regarded as the author»³⁸.

Quanto al sistema messicano, in applicazione del Federal Law of Author's Right (Ley Federal del Derecho de Autor - LFDA) la Suprema Corte di Giustizia della Nazione (SCJN) del Messico, Segunda Sala, ha sottolineato la dimensione antropocentrica del diritto d'autore, escludendo ogni possibile riconoscimento di status autoriale ai sistemi di intelligenza artificiale³⁹.

E qui non può mancare un cenno all'art. 15 della già citata Convenzione di Berna, il quale prevede espressamente che «affinché gli autori di opere letterarie ed artistiche protette dalla presente Convenzione siano fino a prova contraria ritenuti tali, ed ammessi in conseguenza ad agire contro i contraffattori davanti ai tribunali dei Paesi dell'Unione, è sufficiente che il nome dell'autore sia indicato sull'opera nei modi d'uso» e che «[i]l presente alinea è applicabile anche se il nome sia uno pseudonimo, purché questo non lasci dubbi sull'identità dell'autore» (I comma): la disposizione, come è evidente, dà per presupposto - risalendo la sua firma al 9 settembre del 1886 - che l'autore sia una persona fisica.

Infine, nemmeno può omettersi un richiamo all'art. 27 della Dichiarazione Universale dei Diritti Umani secondo cui «ogni individuo ha diritto di prendere parte liberamente alla vita culturale della comunità, di godere delle arti e di partecipare al progresso scientifico ed ai suoi

benefici» (I comma) e secondo cui, ancora, «ogni individuo ha diritto alla protezione degli interessi morali e materiali derivanti da ogni produzione scientifica, letteraria e artistica di cui egli sia autore».

In ultimo, ancorché il diritto industriale rimanga fuori dal tracciato del presente lavoro, sulla linea indicata si collocano altresì le decisioni dell'EPO (*European Patent Office*), che hanno rifiutato le domande di brevetto europeo in cui un sistema di IA è stato designato come inventore⁴⁰: mentre l'ordinamento inglese (e derivati), per le ragioni che verranno di seguito esposte, merita un discorso a parte, da svolgere in una con l'illustrazione di differenti punti di vista.

6. (segue:) Le opinioni favorevoli

Con riguardo alle opinioni che si mostrano invece meno refrattarie al riconoscimento di una qualche soggettività a talune espressioni dell'IA, va intanto notato che non coglierebbe nel segno l'obiezione a tenore della quale il tema si esaurisce in una prospettiva *de iure condendo*⁴¹. Invero, nella Risoluzione del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica, il Parlamento Europeo ha invitato la stessa Commissione a valutare «l'istituzione di uno *status* giuridico specifico per i robot nel lungo termine, di modo che almeno i robot autonomi più sofisticati possano essere considerati come persone elettroniche responsabili di risarcire

³⁸ Per un'applicazione giurisprudenziale di tale normativa si veda la pronuncia resa dal Tribunale di Internet di Pechino (Beijing Internet Court o BIC) 27 novembre 2023, n. 11279, che può essere letta in lingua inglese presso <https://english.bjinternetcourt.gov.cn/pdf/BeijingInternetCourtCivilJudgment112792023.pdf>

³⁹ Suprema Corte di Giustizia della Nazione, 2 luglio 2025, Amparo Directo 6/2025, reperibile presso transparencia-ciudadana.scjn.gob.mx.

⁴⁰ Ci si riferisce alle domande EP 18 275 163 ed EP 18 275 174, le quali sono state respinte dall'EPO a seguito di un procedimento orale con il richiedente nel novembre 2019, in quanto non si è ritenuto che esse soddisfacessero i requisiti stabiliti dalla Convenzione sul brevetto europeo (EPC) per cui l'inventore designato nella domanda deve essere un essere umano e non una macchina, di cui riferisce M. CAPPARELLI, *Le invenzioni della Intelligenza Artificiale: questioni aperte di tutela autoriale e*

brevettabilità, in *Intelligenza artificiale*, in U. RUFFOLO (a cura di), *Intelligenza artificiale – Il diritto, i diritti, l'etica*, Milano, 2020, 345 ss. Sul rapporto fra IA e diritto industriale si vedano anche F. BANTERLE, *Ownership and Inventions Created by Artificial Intelligence*, in *AIDA*, 2018, 68 ss., nonché G. SANSEVERINO, *The Ability to Chase Down Our Dreams. Inventive Steps and Artificial Intelligence*, 12 dicembre 2018, disponibile presso www.ssrn.com, e ID., *Ex Machina. La novità e l'originalità dell'invenzione "prodotta" dall'IA*, in *AIDA*, 2018, 3 ss.

⁴¹ Per vari riferimenti fin da ora M. TAMPIERI, *L'intelligenza artificiale e le sue evoluzioni*, Padova 2022, 93 ss. Al riguardo, è appena il caso di notare che nella letteratura di *Common Law* il tema è trattato da più tempo, come emerge ad esempio da L.B. SOLUM, *Legal Personhood for Artificial Intelligences*, in *North Carolina Law Review*, 70, 1992, 1231 ss.

qualsiasi danno da loro causato, nonché eventualmente il riconoscimento della personalità elettronica dei robot che prendono decisioni autonome o che interagiscono in modo indipendente con terzi»⁴²: anche se, per vero, un successivo documento del Comitato Economico e Sociale Europeo e un'ulteriore Risoluzione del Parlamento Europeo sono orientati in senso diverso⁴³.

Di poi, uno sguardo più attento dimostra che accanto al dibattito sviluppatosi con riferimento al diritto d'autore vi sono vari contributi che affrontano il tema rispetto ad alcune specifiche aree del sistema, come quelle del diritto societario, dei contratti e della responsabilità civile⁴⁴.

Poiché non se ne può dare conto in questa sede e rimanendo quindi sul *copyright*, è intanto da notare che si tratta dell'ambito elettivo dal quale vengono tratti gli esempi di programmi di IA dotati di maggior autonomia all'interno di una fenomenologia assai variegata.

In proposito, è ad esempio noto che nel 2018 Christie's ha messo in vendita uno degli undici ritratti rientranti nell'opera "Edmond de Belamy", non distinguibile da un quadro dipinto da un essere umano e di valore stimato fra 7.000 e 10.000 dollari americani, ma poi venduto per 400.000 dollari: il quale è stato realizzato da un algoritmo programmato dal "collettivo" Obvious

dopo aver elaborato oltre 15.000 immagini risalenti ai secoli dal XIV al XVI Secolo⁴⁵.

Passando al settore della musica si può menzionare l'opera di Nao Tokui "Composing Melodies Inspired by Musical Scores" realizzata da Qosmo – uno dei più creativi gruppi di sperimentazione artistica audiovisuale – in collaborazione con Yamaha, la quale consente la creazione di spartiti effettivamente suonabili⁴⁶. Similmente, la startup "Jukedek" (nel 2019 acquisita da TikTok, dopo di che poco si sa di essa) ha sviluppato una rete neurale in grado di creare musica usando l'IA e persino di interpretare i video e impostarne da sola la colonna sonora⁴⁷.

Sempre in tale contesto una dottrina di casa nostra segnala le applicazioni che, dopo aver processato tutti i brani di Chopin, ne vengono a creare uno caratterizzato dalla medesima tecnica compositiva e che, facendo la stessa cosa con il repertorio dei Beatles, sono in grado di proporre una nuova canzone con musica e testo⁴⁸.

Quanto al cinema, la rete neurale 'Benjamin' ha scritto la sceneggiatura di un cortometraggio chiamato "Sunspring" presentato al festival di fantascienza di Londra nel 2016, il cui "nome" è

⁴² 2015/2103 (INL), Punto 59, Lettera f).

⁴³ Ci si riferisce al parere del Comitato Economico e Sociale Europeo dal titolo «L'intelligenza artificiale – Le ricadute dell'intelligenza artificiale su mercato unico (digitale), sulla produzione, sul consumo, sull'occupazione e sulla società», approvato all'esito della sessione plenaria del 31 maggio e del 1° giugno 2017, Punto 3.3, 2017/C 288/01, nonché alla Risoluzione del Parlamento Europeo recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale, del 5 ottobre 2020, punto 7, 2020/2014(INL).

⁴⁴ Sugli argomenti ritraibili delle altre aree del diritto di cui nel testo, TASSONE, *Riflessioni su intelligenza artificiale e soggettività giuridica*, cit., 214 ss.

⁴⁵ Si veda l'articolo di V. POLI, *Christie's mette all'asta Ritratto di Edmond de Belamy, opera dipinta da un'Intelligenza Artificiale*, 30 agosto 2018, disponibile presso www.tribune.com. Dà notizia del prezzo di vendita M. CAPPARELLI, *Le nuove frontiere del diritto d'autore alla prova dell'intelligenza artificiale*, in RUFFOLO, *Intelligenza artificiale*, cit., 336 ss. Altra peculiarità è

quella per cui l'IA ha anche creato un albero genealogico della immaginaria famiglia "Belamy" che sul piano verticale collega il quadro di ogni personaggio ad un diverso periodo della storia dell'arte e, sul piano orizzontale, ad una diramazione della famiglia che rappresenta «un futuro possibile» della medesima elaborata sempre dall'algoritmo. Per una più dettagliata descrizione si leggano gli stessi soggetti che compongono il citato collettivo e che si firmano con il relativo pseudonimo, cioè Obvious, *La famiglia Belamy e i sogni di Ukiyo*, in A. BARALE (a cura di), *Arte e intelligenza artificiale: Be My GAN*, Milano 2020, 167 ss.

⁴⁶ V. MINAZZI, *Gan e musica, on the road*, in BARALE, *Arte e intelligenza artificiale*, cit., 195 ss.

⁴⁷ Si veda l'articolo *TikTok Acquisisce Jukedek: Crea Musica con l'Intelligenza Artificiale*, 31 luglio 2019, disponibile presso www.musicpromoter.it.

⁴⁸ Ne riferisce M. FRANZONI, *Lesione dei diritti della persona, tutela della privacy e intelligenza artificiale*, in U. RUFFOLO (a cura di), *XXVI Lezioni di diritto dell'intelligenza artificiale*, Torino, 2021, 349-351.

stato peraltro indicato accanto agli altri autori del film⁴⁹.

Ancora, durante l'IDIA Fest tenutosi a Roma in data 18 novembre 2023 in tema di IA i vari contributori invitati a occuparsi di profili determinati hanno proceduto a interventi liberi tutti registrati da un programma di IA con il compito di redigere un documento congiunto che non costituiva la mera trascrizione, bensì una incisiva rielaborazione, del *brainstorming* collettivo lì avvenuto (con evidenti ricadute sul modo in cui, in futuro, verranno stesi gli atti dei convegni).

Infine, più di recente si è messo in luce che un'applicazione di IA ha scritto in Giappone "The Day a Computer Writes a Novel", il quale ha superato il primo round dello Hoshi Shinichi Literary Award del 2016, che lo Iamus Computer ha realizzato la partitura "Hello World" per piano, clarinetto e violino già nel 2011 all'esito dell'elaborazione di innumerevoli dati di input consistenti nella traduzione in linguaggio "machine-readable" di ritmi, melodie e armonie, e che il sistema "The Next Rembrandt"

sviluppato da Microsoft è in grado di apprendere lo stile del pittore olandese e di riprodurlo per creare non copie, ma dipinti nuovi⁵⁰.

Ciò premesso, è intanto chiaro che le opere create dall'IA possono essere agevolmente riconducibili nel novero di quelle proteggibili da parte del diritto d'autore in considerazione del fatto che il gradiente di creatività richiesto dal diritto italiano è (come anticipato) assai basso⁵¹. Invero, secondo la tesi maggiormente seguita, anche in sede unionale, è opera dell'ingegno pure quella che è frutto di una creatività "semplice", che non solo – cioè – porta ad esprimere idee semplici, ma che è la mera risultante di una scelta discrezionale fra le varie alternative possibili: il che pare più in linea con le applicazioni giurisprudenziali che, ad esempio, ritengono proteggibili le traduzioni, le trascrizioni delle lezioni universitarie o le stesse banche di dati e che, per quanto qui interessa, potrebbe agevolmente portare a concedere protezione al risultato cui conduce la IA riguardato, per così dire, oggettivamente⁵².

⁴⁹ Si veda l'articolo di M. TESTA, *Sunspring, il primo film di fantascienza scritto da una AI*, 13 giugno 2016, disponibile presso www.datamanager.it.

⁵⁰ M. FERRO, I. MARINI, L. NAZZICONE e P.L. PARENTI, *Presentazione del Corso "Le sfide dell'intelligenza artificiale al diritto, tra neopositivistica fiducia nella tecnica e rimeditazione delle categorie giuridiche"*, Scuola Superiore della Magistratura, 19-21 marzo 2026.

⁵¹ Sul punto, di nuovo, CASSANO e TASSONE, *La disciplina del plagio evolutivo*, cit., 214 ss.

⁵² In effetti, secondo un consolidato approdo della giurisprudenza la protezione del diritto d'autore postula i requisiti della originalità e della creatività che devono dirsi sussistere anche quando l'opera sia composta da idee e nozioni semplici, comprese nel patrimonio intellettuale di persone aventi esperienza nella materia propria dell'opera stessa, purché formulate ed organizzate in modo personale ed autonomo rispetto alle precedenti (Cass. 13 giugno 2014, n. 13524, in *Repertorio Foro Italiano* 2014, *Diritti d'autore*, n.° 76; Cass. 28 novembre 2011, n. 25173, in *OneLegale*; Cass. 12 marzo 2004, n. 5089, in *Foro it.*, 2004, I, 2441; nonché Cass. 2 dicembre 1993, n. 11953, in *Foro it.* 1994, I, 2416). In particolare, la creatività consiste non già nell'idea che è alla base della sua realizzazione, ma nella forma della sua espressione, ovvero nella sua soggettività, di modo che la stessa idea può essere alla base di diverse opere d'autore, come è ovvio nelle opere degli artisti, le quali tuttavia sono (o

possono essere) diverse per la creatività soggettiva che ciascuno degli autori spende, e che in quanto tale rileva per l'ottenimento della protezione (Cass. 29 maggio 2020, n. 10300, in *Rep. Foro It.*, 2020, *Diritti d'autore*, n.° 70). Peraltro, tali principi di diritto, enunciati dalla giurisprudenza nazionale (da ultimo Cass. 10 febbraio 2025, n. 3393, in *Ced Cass. civ.*, 2025, rv 67395101 (m)) sono conformi ai principi elaborati, in materia, dalla giurisprudenza della CGUE. Secondo quest'ultima, muovendo dal testo della Dir. 2001/29/CE del Parlamento Europeo e del Consiglio del 22 maggio 2001 (sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione), la nozione di "opera" implica che esista un oggetto originale, nel senso che detto oggetto rappresenta una creazione intellettuale propria del suo autore, anche se la qualifica di opera è riservata agli elementi che sono espressione di tale creazione (CGUE 12 settembre 2019, Cofemel, C-683/17, punto 29, nonché CGUE 13 novembre 2018, Levola Hengelo, C-310/17, punti 33 e da 35 a 37, entrambe in *eur-lex.europa.eu*). Dunque, affinché un oggetto possa essere considerato originale è necessario ma al contempo sufficiente che rifletta la personalità del suo autore, manifestando le scelte libere e creative di quest'ultimo (CGUE 12 settembre 2019, Cofemel, cit., punto 30, nonché CGUE 7 agosto 2018, Renckhoff, C-161/17, punto 14, e CGUE 1° dicembre 2011, Painer, C 145/10, punti 88, 89 e 94, tutte in *eur-lex.europa.eu*). Infine, la nozione di

Del resto, uno sguardo ai Paesi di *Common Law* o a sistemi misti fa rilevare che in alcuni di essi si ammette che un'opera dell'ingegno possa essere creata da un computer, sulla scorta della Section 178 dello UK Copyright, Designs and Patents Act del 1988, alla base delle normative di settore di vari ordinamenti⁵³: il quale poi precisa (attraverso una formula di non agevole interpretazione) che i diritti sull'opera vengono comunque attribuiti all'uomo⁵⁴.

Secondo altra tesi, invece, il carattere creativo va inteso come un particolare modo di espressione dell'autore, sicché l'opera deve essere ottenuta con un lavoro intellettuale dell'autore stesso che porta all'estrinsecazione della sua individualità e a imprimere in essa la sua impronta personale, cioè l'espressione del suo personale modo di sentire o interpretare la vita.

"opera" importa necessariamente l'esistenza di un oggetto identificabile con sufficiente precisione e oggettività (CGUE 12 settembre 2019, Cofemel, cit., punto 32, e CGUE 13 novembre 2018, Levola Hengelo, cit., punto 40). In proposito, è stato infatti spiegato che «[g]li articoli da 2 a 5 della direttiva 2001/29 del Parlamento europeo e del Consiglio, del 22 maggio 2001, sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione, devono essere interpretati nel senso che la protezione a titolo del diritto d'autore da essi prevista si applica a un prodotto la cui forma è, quantomeno in parte, necessaria per ottenere un risultato tecnico, qualora tale prodotto costituisca un'opera originale risultante da una creazione intellettuale in quanto, mediante tale forma, il suo autore esprime la propria capacità creativa in maniera originale effettuando scelte libere e creative, di modo che detta forma riflette la sua personalità, circostanza che spetta al giudice del rinvio verificare tenendo conto di tutti gli elementi pertinenti della controversia principale» (CGUE, sez. V, 11 giugno 2020, n. 833/18, in *eur-lex.europa.eu*). Ciò con la precisazione, infine, che nella materia che qui ci occupa si ha violazione dell'esclusiva non solo quando un'opera è copiata integralmente, cioè quando vi sia riproduzione abusiva, ma anche nel caso della contraffazione che ricorre quando i tratti essenziali dell'opera anteriore si ripetono in quella successiva. E tali tratti essenziali dell'opera (che dunque non sono passibili di replicazione) vanno individuati muovendo da quanto è frutto della creatività del suo autore; quel che rileva, dunque, è la forma espressiva intesa nella sua soggettività, vale a dire la scelta dell'autore stesso quanto alla rappresentazione dell'idea, non l'idea in quanto tale (Cass. 29 luglio 2025, n. 21851, in *ForoPlus*).

E per tale via vengono in rilievo profili per così dire soggettivi, che cioè attengono non tanto alle caratteristiche dell'opera di per sé considerata – aspetto in definitiva non critico –, quanto al problema inerente al se l'autore debba essere o no, necessariamente, una persona fisica⁵⁵.

Di fronte ad un tale scenario, una volta distinta la fattispecie della IA usata come strumento (analogamente alla macchina fotografica) e quella della IA che crea (stando per così dire "in regia"), l'opzione è in linea di principio netta: attribuire a chi sia titolare dei diritti di sfruttamento dell'IA stessa i diritti sull'opera dell'ingegno oppure attribuirli alla medesima IA. In proposito, a sostegno della tesi che nega la possibilità di attribuire agli algoritmi i diritti previsti dalla legge sul diritto d'autore vi sono, apparentemente, gli artt. 11 e 29 L.d.a., dai quali

⁵³ La norma così dispone: «'computer-generated', in relation to a work, means that the work is generated by computer in circumstances such that there is no human author of the work». Similmente, secondo la Section 2(1) dell'Irish Copyright and Related Rights Act del 2000 «'computer-generated', in relation to a work, means that the work is generated by computer in circumstances where the author of the work is not an individual». Ancora, per la Section 2 dell'Indian Copyright Act del 1957 (per come novellata nel 1994) «(d) 'author' means [...] (vi) in relation to any literary, dramatic, musical or artistic work which is computer-generated, the person who causes the work to be created». Sulla soluzione adottata nel Regno Unito, R.C. DENICOLA, *Ex Machina: Copyright Protection for Computer-Generated Works*, in *Rutgers University Law Review*, 69, 2016, 251 ss., il quale si sofferma anche sulle normative di settore neozelandese, sudafricana e di Hong Kong.

⁵⁴ Secondo la Section 9.3 «[i]n the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken».

⁵⁵ In effetti, è ritenuto di soluzione assai più semplice il problema inerente alla possibilità di considerare il lavoro di una IA "creativo" nel senso indicato e molti ritengono tutelabile l'opera generata da una macchina laddove il sistema di IA sia in grado di operare scelte espressive autonome, non predeterminate né predeterminabili dall'intervento dell'uomo: per tutti, S. GUIZZARDI, *La protezione dell'autore dell'opera dell'ingegno creata dall'intelligenza artificiale*, in *AIDA*, 2018, 42 ss. e, più di recente, CAPPARELLI, *Le nuove frontiere del diritto d'autore*, cit., 335 ss., seppur rimanendo scettica sulla possibilità di riconoscere i relativi diritti in capo all'IA.

si desume che a talune persone giuridiche spetta «il diritto di autore sulle opere create e pubblicate sotto il loro nome ed a loro conto e spese», quindi entro confini e in casi limitatissimi⁵⁶.

Emerge dunque che un tratto caratterizzante l'ordinamento italiano e quelli continentali più in generale risiede nel fatto che solo l'autore può essere titolare dei diritti morali, mentre per quelli patrimoniali l'acquisto in capo alla persona giuridica è – tendenzialmente – solo derivativo⁵⁷.

Di contro, i sistemi di *Common Law* non prevedono in linea di principio i diritti morali e attraverso meccanismi di pubblicità costitutiva sostanzialmente estranei alla nostra tradizione giuridica – piuttosto tributari della vicinanza del *Copyright Law* al *Patent Law* – configurano come naturale la nascita dei diritti patrimoniali direttamente in capo alla persona giuridica committente⁵⁸.

Pertanto, è assai chiaro che quando la legge italiana sul diritto d'autore poneva l'accento sul fatto che l'autore è solo la persona fisica, la stessa – prima della modifica di cui alla Legge 132/2025, sulla quale si tornerà in appresso – voleva delimitare l'efficacia della fattispecie acquisitiva rispetto alla persona giuridica, non anche rispetto all'IA.

In altri termini, se per il diritto privato generale l'equiparazione persona fisica / persona giuridica – nel tempo estesa ad altri enti non muniti della personalità giuridica e indicati come meri soggetti di diritto – è piena con il limite della compatibilità, un analogo discorso non vale per la normativa di settore di cui si tratta: gli enti sono giuridicamente incapaci rispetto ai diritti morali e possono acquisire quelli patrimoniali, salvo eccezioni, solo a titolo derivativo.

È allora evidente che l'assimilabilità dell'IA munita di certe caratteristiche alla persona fisica bypassa in radice le suddette obiezioni: non potendosi far dire alla legge italiana sul diritto d'autore nel testo vigente prima della Legge n. 132/2025 quanto essa *non* diceva, ossia che la indicazione della persona fisica come autore valeva *ad excludendum* anche rispetto alla stessa IA.

Del resto, nella direzione indicata va la riflessione in corso sul concetto di creatività, nel senso che pare ormai obsoleta l'immagine romantica dell'autore che compone l'opera in solitudine che ha ispirato suddetta concezione, la quale mette in ombra i reali processi che oggi stanno alla base della produzione delle opere dell'ingegno e della cultura in generale⁵⁹: mentre altri studiosi hanno registrato l'ampliamento dell'ambito oggettivo della

⁵⁶ Mentre l'art. 29 regola con apposite norme la durata dei diritti patrimoniali, secondo l'art. 11 «[a]lle amministrazioni dello Stato, alle Province ed ai Comuni, spetta il diritto di autore sulle opere create e pubblicate sotto il loro nome ed a loro conto e spese. Lo stesso diritto spetta agli enti privati che non perseguano scopi di lucro, salvo diverso accordo con gli autori delle opere pubblicate, nonché alle accademie e agli altri enti pubblici culturali sulla raccolta dei loro atti e sulle loro pubblicazioni».

⁵⁷ Ancora assai attuali le considerazioni di V. M. DE SANCTIS, *I soggetti del diritto d'autore*, Milano, 2005, *passim*, per il quale, d'altronde, il diritto d'autore rientrava a pieno titolo all'interno del diritto civile. Nel senso che le norme di cui nel testo sono ora intese nel senso che il diritto di paternità rimane in capo all'autore, mentre i diritti di sfruttamento economico sono acquistati a titolo derivativo, essendo stata superata la tesi che riteneva un tale acquisto a titolo originario e dibatteva sulle sorti del diritto morale: per tutti P. GALLI, *Ad art. 6 L. 22 aprile 1941, n. 633 in UBERTAZZI, Commentario breve alle leggi su proprietà intellettuale e concorrenza*, cit., 1676 ss.

⁵⁸ Sui temi di cui nel testo, U. IZZO, *Alle origini del copyright e del diritto d'autore - Tecnologia, interessi e cambiamento giuridico*, Bari 2010, 45 ss. e *passim*.

⁵⁹ Sul punto G. AJANI, *Contemporary Artificial Art and the Law: Searching for an Author*, collana «Art and Law», 3.4, Leiden 2020, 44, che cita M. ROSE, *Authors and Owners. The Invention of Copyright*, Cambridge, 1993, 3 ss. (secondo il quale «the figure of the romantic author sits monolithically at the core of copyright law, obscuring important realities about the collective nature of creativity and misrepresenting the actual processes of cultural production»), nonché A. BRIDY, *Coding Creativity: Copyright and the Artificially Intelligent Author*, in *Stanford Technology Law Review*, 5, 2012, 5 ss. e C. CRAIG e I. KERR, *The Death of AI Author*, in *Osgoode Legal Studies Research Paper*, 17 maggio 2019. Proprio sulla scorta delle riflessioni di cui nel testo AJANI, *Contemporary Artificial Art and the Law*, cit., 45, afferma che «a reflection on authorship in AI-generated artworks is particularly valuable not only for its capacity to propose new solutions to fill some gaps in the law, but also because it brings scholars and practitioners with different expertise (like

protezione e l'estensione di essa a nuovi soggetti che caratterizzano l'evoluzione del diritto d'autore in connessione con il mutamento delle forme di espressione, di diffusione e di fruizione delle opere dell'ingegno⁶⁰.

Infine, in parziale contrasto con la tesi che nega il riconoscimento sembra collocarsi la citata pronuncia della Suprema Corte di Cassazione n. 1107/2023 per la quale, come si è visto, l'utilizzo di un *software* «per generare l'immagine [...] è pur sempre compatibile con l'elaborazione di un'opera dell'ingegno con un tasso di creatività che andrebbe solo scrutinato con maggior rigore»⁶¹.

Ovviamente, anche nella prospettiva che mostra una maggiore apertura rimane fermo il cennato limite della compatibilità, nel senso che vi saranno alcune norme – si pensi a quelle che regolano il rapporto con gli eredi – non applicabili. Ma qui la dinamica torna ad essere quella tradizionale, che attribuisce una capacità giuridica piena solo alle persone fisiche.

Inoltre, il discorso svolto per il diritto italiano vale anche per il diritto euro-unitario. In effetti, sembra che quando lo stesso definisce in varie direttive – quali quelle sopra citate – l'autore come persona fisica, ciò accade perlopiù all'interno della contrapposizione tra persona fisica-persona giuridica che ha accompagnato la riflessione sulla titolarità dei diritti d'autore,

stante la presa di coscienza del maggior peso assunto dalle industrie culturali a scapito dell'individuo e, di nuovo, non *ad excludendum*⁶²: né gli ostacoli posti dall'EPO appaiono insuperabili, considerando che la tutela del *software* è affidata primariamente al diritto d'autore, nel contesto del quale il problema deve essere risolto, di là dalle poche aperture che il diritto brevettuale ha avuto con il noto caso Dabus in Australia e Sud Africa⁶³.

Infine, vero è che nell'ordinamento nordamericano è da molto più tempo che si è negata la possibilità di attribuire il *copyright* all'IA, ma proprio gli argomenti usati portano a chiedersi quanto l'evoluzione tecnologica debba farne rivedere gli esiti, dacché – tanto per fare un esempio – una decisione della Corte Suprema spesso citata e che ha ormai più di trent'anni, fa leva sul concetto di assemblaggio ormai superato⁶⁴.

D'altronde, che le opere create dall'IA – così come una risposta dicotomica agli interrogativi posti – creino non pochi problemi emerge indirettamente dalla voce che propone di introdurre per le stesse un regime analogo a quello delle opere cinematografiche al fine di attribuire al produttore un diritto connesso sulle eventuali creazioni delle macchine intelligenti con una durata più breve delle opere dell'ingegno classiche⁶⁵.

software engineers, philosophers, lawyers, curators) to review their positions on the identification of art making as a practice that is less and less made 'by hand of the artist'», concludendo che laddove non si occupa del problema delle opere «generated and selected» dall'IA senza coinvolgimento umano «the current law is not satisfying» (p. 76).

⁶⁰ V. FRANCESCHELLI, *Gli ottant'anni della legge sul diritto d'autore*, in *Riv. dir. ind.*, 2021, 267 ss.

⁶¹ Cass. 16 gennaio 2023, n. 1107, cit.

⁶² Per l'esame delle fonti unionali e il rilievo di cui nel testo, C. VENANZONI, *Intelligenze umane e artificiali. Profili giuridici nella circolazione delle opere dell'ingegno e problematiche sull'identità autorale*, in ALPA, *Diritto e intelligenza artificiale*, cit., 305, la quale però si mostra contraria al riconoscimento dei diritti d'autore in capo alle macchine, diversamente da GUIZZARDI, *La protezione dell'autore dell'opera dell'ingegno*, cit., 42 ss.

⁶³ Sul tema L. SCHIUMA, *Il software tra brevetto e diritto d'autore*, in *Riv. dir. civ.*, 2007, 683 ss., mentre sul caso

Dabus si veda A. NICO, *Brevetti – Intelligenza artificiale: diritto morale e designazione di inventore*, in *La Nuova Giurisprudenza Civile Commentata*, 2023, 227 ss.

⁶⁴ Notissimo il caso *Feist Publication-Rural Telephone Service Company, Inc.* 499 U.S. 340 (1991), nel quale la Corte Suprema degli Stati Uniti ha affermato che «il lavoro intellettuale è fondato sui poteri creativi della mente» (nostra traduzione), così negando che un assemblaggio di dati, effettuato utilizzando un *software*, potesse contenere qualche forma espressiva unica e irripetibile, tanto da acquisire dignità di opera intellettuale ed essere tutelata dalla normativa sul diritto d'autore. Il *Federal Copyright Office* ha allora dichiarato nel suo *Compendium of U.S. Copyright Office Practices* che avrebbe consentito la registrazione di opere originali solo se queste fossero state create da un essere umano. In dottrina, cfr. J. GINSBURG, *The Concept of Authorship in Comparative Copyright Law*, in *De Paul Law Review*, 52, 2003, 1063 ss.

⁶⁵ E. AREZZO, *Gli investimenti in "attrezzature per produzioni di avanguardia tecnologica" come strumento*

E volendo solo per un attimo legare il dibattito sviluppatosi in tema di diritto d'autore a quello più generale, proprio dalla considerazione di tale area del diritto viene l'idea per cui il problema dell'attribuzione dei diritti all'IA può essere affrontato considerando vari modelli, anche se non tutti esportabili al di fuori di tale contesto, come ad esempio quello che in certi casi attribuisce i diritti, paradossalmente, all'utilizzatore delle opere stesse⁶⁶.

In ultimo, sempre nel contesto appena indicato si è portata a sostegno della tesi incline ad attribuire all'IA la soggettività una possibile analogia fra essa e le società di capitali⁶⁷: il che ci conduce a tratteggiare, seppur rapidamente, i termini di quella che è stata indicata come "soggettività dimezzata".

7. La "soggettività dimezzata"

Che l'uomo sia al centro del sistema e che tale debba rimanere non è discutibile⁶⁸.

E l'esigenza di protezione dell'IA, se del caso tramite il modello della soggettività giuridica o della personalità elettronica, nel settore del

diritto d'autore o in altri, nasce sempre da un interesse dell'uomo⁶⁹.

Un tale interesse può essere legato al fatto che il mancato impiego di tale modello contrasta con il comune sentire, che esso si ritenga il più confacente anche sotto il profilo della tutela degli interessi patrimoniali della collettività o che sia supportato da altre ragioni ancora⁷⁰.

Al riguardo, si è visto che l'IA ha già un grado di autonomia prima impensabile e – se si guarda alla storia del progresso tecnologico – è facile presagire che essa sarà presto dotata di caratteristiche che oggi non riusciamo nemmeno ad immaginare compiutamente, anche in base alla nota legge di Moore⁷¹.

Dunque, messe da parte tanto le paure quasi paranoiche quanto una cieca fiducia, ossia gli atteggiamenti che spesso si hanno da tempo di fronte a tale progresso e che sono probabilmente figlie delle «utopie di angoscia» e delle «utopie di desiderio» che hanno caratterizzato il XX Secolo, un approccio che tenda a regolare preventivamente la soggettività

di ripartenza economica. La tutela dell'IA e dei suoi prodotti tramite diritto d'autore e diritti connessi, in AIDA, 2022, 333 ss.

⁶⁶ Su tale fenomeno G. D'IPPOLITO, *L'esigenza di un nuovo bilanciamento per il diritto d'autore: gli user generated content e l'ipotesi di un'eccezione per le opere creative e trasformative*, in *Cyberspazio e diritto*, 2017, 513 ss., mentre per la rassegna dei variegati modelli che il diritto d'autore propone e che sono ulteriori rispetto alla secca alternativa attribuzione piena / negazione totale della soggettività si veda B. TASSONE, *Intelligenza artificiale, soggettività giuridica e personalità elettronica in prospettiva di comparazione*, in C. CASCIONE, G. GIANNONE CODIGLIONE e P. PARDOLESI (a cura di), *Public and Private in Contemporary Societies - Atti del Colloquio Biennale della Associazione Italiana di Diritto Comparato*, Bari e di Taranto, 25-27 maggio 2023, 2020, Roma, 587 ss.

⁶⁷ Si veda ancora AJANI, *Contemporary Artificial Art and the Law*, cit., 59 ss.

⁶⁸ In proposito, E. DEL PRATO, *Le basi del diritto civile*, Torino, 2020, 236, dopo aver notato che la capacità giuridica rientra fra le tecniche attraverso cui l'ordinamento dà rilevanza giuridica alla realtà, precisa che comunque «il diritto è in funzione dell'uomo per garantire la migliore convivenza possibile, non per proteggere interessi in sé diversi da quelli dell'uomo» e che «[a]nche gli interessi della collettività e di enti rispecchiano, direttamente o indirettamente, interessi

delle persone fisiche», il che «deve costituire il canone interpretativo con cui risolvere le questioni collegate a realtà diverse dall'uomo». Dunque, come rileva M. COSTANZA, *L'AI: de iure condito e de iure condendo*, in RUFFOLO, *Intelligenza artificiale*, cit., 410-412, il punto non è quello di «creare una nuova categoria di essere umani o quasi umani», né quello di «modificare le assiologie delle leggi esistenti», sicché anche di fronte all'IA «la valutazione d'adeguatezza o meno della normativa non implica un giudizio sui fondamentali».

⁶⁹ Peraltro, nel contesto del diritto degli animali non è mancato chi ha cercato di attribuire agli stessi la soggettività giuridica, come emerge dalla letteratura citata in R. GARRETO, *Soggettività giuridica dell'animale. Rilevi critici*, in F. BILOTTA e F. RAIMONDI (a cura di), *Il soggetto di diritto – Storia ed evoluzione di un concetto nel diritto privato*, Napoli, 2020, 165 ss., cui adde F.P. TRASCIA, *Animali e umani: il tentativo di un inquadramento razionale e unitario dell'animale anche nel nostro sistema giuridico*, Napoli, 2021.

⁷⁰ In proposito, vale ad esempio la pena di ricordare la teoria di Lawrence Lessig in base alla quale la tecnologia è una forza in grado di produrre un proprio codice, che si impone sulla regola giuridica, scardinandone gli effetti.

⁷¹ Secondo la quale il numero di componenti elettronici (ad esempio i transistor) che formano un chip raddoppia ogni anno.

dell'IA potrebbe anche valere a impedire che essa «si imponga» per via tecnologica⁷².

D'altronde, sulla scia delle leggi della robotica di Isaac Asimov e grazie ad un eccezionale incrocio fra diritto e letteratura, la roboetica da tempo si confronta con il problema di come "educare" l'IA, il che apre la strada all'affascinante costruzione di un separato sistema di regole tecnologiche volte a disciplinare il modo in cui la stessa prende le proprie decisioni⁷³. E – in prospettiva – non si deve scartare l'idea che la medesima possa imparare dai propri "errori" anche grazie al fatto che i medesimi incidono sulla sua sfera giuridica⁷⁴: pur se ad oggi l'assenza di coscienza e istinto di conservazione rende problematica l'applicazione delle regole di responsabilità a tale fine⁷⁵.

Tuttavia, tali embrionali considerazioni portano paradossalmente ad esaltare il ruolo del diritto privato, laddove lo stesso può assai meglio di

quello penale e amministrativo approntare idonei strumenti per la tutela delle vittime e lo spostamento di risorse dal patrimonio del soggetto giuridico IA a quello del danneggiato. E un passo più in là si colloca l'interrogativo inerente al se la tassazione volta alla creazione di fondi pubblici (al netto del fallimento del progetto relativo alla c.d. "Robotax") o comunque sistemi di *no-fault compensation* come quello neozelandese possano essere ancora più idonei, secondo un interrogativo sul quale mette conto tornare in altra sede⁷⁶.

Qui, più che soffermarsi su tali profili, è utile formulare una considerazione di fondo che emerge da un complessivo sguardo all'evoluzione tecnologica in materia.

Pur senza riprendere tale riflessione nella terza edizione del suo testo, un Autore ricorda in quella precedente che la Dichiarazione Universale dei Diritti Umani del 1948 riconosce

⁷² Il riferimento è a S. RODOTÀ, *Protezione dei dati e circolazione delle informazioni*, in *Riv. crit. dir. priv.*, 1984, 721-722, il quale rileva che «il passaggio dalle utopie di 'desiderio' alle utopie di 'angoscia' sembra avere caratterizzato in maniera incisiva l'intero corso del XX secolo. Una attitudine ondivaga nei confronti del progresso tecnologico ha portato, da un lato, a una integrazione sempre più forte, ai limiti della dipendenza, tra l'individuo e le nuove possibilità offerte dalla scienza applicata, e, dall'altro lato, a una paura diffusa sui rischi, reali o presagiti, che tale integrazione reca con sé. Ma l'angoscia del futuro non porta al rifiuto del futuro: accanto alla percezione, sempre più netta, dei rischi del progresso tecnologico, c'è la consapevolezza della impossibilità di arrestare tale progresso, anche se non si presenta più con stimate soltanto positive».

⁷³ Premesso che non è questa la sede per esaminare i temi e problemi della roboetica – per tutti F. PASQUALE, *Le nuove leggi della robotica*, Roma, 2021, nonché L. FLORIDI, *Etica dell'intelligenza artificiale*, Milano, 2022 –, è frequente in dottrina (ad esempio A.B. SUMAN, *Intelligenza artificiale e soggettività giuridica: quali diritti (e doveri) dei robot?*, in ALPA, *Diritto e intelligenza artificiale*, cit., 267 ss.) la notazione per cui essa ha origine prendendo spunto dalle prime tre leggi della robotica di Isaac Asimov (Id., *Io, robot*, Milano, 1987) cui è stata poi aggiunta "la legge 0" (Id., *Robots and Empire*, New York, 1985). Si tratta in specie di quelle seguenti «0. Un robot non può recare danno all'umanità, né può permettere che, a causa del proprio mancato intervento, l'umanità riceva danno 1. Un robot non può recare danno a un essere umano né può permettere che, a causa del proprio mancato intervento, un essere umano riceva danno. 2. Un robot deve obbedire

agli ordini impartiti dagli esseri umani, purché tali ordini non contravvengano alla Prima Legge. 3. Un robot deve proteggere la propria esistenza, purché questa autodifesa non contrasti con la Prima e con la Seconda Legge».

⁷⁴ Profila alcuni modi in cui la rieducazione a seguito dell'errore potrebbe avvenire U. RUFFOLO, *Machina delinquere potest? Responsabilità ed "illeciti" (anche penali?) della "persona elettronica" e tutele per gli agenti software autonomi*, in Id., *XXVI Lezioni*, cit., 301.

⁷⁵ Proprio qui, in effetti, viene in esponente il problema della percezione delle conseguenze delle violazioni, perché la minaccia della sanzione e la stessa punizione hanno per l'essere umano un significato e funzioni in principio non replicabili. Per ulteriori considerazioni sul tema CINGOLANI e ANDRESCIANI, *Robot, macchine intelligenti e sistemi autonomi*, cit., 40 ss., nonché W. WALLACH e C. ALLEN, *Moral Machines, Teaching Robots Right from Wrong*, New York, 2010, *passim*.

⁷⁶ Si veda K. ABNEY, *Robotics, ethical theory, and metaethics: a guide for the perplexed*, in P. LIN, K. ABNEY e G.A. BEKEY (a cura di), *Robot Ethics: The Ethical and Social Implications of Robotics*, London 2012, 35 ss. Da notare che nella proposta di Risoluzione recante raccomandazioni alla Commissione Europea concernente norme di diritto civile sulla robotica, con un considerando poi scomparso nel testo finale, si esprimeva il timore relativo all'impatto che la robotica avrebbe potuto avere sul tasso occupazionale, tanto da ritenere opportuna l'introduzione di una tassa, la c.d. *Robotax*, che avrebbe dovuto colpire i redditi prodotti da entrate "digitali", per poi costituire un fondo per la tutela e la riqualificazione dei lavoratori licenziati a causa della tecnologia digitale.

vari diritti e libertà fondamentali a tutti gli uomini indicandoli come dotati di ragione e coscienza, chiedendosi se quando saremo in grado di munire le macchine di tali caratteristiche (in misura simile all'essere umano) le macchine stesse «dovrebbero essere, per definizione, escluse dal club degli umani»: soprattutto a fronte di confini biologici dell'individuo che vengono continuamente ridefiniti, come dimostrano gli strumenti che consentono l'idratazione, la nutrizione e la respirazione artificiali, le tecniche di riproduzione assistita e le operazioni tese al cambio di sesso⁷⁷; cui aggiungere ora la progettazione di impianti che possono entrare nel midollo spinale per ripristinare il movimento di soggetti paralizzati o dare la vista a persone cieche dalla nascita⁷⁸.

Del resto, proprio in questa direzione una recente voce ha messo in evidenza come la distinzione fra siliceo e biologico non corrisponda più a quella fra artificiale e naturale in ragione dei progetti di ricerca volti a sviluppare «sistemi wetware [...] che cercano di duplicare cervelli biologici per produrre un cervello sintetico»⁷⁹: senza contare che le neuroscienze e la neurobiologia si stanno cimentando nella coltivazione di cellule cerebrali nella prospettiva di arrivare alla costruzione di un cervello totalmente biologico quanto alla sua materia, ma evidentemente 'artificiale' quanto alla sua origine⁸⁰.

Di poi, un argomento un po' forzato e radicale che la voce per prima citata propone (per profilare una possibile discriminazione di rilievo costituzionale in danno delle macchine), induce

a domandarsi (dopo averlo riadattato) come il diritto si dovrà porre in futuro di fronte a esseri umani che quelle caratteristiche abbiano in misura così bassa da dover essere muniti, grazie alla IA, di ragione e coscienza perlopiù artificiali⁸¹.

E gli ultimi sviluppi del progetto Neuralink che ha per ora superato il vaglio della *Food and Drug Administration* dimostrano che potrebbe trattarsi di un futuro non troppo lontano⁸²: dacché è del 2024 la notizia della installazione di un impianto cerebrale sul primo essere umano con il fine di «potenziare le capacità umane, curare disturbi neurologici come la SLA o il Parkinson e forse un giorno realizzare una relazione simbiotica tra uomo e intelligenza artificiale»⁸³.

Inoltre, le tecniche di *human enhancement* e di ibridazione uomo-macchina che si sono menzionate inducono a profilare almeno altri due scenari che possono stimolare una più compiuta riflessione in materia, sempre nell'ottica di valutare l'esigenza dell'ordinamento (e quindi della collettività) di estendere il concetto di soggettività o di costruirne uno su misura: anche perché se l'IA è in grado di imparare, prendere decisioni al di fuori degli *input* originari e agire con relativa autonomia, ciò vuol dire che essa ha oggettivamente una storia e non rimane mai uguale a se stessa.

Il primo scenario riguarda l'interrogativo inerente alla misura in cui il regime giuridico che riguarda gli animali andrà ripensato ove gli strumenti testé menzionati possano incrementare non solo la loro capacità di

⁷⁷ A. SANTOSUOSSO, *Diritto, scienza e nuove tecnologie*, Padova, 2016², 354 e 356.

⁷⁸ Riferisce dei due ultimi progetti di cui nel testo, F. SARACINA, *Neuralink, interfaccia computer-cervello, in sei mesi*, 1° dicembre 2022, disponibile su www.ilsole24ore.it.

⁷⁹ C. PERLINGIERI, *Responsabilità civile e robotica medica*, in *Tecnologie e diritto*, 2020, 178-179.

⁸⁰ Per vari riferimenti RUFFOLO, *La personalità elettronica*, in Id., *Intelligenza artificiale*, cit., 119.

⁸¹ SANTOSUOSSO, *Diritto, scienza e nuove tecnologie*, cit., 357, che formula il radicale esempio di esseri umani totalmente privi di ragione e coscienza che la acquisiscano solo grazie alle macchine, per profilare una violazione dei diritti fondamentali in danno delle «entità artificiali».

⁸² Quanto al progetto di cui nel testo, si veda l'articolo di L. FORLANI, *Chip nel cervello, primi test in 6 mesi: il piano di Neuralink*, 1° dicembre 2022, disponibile presso www.corrierecomunicazioni.it, il quale spiega che il chip fatto costruire da Elon Musk contiene un impianto cerebrale informatico in grado di leggere le onde cerebrali e trasmettere informazioni, ad esempio per consentire che una persona affetta da sclerosi laterale amiotrofica (SLA) possa comunicare con l'esterno attraverso i propri pensieri.

⁸³ Nel senso di cui nel testo l'articolo a firma della redazione *Musk: Neuralink ha installato il primo impianto cerebrale su un uomo* disponibile presso www.corriere.it.

trasmettere informazioni all'esterno, ma anche di ragionare e avere coscienza di sé.

Il secondo scenario riguarda la sempre crescente possibilità di interazione fra macchina e sistema nervoso fino al punto di ipotizzare il trasferimento di una coscienza umana su supporti artificiali, magari per garantire la sopravvivenza ad una mente contenuta in un corpo in decadimento⁸⁴: se ciò divenisse possibile, una volta affermato che il soggetto continua ad avere la sua precedente soggettività e capacità giuridica sebbene sia venuto meno il suo sostrato biologico, sarebbe più difficile negarlo ad una mente elettronica che semplicemente “nasce” – volendo così evocare il tema dell'applicazione estensiva o analogica dell'art. 1 c.c. – senza averlo mai avuto⁸⁵.

Tornando ad un futuro più immediato – ed ecco il terzo scenario – l'ipotesi più ragionevole sembra quella di pensare ad uno statuto speciale, segnatamente ad una capacità asimmetrica nei diritti e doveri – o comunque caratterizzata da un regime *ad hoc* – da declinare nelle diverse aree dell'ordinamento.

E riprendendo il filo della esposizione nel settore del diritto d'autore il diritto morale potrebbe diventare allora una forma di garanzia per il consumatore senza che all'IA sia attribuita la paternità dell'opera nel senso tradizionale del

termine⁸⁶: sì da potersi profilare l'ipotesi della costruzione di una “soggettività dimezzata”⁸⁷.

Tuttavia, proprio tale ipotesi è da ritenersi “scartata”, almeno per ora, per via delle scelte legislative che hanno informato la Legge n. 132/2025 e la conseguente novella della legge sul diritto d'autore.

8. L'IA, le modifiche introdotte dal Legislatore con la Legge 132/2025 e la giurisprudenza della UE.

A seguito della Legge 23 settembre 2025, n. 132 (Disposizioni e deleghe al Governo in materia di intelligenza artificiale) il nuovo testo dell'art. 1 della Legge n. 633/1941 vuole ora che «sono protette ai sensi di questa legge le opere dell'ingegno umano di carattere creativo che appartengono alla letteratura, alla musica, alle arti figurative, all'architettura, al teatro ed alla cinematografia, qualunque ne sia il modo o la forma di espressione, anche laddove create con l'ausilio di strumenti di intelligenza artificiale, purché costituenti risultato del lavoro intellettuale dell'autore» (I comma).

Il Legislatore del 2025 – che, curiosamente, non ha modificato anche e di conseguenza l'art. 2575 c.c. – accorda così tutela anche alle opere create con l'ausilio di strumenti di IA, a condizione che esse siano il “risultato del lavoro intellettuale dell'autore”.

⁸⁴ Ne riferiscono U. RUFFOLO e A. AMIDEI, *Intelligenza artificiale, human enhancement e diritti della persona*, in RUFFOLO, *Intelligenza artificiale*, cit., 179, nt. 1.

⁸⁵ Nel rinviare ancora a TASSONE, *Riflessioni su intelligenza artificiale e soggettività giuridica*, cit., 214 ss., sia per l'ipotesi di cui nel testo, sia per una rassegna delle varie ricostruzioni proposte dalla dottrina italiana sul tema della soggettività giuridica, si dirà per il resto che il problema è ancora di là da venire e non per poco. Ma ciò non costituisce motivo sufficiente a non porcelo affatto, almeno in prospettiva, anche in considerazione di altri temi che il progresso tecnologico porta ad affrontare: come ad esempio dimostra il fatto che, ove non fosse più vietato e ove fosse medicalmente possibile procedere ad un espianto della testa o del cervello – o addirittura coltivare la distopica prospettiva di un essere umano clonato o letteralmente fabbricato in laboratorio –, ecco che vi sarebbero ulteriori ragioni per riflettere sulla relativizzazione del sostrato biologico cui la capacità giuridica e la soggettività vengono tradizionalmente legate. Sulla clonazione e «fabbricazione» dell'essere

umano C. CORVI, *La clonazione. Profili giuridici*, Milano, 2022, mentre richiamano le problematiche giuridiche che gli interventi di cui nel testo solleverebbero, sempre RUFFOLO e AMIDEI, *Intelligenza artificiale*, cit., 208-209.

⁸⁶ Nel settore della responsabilità civile ad obblighi risarcitori pieni potrebbero corrispondere varie limitazioni rispetto alla gestione del patrimonio, magari primariamente funzionalizzato al ristoro della vittima e ancor prima a sostenere i costi di un'assicurazione obbligatoria, mentre nel settore del diritto societario un algoritmo che sedesse nel C.d.A. potrebbe avere diritti di voto solo su certe materie o, abbracciando la più verosimile prospettiva della delega parziale, all'IA potrebbero essere affidati compiti di *compliance* (si pensi alla *privacy* o al controllo sulla sicurezza informatica) in settori specifici.

⁸⁷ Per gli ulteriori elementi che caratterizzano tale proposta, TASSONE, *Intelligenza artificiale, soggettività giuridica e personalità elettronica in prospettiva di comparazione*, cit., 587 ss.

La prospettiva è dunque quella antropocentrica. Né avrebbe potuto essere diversamente, giacché tutto il sistema della Legge n. 132 è improntato su una tale “dimensione” dell’IA.

Ciò detto, dalla lettura del novellato art. 1 della Legge n. 633/1941 sembra a prima vista che, con riferimento al tema qui trattato non sorgano questioni nuove per le opere frutto dell’esclusivo apporto umano (perché esse non si confrontano punto con la GenAI) così come per quelle che sono frutto dell’esclusivo apporto della IA, giacché, queste ultime, sono messe fuori gioco dall’impianto normativo (perché gli strumenti di intelligenza artificiale, come visto, devono essere solo di ausilio).

Tuttavia, se si va oltre le considerazioni generali appena esposte, sotto l’apparente superficie tranquilla del disposto normativo si agitano le non poche “correnti” che una analisi più approfondita fa emergere, come già messo in luce in altra sede⁸⁸.

Al riguardo e in primo luogo, non è agevole comprendere cosa debba intendersi, in concreto, per opera dell’ingegno *umano*. Come si vedrà fra un momento, in senso ampio ogni creazione mediata dall’IA potrebbe dirsi comunque frutto dell’ingegno umano perché l’IA è, in ultima analisi, costruita dall’uomo e l’input da cui prende avvio il procedimento creativo, ad oggi, proviene sempre dall’uomo.

In questa prospettiva, analogo margine di incertezza riguarda la nozione di «ausilio» e l’inciso «purché costituenti risultato del lavoro intellettuale dell’autore»: non è infatti agevole stabilire dove si debba tracciare la linea di confine tra quello che può qualificarsi come *lavoro intellettuale dell’autore* e quello che – per via di un intervento artificiale che abbia sconfinato dal perimetro dell’ausilio, espropriando l’autore del controllo sul processo creativo e declassandolo a mero “utente” – non possa più considerarsi tale. Sicché un ulteriore interrogativo riguarda la sorte di tutte quelle “opere” che, in conseguenza di ciò, non ricadono in detto schema.

La norma, lasciando spazio a scenari di collaborazione uomo-macchina più o meno estesi, impone allora di prestare particolare attenzione alla progettazione del flusso creativo e alla documentazione del contributo umano, sì da poter garantire, in concreto, la tutela autoriale dell’opera così generata⁸⁹.

Al riguardo, è evidente che una risposta definitiva alle questioni interpretative poste dalla norma in commento potrà essere data solo con il tempo, quando la giurisprudenza sarà – inevitabilmente – chiamata a definire il perimetro della tutela delle opere AI-assisted, di là dagli spunti che provengono di per sé dalla “storia” della norma⁹⁰: anche perché sul punto

⁸⁸ Si vedano B. TASSONE e S. MANNI, *Commento all’art. 25 L. IA: creatività e ausilio artificiale*, in G. CASSANO e E.M. TRIPODI, *Intelligenza artificiale, la disciplina italiana, Legge 23 settembre 2025, n. 132*, Pisa, 2025, 383 ss.

⁸⁹ Stante la regola aurea per cui l’insorgenza del diritto d’autore non dipende dall’adempimento di alcuna formalità e nasce al momento della creazione dell’opera (art. 5 par. 2 della citata Convenzione di Berna), la progettazione e documentazione appena cennate avranno ruolo fondamentale in un eventuale, quanto probabile, momento contenzioso. E allora, se come affermato da V. IAIA, *Prova e carattere creativo dell’opera dell’ingegno*, in G. CASSANO, V. FRANCESCHELLI e B. TASSONE, *Prova e tutele nel diritto d’autore e nel diritto industriale*, Milano, 2023, 18, l’onere probatorio in capo all’autore anteriore è più rigoroso in caso di opere assistite da *software* (e ancora più da IA), una documentazione e progettazione in chiave “difensiva” del processo creativo può risultare quanto mai utile. La paternità invece, essendo di ben più difficile dimostrazione in quanto conseguenza di un processo interno che ha luogo nella

mente dell’autore, tende ad essere, di fatto, presunta, con onere in capo a chi la contesta di assolvere alla relativa prova, come rammenta L. CHIMIENTI, *Prova della paternità dell’opera dell’ingegno*, in CASSANO, FRANCESCHELLI e TASSONE, *Prova e tutele nel diritto d’autore e nel diritto industriale*, cit., 24-29. V’è dunque da chiedersi se tale conclusione valga anche in caso di una paternità di Opera AI-assisted, potenzialmente contestata al fine di ricondurla al pubblico dominio, piuttosto che per attribuirla a chi la rivendica, secondo considerazioni più ampiamente sviluppate in TASSONE e MANNI, *Commento all’art. 25 L. IA.*, cit., 383 ss.

⁹⁰ Invero, la formulazione della disposizione nella stesura originaria del DDL che prevedeva, dopo le parole «anche laddove create con l’ausilio di strumenti di intelligenza artificiale», l’aggiunta dell’inciso «purché il contributo umano sia creativo, rilevante e dimostrabile», avrebbe quanto meno tracciato i binari guida per individuare il contributo umano necessario (e sufficiente). Tuttavia, come segnalato in G. CASSANO, *Il DDL italiano in materia di IA*, in CASSANO e TRIPODI, *Il regolamento europeo*

non emergono rilievi di particolare interesse dal relativo *dossier* di controllo parlamentare⁹¹.

Ciò detto, se un merito va riconosciuto al nuovo art. 1 L.D.A. – almeno sotto il profilo della chiarezza del testo normativo –, è quello di aver incontrovertibilmente circoscritto la tutela autoriale alle opere frutto dell'ingegno umano. In proposito, senza qui esporre le considerazioni inerenti all'opportunità o no di una tale scelta, essa appare coerente con l'impostazione umanistica che, come si diceva, informa l'intero approccio unionale alla regolazione dell'IA⁹²: anche se – come si è detto – essa non è necessariamente in contrasto con soluzioni interpretative atte a ipotizzare l'attribuzione dei diritti d'autore, di alcuni di essi o di diritti *ad hoc* previsti, a centri di imputazione diversi dall'essere umano.

In particolare, la selezione operata dalla norma determina la formazione di una matrice categoriale di opere calibrata sul diverso grado di intervento dell'IA nel processo creativo, tra le quali solo le prime due meritano la "O" maiuscola, mentre le seconde due vengono lasciate in una sorta di vuoto normativo, venendosi a dover distinguere fra Opere interamente create dall'uomo, Opere create dall'uomo con l'ausilio di strumenti di IA, opere "co-create" e opere create dall'IA.

Ebbene, messe "da parte" come già detto sia le opere interamente create dall'uomo sia quelle interamente generate dall'IA (escluse in radice dall'uso del termine «ausilio») e che, per ragioni opposte, non pongono particolari problemi, l'attenzione va rivolta necessariamente alla perimetrazione delle categorie intermedie e ai criteri che determinano la ricaduta di un'opera

nella seconda ovvero nella terza categoria (con conseguente attrazione o esclusione dalla tutela)⁹³: non tanto alla «umanità» dell'ingegno, che appare più che altro una dichiarazione d'intenti, quanto agli elementi che consentono di ritenere che una creazione costituisca «risultato del lavoro intellettuale dell'autore» nonostante l'impiego ausiliario di strumenti di IA⁹⁴.

Orbene, l'impiego di strumenti, anche complessi, non incide di per sé sulla titolarità del diritto d'autore, *purché l'apporto dell'autore si configuri come espressione della sua capacità intellettuale e creativa*. Tanto è stato chiarito, in relazione all'arte digitale, dalla già citata pronuncia n. 1107/2023, laddove la Suprema Corte non ha solo osservato che la generazione di un'immagine mediante un software «compatibile con l'elaborazione di un'opera dell'ingegno con un tasso di creatività *che andrebbe solo scrutinato con maggior rigore*» (corsivo aggiunto), ma anche che, in relazione a tali tipi di opere, si rende «necessario un accertamento di fatto per verificare se e in qual misura l'utilizzo dello strumento avesse assorbito l'elaborazione creativa dell'artista che se ne era avvalsa»⁹⁵.

In definitiva, la creatività, individuata come tratto distintivo della forma espressiva e condizione necessaria per ottenere tutela, diventa il metro di misurazione dell'apporto intellettuale dell'autore, rendendo necessario individuare i tratti distintivi di essa prima di poterla rinvenire o no nelle sue manifestazioni più inedite. Dunque, il requisito di creatività, già elemento fondamentale per concedere la protezione voluta dal diritto d'autore e

sull'intelligenza artificiale, cit., 974, l'inciso in questione era stato foriero di molteplici perplessità in dottrina.

⁹¹ *Disposizioni e deleghe al Governo in materia di intelligenza artificiale A.C. 2316 Dossier n. 206*, cit., 2.

⁹² Per vari profili di criticità della scelta operata dal Legislatore italiano, di nuovo TASSONE e MANNI, *Commento all'art. 25 L. IA.*, cit., 383 ss.

⁹³ Relativamente alla quarta categoria è d'uopo rilevare che, giacché l'IA non è ancora capace di una vera iniziativa creativa autonoma, per «opere interamente generate dall'IA» ci si riferisce necessariamente a quelle in cui l'essere umano fornisce un input non creativo né

originale, apportando un contributo pari al "premere un pulsante".

⁹⁴ In tal senso, si può ragionevolmente ritenere che l'apporto umano debba essere almeno *prevalente*, poiché la lettura sistematica della norma di cui nel testo richiede di tener conto dei principi e delle disposizioni contenute negli altri articoli della normativa sull'IA, fra cui l'art. 13, il quale relega l'utilizzo dell'IA nell'esercizio della professione intellettuale ad attività «strumentali e di supporto [...] con *prevalenza* del lavoro intellettuale», come messo in luce sempre in TASSONE e MANNI, *Commento all'art. 25 L. IA.*, cit., 383 ss.

⁹⁵ Cass. 16 gennaio 2023, n. 1107, cit.

parametro d'accertamento delle violazioni, riconferma il suo ruolo fondamentale di criterio-soglia.

In proposito, si è già visto che per essere creativa un'opera non deve essere *in toto* inventiva, unica o avere un particolare valore artistico⁹⁶. E a quanto detto circa le due principali elaborazioni dottrinali e pretorie in punto di creatività – l'una che fa leva su una creatività "semplice", intesa come «scelta discrezionale all'interno di un numero sufficientemente ampio di varianti con cui esprimere un'idea», l'altra propensa a postulare una creatività "qualificata", che presuppone un apporto personale dell'autore, tale da rifletterne la personalità o comunque da configurarsi come «attività superiore rispetto alla normale iterazione della vita quotidiana» – si può aggiungere che esse non trovano un preciso riscontro nel diritto dell'Unione: il quale sembra propendere per la prima accezione di creatività, spesso sovrapponendola al concetto di originalità, con una soglia fissata dalla CGUE che è anch'essa, comunque, piuttosto bassa⁹⁷. Come anticipato, la possibilità che un'opera dell'ingegno venga realizzata con il contributo di strumenti tecnologici più o meno sofisticati costituisce un elemento intrinseco all'evoluzione culturale e tecnica, e proprio in tale prospettiva appare utile rammentare i principi affermati

dalla Corte di Giustizia dell'Unione Europea nel caso *Eva-Maria Painer v. Standard VerlagsGmbH and Others*, il quale offre un paradigma utile anche nell'attuale contesto della creatività *AI-assisted*. La Corte di Giustizia, nel caso *Painer*, ha infatti chiarito che la creatività si manifesta attraverso una serie di "scelte libere e creative", che l'autore può compiere in diverse fasi del processo produttivo⁹⁸. Il fotografo, quindi, può scegliere lo sfondo, la posa del soggetto, l'illuminazione, l'inquadratura, l'angolo di ripresa e l'atmosfera; e – nella fase della selezione – può determinare quale scatto utilizzare e quali tecniche di sviluppo o software impiegare. Attraverso tali scelte, dunque, egli imprime all'opera il proprio "tocco personale", rendendola espressione della sua individualità⁹⁹. Analogamente, nella produzione di opere realizzate con il supporto dell'intelligenza artificiale il discrimine tra semplice ausilio tecnico e vera e propria co-creazione dovrebbe risiedere nell'individuazione di quel nucleo di scelte creative che riflettono la personalità dell'autore umano.

Con riguardo agli ulteriori tratti distintivi della creatività, talvolta definita *a contrario*, la Corte ha escluso che possano essere creative scelte dettate da necessità tecniche¹⁰⁰; ovvero

⁹⁶ L. BENTLY, B. SHERMAN, D. GANGJEE e P. JOHNSON, *Intellectual Property Law*, Oxford, 2022, 101, nonché P. GALLI, *Ad Art. 1 Legge 633/1941*, in P. MARCHETTI e L. C. UBERTAZZI (a cura di), *Commentario breve alle leggi su proprietà intellettuale e concorrenza*, V ed., Padova, 2012, 1331.

⁹⁷ Ad esempio, nella sentenza della CGUE, 16 luglio 2009, C-5/08, *Infopaq International A/S c. Danske Dagblades Forening*, fra le altre in *ForoPlus*, la Corte di giustizia dell'Unione europea ha ritenuto che una sequenza di appena undici parole possa essere considerata originale, a condizione che rifletta l'espressione creativa dell'autore. Per una recente e più ampia rassegna sul punto si vedano sempre TASSONE e MANNI, *Commento all'art. 25 L. IA.*, cit., 383 ss.

⁹⁸ È interessante rilevare che l'Avvocato generale Trstenjak aveva ritenuto che fosse lo spazio creativo disponibile a determinare l'originalità, sicché più lo spazio creativo è limitato, meno l'opera risulterà originale, rilevando altresì che "sforzo creativo" significa utilizzare la libertà formativa disponibile (caso C-145/10, Opinione dell'Avvocato generale Trstenjak

del 12 aprile 2011, ECLI:EU:C:2011:239, punto 111 e 122).

⁹⁹ CGUE, 1° dicembre 2011, C-145/10, *Eva-Maria Painer c. Standard VerlagsGmbH e Altri*, parr. 90-92, in *ForoPlus*.

¹⁰⁰ Nel caso *BSA* la Corte ha stabilito che la creatività non deve essere dettata o vincolata da necessità tecniche (CGUE, 22 dicembre 2010, C-393/09, *Bezpečnostní softwarová asociace - Svaz softwarové ochrany v. Ministerstvo kultury*, parr. 29 e 48, in *ForoPlus*). Anche in questo caso è interessante rilevare l'opinione dell'Avvocato generale, Yves Bot, il quale ha affermato che, affinché sia presente la creatività necessaria a soddisfare il requisito di originalità, l'autore non deve essere vincolato da necessità tecniche. In tali casi, infatti, «l'espressione è dettata dalla funzione tecnica di tali elementi» e «il criterio dell'originalità non [è] soddisfatto, in quanto i vari modi di realizzare un'idea sono così limitati che l'idea e l'espressione si confondono» (caso C-393/09, Opinione dell'Avvocato generale Bot del 14 ottobre 2010, ECLI:EU:C:2010:611, 75-76).

vincolate da regole¹⁰¹. Ne consegue che la creatività non può essere determinata da alcun fattore esterno che ne limiti l'autonomia espressiva: ciò che rileva, in definitiva, è che la scelta compiuta dall'autore sia effettivamente libera, frutto di un processo decisionale realmente personale¹⁰².

Ebbene, alla luce di tali orientamenti ci si può allora domandare se una scelta creativa effettuata utilizzando un sistema di *Generative AI* possa essere considerata effettivamente libera o se, al contrario, l'impiego di tali strumenti sottragga all'utente il controllo sul processo creativo.

Al riguardo, la GenAI è spesso descritta come una sorta di "scatola" nella quale l'utente inserisce un input per ottenere un output¹⁰³. Tuttavia, non di rado tale scatola assume le caratteristiche di una *black box*, poiché i processi interni che conducono al risultato non sono pienamente conoscibili né prevedibili¹⁰⁴. In un simile contesto, diventa complesso sostenere che l'autore mantenga un controllo effettivo sul procedimento creativo e, conseguentemente, che la sua scelta possa qualificarsi come realmente libera.

D'altronde, che le scelte libere e creative e l'*imprimatur* personale dato all'opera dall'autore debbano costituire il nucleo essenziale del requisito della originalità è poi stato ulteriormente ribadito dalla Corte di Giustizia, ad esempio nel caso *Football Dataco Ltd and*

Others v. Yahoo! UK Ltd and Others, in riferimento all'allestimento di database¹⁰⁵: sicché occorre verificare se dalla stessa elaborazione dei Giudici di Lussemburgo si possano trarre elementi atti a proporre una *pars construens* del ragionamento interpretativo da compiere sul citato art. 25, come sempre messo in luce in altra sede¹⁰⁶.

In proposito, proprio nella occasione appena richiamata la Corte ha precisato che l'originalità non può essere desunta dal mero impegno, dalla competenza o dall'abilità tecnica profusa nella realizzazione dell'opera. Invero, come si legge nella decisione, «il lavoro e l'abilità non possono di per sé giustificare la protezione [del database] ai sensi della Direttiva 96/9, se quel lavoro e quella abilità non esprimono alcuna originalità nella selezione o nella disposizione dei dati»¹⁰⁷.

La rilevanza del passaggio è triplice.

Primo. Se si configurassero abilità, lavoro o sforzo quali elementi necessari e sufficienti ad ottenere la tutela autoriale ai fini dell'originalità, si finirebbe per reintrodurre – sia pure in forma implicita – il criterio del *labour or skill*, o dello *sweat of the brow*, estraneo alla tradizione italiana se non anche europea e ormai, generalmente, obsoleto anche in quella di origine.

Secondo. Se tali criteri rilevassero, sarebbe impossibile riconoscere tutela autoriale a un'opera AI-assisted, poiché, nella maggior parte dei casi, l'utilizzo di tali strumenti mira proprio a

¹⁰¹ Nel caso *Football Association Premier League Ltd and Others v. QC Leisure and Others*, la Corte ha ampliato la soluzione adottata nel caso BSA, stabilendo che la creatività è assente non solo quando è vincolata da necessità tecniche, ma anche quando è vincolata dalle regole di un gioco o di un processo (CGUE, 4 ottobre 2011, C-403/08, *Football Association Premier League Ltd and Others v QC Leisure and Others*, e C-429/08, *Karen Murphy v Media Protection Services Ltd*, parr. 96-98, in *ForoPlus*).

¹⁰² Per maggiori approfondimenti sul tema si vedano sempre TASSONE e MANNI, *Commento all'art. 25 L. IA.*, cit., 383 ss.

¹⁰³ A. AMIDEI e M. MAGGIOLINO, *Intelligenza Artificiale, dati digitali e proprietà intellettuale*, in F. DONATI, A. PAJNO e A. PERRUCCI (a cura di), *Intelligenza Artificiale e diritto: una rivoluzione?*, Vol. 3, *Proprietà intellettuale, società e finanza*, Bologna, 2022, 54.

¹⁰⁴ Una *black-box* è un modello di *machine learning* che opera come sistema opaco in cui il funzionamento interno non è agevolmente accessibile né interpretabile. Tali modelli svolgono predizioni basate su dati di input, ma il processo decisionale e il ragionamento retrostante a tali decisioni non sono accessibili in modo trasparente per l'utente. Per tutti, V. HASSIJA, V. CHAMOLA, A. MAHAPATRA et Al., *Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence*, in *Cognitive Computation*, 2024, 16, 47.

¹⁰⁵ CGUE, sentenza del 1° marzo 2012, *Football Dataco Ltd and Others v. Yahoo! UK Ltd e Altri*, C-604/10, ECLI:EU:C:2012:115, parr. 38-39, in *ForoPlus*.

¹⁰⁶ Per ulteriori riflessioni si vedano sempre TASSONE e MANNI, *Commento all'art. 25 L. IA.*, cit., 383 ss.

¹⁰⁷ CGUE, sentenza del 1° marzo 2012, C-604/10, cit., par. 42.

snellire il lavoro e a ridurre il carico “operativo” che grava sull’autore. Escludendo dunque la rilevanza dell’abilità tecnica e dello sforzo materiale, diviene ammissibile che un’opera assistita dall’IA possa beneficiare della protezione del diritto d’autore anche in assenza di un particolare “sforzo” (diverso da quello strettamente creativo)¹⁰⁸.

Terzo. Il principio affermato in *Football Dataco* può essere mutuato nel contesto delle opere generate con l’ausilio di sistemi di GenAI. Il tocco personale dell’autore può, infatti, manifestarsi sia nella fase di predisposizione dell’input - attraverso istruzioni dettagliate, mirate e frutto di scelte autonome - sia nella successiva elaborazione dell’output, mediante attività di selezione, modifica o rielaborazione del materiale prodotto dal sistema. È allora in tale processo di interazione consapevole e selettiva che potrebbe continuare a risiedere la dimensione autenticamente creativa, idonea a fondare la protezione autoriale.

Del resto, anche in altre occasioni la Corte di Giustizia è tornata sul concetto di libera scelta creativa, talvolta per escludere la rilevanza di ulteriori elementi, come la qualità estetica dell’opera¹⁰⁹; talaltra per ribadire che, affinché un’opera sia considerata originale ai fini della attribuzione della tutela autoriale, «è necessario e sufficiente che rifletta la personalità del suo autore, manifestando le scelte libere e creative di quest’ultimo»¹¹⁰.

Pertanto, così fissati gli elementi caratterizzanti dell’autorialità nel diritto euro-unitario e interno, ne consegue che, per mantenere l’IA relegata a funzione di mero ausilio, essi debbano rimanere prerogativa dell’autore umano, e che non possano essere ad essa “outsourced”. L’opera AI-assisted tutelabile è allora un’opera rispetto alla quale l’autore ha fatto delle scelte creative libere e alla quale ha in qualche modo impresso il proprio tocco personale.

In questa prospettiva, allora, la creatività risiede nelle scelte dell’utente che orientano, filtrano e valorizzano i risultati prodotti dal sistema senza che il processo algoritmico di generazione automatica possa usurparla. Il fatto di selezionare tra molteplici possibilità, nonché di combinare o modificare gli output per realizzare una determinata visione estetica o concettuale, rappresenta il momento in cui l’autore dà all’opera il proprio *imprimatur*. Proprio in tale spazio decisionale continua a concretizzarsi l’apporto creativo tutelato dal diritto d’autore¹¹¹.

9. (segue:) Originalità progressiva, fasi del processo creativo e opere in pubblico dominio

La libera scelta creativa si esprime all’interno di un processo che coinvolge strumenti di intelligenza artificiale e può manifestarsi grazie a - e nonostante - l’uso dell’IA. In particolare, è stato teorizzato che il processo creativo si articoli in tre fasi - concezione, esecuzione e redazione - , all’interno delle quali il controllo creativo può rendersi evidente in modi diversi¹¹².

¹⁰⁸ La dottrina dello *sweat of the brow* o della “industriosità” si concentra sull’abilità e sul lavoro necessari per creare l’opera e ritiene che essi possano costituire fondamento dell’originalità dell’opera. Sul punto si rinvia a Drassinower, *Sweat of the Brow, Creativity and Authorship: On Originality in Canadian Copyright Law*, in *University of Ottawa Law and Technology Journal*, (2003–2004) 1 UOLTJ 105, 107.

¹⁰⁹ Nel caso *Cofemel c. G-Star Raw* la Corte ha stabilito che «è certamente vero che considerazioni di ordine estetico partecipano dell’attività creativa. Tuttavia, la circostanza che un modello generi un effetto estetico non consente, di per sé, di determinare se tale modello costituisca una creazione intellettuale che riflette la libertà di scelta e la personalità del suo autore, che soddisfa pertanto l’esigenza di originalità evocata ai punti 30 e 31 della presente sentenza» (CGUE, 12

settembre 2019, C-683/17, *Cofemel – Sociedade de Vestuário SA v G-Star Raw CV*, par. 54, in *ForoPlus*).

¹¹⁰ CGUE, 11 giugno 2020, C-833/18, *SI, Brompton Bicycle Ltd c. Chedech / Get2Get*, par. 23, in *ForoPlus*.

¹¹¹ Si tratta delle conclusioni raggiunte sempre in TASSONE e MANNI, *Commento all’art. 25 L. IA.*, cit., 383 ss.

¹¹² Il riferimento è al lavoro di P. B. HUGENHOLTZ e J. P. QUINTAIS, *Copyright and Artificial Creation: Does EU Copyright Law Protect AI-Assisted Output?*, in *IIC – International Review of Intellectual Property and Competition Law*, 2021, 52, 1190-1216, mentre K. MILITSYNA, *Human Creative Contribution to AI-Based Output – One Just Can’t Get Enough*, in *GRUR International*; 2023, 72, 10, 939-949, propone un test in cinque parti per stabilire se l’apporto creativo umano sia sufficiente a ottenere la tutela autoriale, e tre diverse soluzioni in base all’esito di tale test.

La fase di concezione concerne l'ideazione dell'opera, ma non si limita alla mera formulazione dell'idea. Essa include tutte le scelte creative che determinano il modo in cui l'opera si esprimerà. È in questa fase che si applica il requisito che condiziona la tutela secondo cui in positivo l'autore deve essere il creatore intellettuale dell'opera e in negativo non deve coinvolgere l'IA nella generazione dell'idea in sé: in altri termini il *prompt*, non inteso come mera idea ma come insieme di istruzioni volte a svilupparla, deve rispondere al "concepimento" originale dell'autore. Qualora l'IA fosse incaricata di generare output a partire da idee generiche, il risultato sarebbe interamente frutto della macchina e quindi non tutelabile per mancanza di paternità umana. Se, invece, l'idea appartiene all'autore e l'*input* è abbastanza soggettivo da riflettere il suo tocco personale, la fase di concezione soddisfa pienamente il requisito ed è riconducibile alla creazione intellettuale dell'autore: anche se è proprio nella zona grigia del *prompt crafting* (*id est*: la predisposizione accurata e dettagliata dei *prompt*) che la linea di confine tra Opere ausiliate (continuando a usare il maiuscolo di cui al precedente paragrafo per quelle protette) e opere co-create è più sfumata¹¹³.

In questa fase, in specie, non si tratta soltanto di sviluppare l'idea fondante, ma anche di compiere tutte quelle scelte che definiranno la direzione dell'opera: dalla determinazione dei dettagli che l'IA dovrà includere, alla lunghezza del testo (se trattasi di opera testuale), allo stile, agli elementi visivi da rappresentare (se trattasi di opere visive), fino alle numerose altre decisioni che rendono l'*input* più ricco e preciso (cui secondo alcuni si aggiunge la stessa scelta del sistema di IA da utilizzare, che costituisce un'espressione creativa propria della fase di concezione)¹¹⁴. Tutte queste scelte confluiscono

in definitiva nel *prompt*, che fornisce all'IA contesto, parametri e indicazioni entro cui operare¹¹⁵.

La fase di esecuzione riguarda invece la traduzione dell'idea e delle modalità di implementazione definite nella fase di concezione della bozza dell'opera finale. È in questa fase che l'IA può assumere un ruolo determinante, elaborando l'*input* dell'autore per produrre un *output*. Qui la principale criticità risiede nella citata opacità del processo, soprattutto nei sistemi di *deep learning*, il cui apprendimento autonomo rende difficile per l'autore comprendere pienamente ciò che la macchina ha appreso, generando il menzionato effetto "black box" e un potenziale rischio di perdita di controllo sull'opera.

Tuttavia, stante l'abbandono della dottrina dello *sweat of the brow*, l'esecuzione dell'opera intesa come mera operazione meccanica svolta dall'IA non pregiudica la possibilità di ottenere tutela, dovendo la paternità essere conferita alla persona che concepisce, dirige e supervisiona il processo creativo¹¹⁶.

Infine, la fase di redazione. Essa può rappresentare un momento di recupero del controllo creativo mediante selezione e aggiustamento dei risultati ottenuti: ad esempio, se durante la fase di esecuzione l'autore richiede all'IA di generare plurimi *output* e successivamente li combina e li modifica per ottenere varianti tra cui selezionare quella definitiva, sta compiendo scelte creative autonome. E ulteriori interventi di manipolazione dell'*output* rafforzano ancora di più il suo *tocco personale*.

Dunque, così collocato il confine tra l'opera che trae supporto dall'IA e risulta protetta e quella che, invece, risulta esclusa dalla protezione, è necessario soffermarsi ora sulla zona grigia delle opere create, o co-create, dall'IA. Trattasi,

¹¹³ F. MAZZI, *Authorship in Artificial Intelligence-generated Works: Exploring Originality in Text Prompts and Artificial Intelligence Outputs Through Philosophical Foundations of Copyright and Collage Protection*, in *The Journal of World Intellectual Property*, 2024, 27, 3, 410-427, si sofferma ampiamente sul ruolo fondamentale del *prompt* di cui nel testo.

¹¹⁴ La criticabile affermazione di cui nel testo si deve a P. B. HUGENHOLTZ – J. P. QUINTAIS, *Copyright and Artificial Creation: Does EU Copyright Law Protect AI-Assisted Output?*, in *International Review of Intellectual Property and Competition Law*, 4 ottobre 2021, Vol. 52, 1190 ss.

¹¹⁵ Così, sempre TASSONE e MANNI, *Commento all'art. 25 L. IA.*, cit., 383 ss.

¹¹⁶ *Ibidem*, 383 ss.

invero, delle opere in cui l'apporto dell'autore è così scarso da non rendere percepibile nel risultato il suo tocco personale e da non determinare la presenza di una scelta creativa vera e propria. A tali opere l'ordinamento non riconosce tutela, sicché viene da chiedersi se tale scelta sia la più opportuna, lasciando in una sorta di "limbo" tutta una serie di opere, il destino delle quali è, allo stato, il pubblico dominio.

Orbene, benché di primo acchito potrebbe sembrare che ci si trovi in presenza di un vuoto normativo, il summenzionato "limbo" pare sia una scelta di sistema, coerente con la tradizione europeo-continentale, in una certa prospettiva connessa al *droit d'auteur* con tutto ciò che ne consegue in punto di diritti morali. Ciò in quanto non esiste nell'ordinamento una soluzione tale per cui potrebbe essere conferita all'autore la titolarità di un'opera che non ha, di fatto, creato. Né è agevole teorizzare una soggettività dei sistemi di IA tale per cui i diritti derivanti dalla titolarità dell'opera ricadrebbero in capo ai sistemi medesimi, come sopra si è detto.

In proposito, giova ricordare una volta ancora il meccanismo di funzionamento dei modelli di intelligenza artificiale (in particolare, dei c.d. Large Language Models), i quali, analogamente a sofisticati "pappagalli stocastici", non sono (quantomeno, allo stato attuale) dotati di coscienza e operano, piuttosto, attraverso complesse inferenze statistiche, limitandosi a prevedere la sequenza di parole più probabile sulla base dei dati che hanno a disposizione: da cui la già citata notazione per cui l'IA non possiede una reale comprensione di ciò che esprime né ha piena coscienza delle azioni che compie¹¹⁷. Né può dirsi libera, poiché, a ben vedere, non possiede la capacità di creare in modo autonomo, in assenza dell'intervento e

della guida dell'essere umano, la cui ingerenza si manifesta tanto nella fase di *training*, quanto in quella di codificazione e predisposizione del sistema, nonché nella formulazione del *prompt*, sicché l'intero procedimento creativo risulta, in definitiva, eterodiretto dall'uomo: perlomeno muovendosi all'interno del paradigma tradizionale della creatività, sul quale si è soffermata l'attenzione ai fini della interpretazione del nuovo testo dell'art. 1 della legge sul diritto d'autore per come novellato¹¹⁸. Di contro, si è pure detto che la figura dell'autore celebrata come genio unico, traboccante di creatività (tipica del Rinascimento e alla base del *droit d'auteur*) è ben superata¹¹⁹; ed è altresì chiaro che opere del tipo suindicato, pur carenti sotto il profilo dell'apporto umano e quindi allo stato non protette dal diritto d'autore, possono comunque avere un significativo valore economico e culturale, il che porta a soffermarsi sulla seconda modifica che la legge n. 132/2025 apporta alla legge n. 633/1941, ossia quella in tema di utilizzazioni libere.

10. Addestramento dei sistemi di IA e tutela autoriale

Come noto, l'IA generativa di immagini con carattere di originalità – facilmente accessibile grazie anche ai suoi costi, spesso, piuttosto contenuti – richiede agli operatori giuridici l'analisi della rilevanza dell'immissione, da parte di un utente, di un *prompt* in un sistema di *deep learning* per la creazione di una immagine nuova e inedita.

Al riguardo, è noto che l'IA generativa ha bisogno di attingere a molte informazioni (nell'ordine di miliardi) per il suo apprendimento e che spesso attinge anche ad opere coperte dal diritto d'autore, ponendo il problema della tutela di

¹¹⁷ Una volta dato l'input, l'IA, sulla base dei dati forniti e di quelli che ha a propria disposizione, determina come abbinare le informazioni nel modo statisticamente più probabile: l'IA imita, non pensa. E ciò spiega anche le allucinazioni, perché non è in grado di capire cosa dice. In proposito, si vedano M. STERPI e F. DI LAZZARO, *Sistemi di IA e proprietà intellettuale*, in CASSANO e TRIPODI, *Il regolamento europeo sull'intelligenza artificiale*, cit., 173, nonché, più di recente, F. COSTANTINO, *Sa tutto, o quasi, ma*

non capisce nulla – L'intelligenza artificiale e il primato del senso, Roma, 2026, *passim*.

¹¹⁸ Quelle qui esposte sono sempre le conclusioni raggiunte in TASSONE e MANNI, *Commento all'art. 25 L. IA.*, cit., 383 ss.

¹¹⁹ Basti pensare all'introduzione del repertorio delle opere digitali nel registro delle opere protette (DM rep. 319 del 3 ottobre 2024) a tutela di *creator digitali* e *influencers* per rendersi conto che è così.

quest'ultimo, con soluzioni normative tra loro differenti.

Così, mentre negli Stati Uniti si ricorre alla generale *doctrine* del "fair use" (disciplinata dalla U.S. Copyright Law, Chapter 1, Section 107) che legittima l'incorporazione di materiale protetto da *copyright* in opere aventi scopo di critica, insegnamento, ricerca e giornalismo, l'esperienza euro-unionale si fonda sull'eccezione per il c.d. "Text and Data Mining" (TDM)¹²⁰.

In Italia, in particolare, il comma 1 bis dell'art. 70 L. n. 633/1941 espressamente consente «la libera pubblicazione attraverso la rete internet, a titolo gratuito, di immagini e musiche a bassa risoluzione o degradate, per uso didattico o scientifico e solo nel caso in cui tale utilizzo non sia a scopo di lucro. Con decreto del Ministro per i beni e le attività culturali, sentiti il Ministro della pubblica istruzione e il Ministro dell'università e della ricerca, previo parere delle Commissioni parlamentari competenti, sono definiti i limiti all'uso didattico o scientifico di cui al presente comma».

A sua volta il successivo art. 70 *ter* consente le riproduzioni compiute da organismi di ricerca, e da istituti di tutela del patrimonio culturale, per scopi di ricerca scientifica, ai fini dell'estrazione di testo e di dati da opere o da altri materiali disponibili in reti o banche di dati cui essi hanno lecitamente accesso, nonché la comunicazione al pubblico degli esiti della ricerca ove espressi in nuove opere originali.

Di poi, a norma dell'art. 70 *quater* «[f]ermo restando quanto previsto dall'articolo 70-ter, sono consentite le riproduzioni e le estrazioni da opere o da altri materiali contenuti in reti o in banche di dati cui si ha legittimamente accesso ai fini dell'estrazione di testo e di dati. L'estrazione di testo e di dati è consentita quando l'utilizzo delle opere e degli altri materiali non è stato espressamente riservato dai titolari del diritto d'autore e dei diritti

connessi nonché dai titolari delle banche dati» (l comma).

Infine, l'art. 70 *septies* stabilisce che «[f]ermo restando quanto previsto dalla Convenzione di Berna per la protezione delle opere letterarie ed artistiche, ratificata e resa esecutiva ai sensi della legge 20 giugno 1978, n. 399, le riproduzioni e le estrazioni da opere o da altri materiali contenuti in rete o in banche di dati a cui si ha legittimamente accesso, ai fini dell'estrazione di testo e di dati attraverso modelli e sistemi di intelligenza artificiale, anche generativa, sono consentite in conformità alle disposizioni di cui agli articoli 70-ter e 70-*quater*»¹²¹.

Orbene, proprio quest'ultima norma, inserita dall'art. 25, l, lett. b), L. n. 132/2025, espressamente include la disciplina del text and data mining nella L. n. 633/1941, dimostrando la volontà del Legislatore di tutelare le esigenze di addestramento dei sistemi di intelligenza artificiale.

Essa, inoltre, trova il suo referente più immediato nel Regolamento del Parlamento europeo e del Consiglio (che stabilisce regole armonizzate sull'intelligenza artificiale) n. 2024/1689/UE, e precisamente nell'art. 53 sugli «Obblighi dei fornitori di modelli di IA per finalità generali» per quanto dal raffronto dei due testi emerga come l'art. 53 non faccia espresso riferimento, a differenza del testo italiano, al text and data mining (che in verità si trova menzionato al Considerando 105).

In proposito, si è già detto che per addestrare i propri modelli i fornitori di intelligenza artificiale di uso generale (GPAI) necessitano di ampi dataset, i quali possono includere anche materiali protetti da diritto d'autore. E qui si deve aggiungere che i modelli di GPAI si "allenano", *inter alia*, su enormi quantità di dati recuperati in rete per mezzo di *web crawlers* (programmi che navigano sul *web* in autonomia per eseguire un insieme indefinito di operazioni,

¹²⁰ Sul tema, più di recente, S. VIGLIAR, *Responsabilità e diritto d'autore nell'era dell'intelligenza artificiale*, in V. D'ANTONIO e B. MEOLI, *La responsabilità per danni da intelligenza artificiale*, Sant'Arcangelo di Romagna, 2026, 365 ss.

¹²¹ Per un più analitico commento alle norme di cui nel testo si vedano sempre TASSONE e MANNI, *Commento all'art. 25 L. IA.*, cit., 383 ss.

fra le quali la “pesca a strascico” di materiali, protetti o no che essi siano), sicché le tecniche di estrazione implicano atti di riproduzione di testo e dati che, ove tutelati dal diritto d'autore, necessitano di un'eccezione per essere legittimamente posti in essere¹²²: di talché viene in rilievo l'art. 4 della Direttiva UE 2019/790, attuata nel nostro ordinamento proprio con gli artt. 70 *ter* e *quater*.

L'art. 25 della Legge 132/2025, con il nuovo art. 70-*septies*, estende dunque l'applicabilità specifica degli artt. 70-*ter* (eccezione per estrazione di testo e dati compiuta da organismi di ricerca e da istituti di tutela del patrimonio culturale per scopi di ricerca scientifica) e 70-*quater* (eccezione per estrazione di testo e dati generale) ai sistemi e modelli di Intelligenza Artificiale.

E l'inclusione della disciplina delle estrazioni di testo e dati (c.d. *text and data mining* o *TDM*) nel tessuto normativo della legge n. 633/1941 dimostra l'attenzione del Legislatore anche alle esigenze di addestramento dei sistemi di IA, quindi all'evoluzione tecnologica, e non solo alla tutela degli autori, di nuovo venendo tentato un delicato bilanciamento tra diritto d'autore e altri diritti.

Ciò detto e volendo ora guardare più da vicino alla norma in questione, essa appare ricalcare il citato art. 53, co. 1, lett. c, Reg. (UE) n. 1689 del 2024, che impone ai fornitori di modelli di attuare «una politica volta ad adempiere al diritto dell'Unione in materia di diritto d'autore e diritti ad esso collegati» e di «*individuare e rispettare, anche attraverso tecnologie all'avanguardia, una riserva di diritti espressa a norma dell'articolo 4, paragrafo 3, della direttiva (UE) 2019/790*» (corsivo aggiunto)¹²³: con ciò richiedendo una particolare attenzione per l'opzione di *opt-out* ivi prevista, ma con una differenza fondamentale.

Difatti, la norma europea prevede direttamente l'obbligo di rispettare il diritto di *opt-out* dei

titolari, senza disporre l'applicazione dell'eccezione per TDM in alcun articolo, benché ne sia menzionata l'applicabilità al Considerando 105 del Regolamento.

La norma italiana, per parte sua, disciplina l'applicabilità dell'eccezione TDM mediante un rinvio agli artt. 70-*ter* e 70-*quater* l.d.a. Tale tecnica di rinvio non equivale, tuttavia, ad un affievolimento dell'obbligo di rispetto della riserva: l'*opt-out* resta operante, perché previsto dallo stesso art. 70-*quater* richiamato e perché l'art. 53, par. 1, lett. c), del Reg. (UE) 2024/1689 ne impone direttamente l'osservanza ai fornitori di modelli.

Né a diversa conclusione conduce il richiamo all'art. 9, par. 2, della Convenzione di Berna. La disposizione - espressamente fatta salva dall'incipit dell'art. 70-*septies* («fermo restando quanto previsto dalla Convenzione di Berna») - enuncia un limite alle eccezioni, non una facoltà di ampliarle: il *three-step test* opera come tetto, subordinando ogni eccezione ai requisiti cumulativi del caso speciale, dell'assenza di conflitto con il normale sfruttamento e dell'assenza di ingiustificato pregiudizio agli interessi del titolare. In questa esatta funzione limitativa esso è stato applicato dal Landgericht München I nel caso GEMA, per negare - e non per estendere - l'operatività dell'eccezione TDM rispetto alla memorizzazione. Letta insieme all'art. 5, par. 2, della medesima Convenzione (protezione automatica, indipendente da formalità), la clausola di salvezza impone anzi di interpretare l'art. 70-*septies* in senso conforme al primato del diritto esclusivo, e non come varco verso una lettura più permissiva dell'*opt-out*.

Ovviamente, a valle di tale affermazione restano poi da precisare le modalità di esercizio di tale opzione, così come la questione della responsabilità in caso di mancato rispetto della medesima - che la novella del 2025 ha inteso presidiare anche sul versante penale, per il tramite della nuova lett. a-*ter* dell'art. 171 l.d.a.,

¹²² EUROPEAN PARLIAMENT THINK TANK, *AI and copyright: The training of general-purpose AI, measure 1.3, Identify and comply with rights reservations when crawling the World Wide Web*, 2025, disponibile presso www.europarl.europa.eu.

¹²³ Nel senso di cui nel testo C. SGANGA, *Art. 53*, in A. MANTELERO, G. RESTA e G. M. RICCIO (a cura di), *Intelligenza artificiale - Commentario*, Milano, 2025, 500.

che punisce chi «riproduce o estrae testo o dati da opere o altri materiali disponibili in rete o in banche dati in violazione degli artt. 70-ter e 70-quater, anche attraverso sistemi di intelligenza artificiale».

In proposito, proprio sulle modalità d'espressione dell'*opt-out* si registra un'ulteriore discrepanza tra diritto italiano ed unionale, almeno al livello potenziale, giacché il riferimento (sebbene solo in via esemplificativa) a «strumenti che consentano la lettura automatizzata in caso di contenuti resi pubblicamente disponibili online» dell'art. 4, par. 3, Dir. (UE) 2019/790 non si rinviene, specularmente, nell'art. 70-quater.

In proposito, i fornitori di modelli di intelligenza artificiale generativa dovranno assicurare *by design e by default* che i processi di addestramento rispettino le eventuali limitazioni d'accesso imposte dai titolari dei diritti sui materiali disponibili *online*¹²⁴. E alcuni spunti utili possono venire dal Codice di Buone Pratiche per l'IA (CPAI) – sul quale si tornerà più oltre –, laddove esso dispone che occorre utilizzare tecnologie all'avanguardia per identificare e rispettare le riserve di diritti operate ai sensi dell'art. 4(3) della citata Direttiva (UE) 2019/790; impiegare *web-crawlers* che sappiano rispettare le istruzioni espresse a mezzo di Robot Exclusion Protocol (*robots.txt*), con ciò riconoscendo la validità di una riserva in tal modo espressa; nonché identificare e rispettare altri protocolli *machine-readable* a mezzo dei quali i diritti siano eventualmente stati riservati, con salvezza della riserva espressa in qualsiasi altro modo appropriato¹²⁵.

Un presupposto condiziona a monte l'operatività dell'eccezione, tanto nella variante di ricerca (art. 70-ter) quanto in quella generale soggetta a *opt-out* (art. 70-quater): il legale accesso alle opere e agli altri materiali, richiesto dall'art. 4, par. 1, della Direttiva 2019/790 e ribadito dalle

norme interne di recepimento. Ne discende che l'estrazione condotta su materiali reperiti presso fonti illecite - siti c.d. pirata, portali che ospitano caricamenti abusivi, copie non autorizzate - resta esclusa dall'eccezione in radice, e ciò a prescindere dall'esercizio o meno della riserva: il difetto di legale accesso preclude l'applicazione dell'eccezione ancor prima che venga in rilievo l'*opt-out*.

In questa direzione depone la giurisprudenza statunitense in *Bartz v. Anthropic*, ove si è affermato che l'illiceità dell'acquisizione - il download da *shadow libraries* - non è sanata dalla successiva trasformatività dell'uso; e vi converge la decisione monacense in *GEMA*, secondo cui né il diritto tedesco né quello dell'Unione conoscono un principio in forza del quale un'opera possa essere sfruttata senza consenso per il solo fatto che un terzo abbia già tenuto una condotta illecita. Nel sistema unionale la medesima conclusione trova fondamento vincolante nella sentenza *VCAST c. RTI*, ove la Corte di giustizia ha ribadito che l'eccezione riproduttiva non priva il titolare del diritto di autorizzare o vietare l'accesso alle opere (punti 34 e 39). Il disvalore si aggrava quando il sistema non si limiti ad attingere da tali fonti, ma orienti verso di esse gli utenti in sede di output.

Infine, una criticità applicativa ulteriore è rappresentata dal fatto che non solo il Legislatore europeo aveva configurato l'eccezione per il *TDM* come rivolta ad esseri umani¹²⁶: ma che – per di più – quando l'estrazione è svolta da sistemi e modelli di IA i dati sono preservati per l'addestramento, laddove la possibilità di riprodurre e conservare il materiale tutelato è invece limitata dalla norma ai fini della mera estrazione di testo e dati, con attività in teoria non sovrapponibili¹²⁷. Inoltre, la conservazione dei dati non appare, ad oggi, costruita in modo tale da evitare che

¹²⁴ Il modello di compliance è accostato a quello della *privacy by design e by default* previsto dal GDPR da SGANGA, Art. 53, cit., 501.

¹²⁵ A. PEUKERT, C. CASTETS-RENARD *et Al.* *Code of Practice for General Purpose AI Models, Copyright Chapter*, disponibile presso www.digital-strategy.ec.europa.eu.

¹²⁶ Secondo l'art. 4(2) della Direttiva (UE) 2019/790 «le riproduzioni e le estrazioni effettuate a norma del paragrafo 1 possono essere conservate per il tempo necessario ai fini dell'estrazione di testo e di dati».

¹²⁷ STERPI e DI LAZZARO, *Sistemi di IA e proprietà intellettuale*, cit., 175.

risultino *output* potenzialmente plagianti: ciò perché la conservazione di materiale consentita in origine al solo fine dell'allenamento potrebbe poi "continuare" in momenti successivi, sicché il materiale tutelato potrebbe essere in parte riprodotto nell'output finale¹²⁸.

Ebbene, poiché un tale utilizzo non è conforme all'eccezione disposta dall'art. 4 della Direttiva 2019/790, esso determina anche l'insorgere di interrogativi circa un'eventuale responsabilità dei fornitori e *deployer* di tali sistemi ai sensi dell'art. 17 della medesima Direttiva qualora questi ultimi si dovessero considerare come piattaforme che, riproducendo – essenzialmente – i dati di *training* nell'output finale, vengono a compiere un atto non consentito di comunicazione al pubblico¹²⁹.

In ultimo, considerando che «[l]e eccezioni e limitazioni di cui ai paragrafi 1, 2, 3 e 4 sono applicate esclusivamente in determinati casi speciali che non siano in contrasto con lo sfruttamento normale dell'opera o degli altri materiali e non arrechino ingiustificato pregiudizio agli interessi legittimi del titolare» (art 5, par. 5, della Direttiva InfoSoc), è chiaro che l'art. 70-*septies* non vale a espandere il margine applicativo delle eccezioni predisposte dagli artt. 70 *ter* e *quater* oltre quanto ivi espressamente previsto: senza dunque che sia in alcun modo ricompresa la fase "successiva" al *training* concernente l'emissione di output riproduttivi, o potenzialmente tali, del materiale protetto da diritto d'autore utilizzato per l'addestramento, che necessiterebbero di apposita ed ulteriore regolazione¹³⁰.

In proposito, è appena il caso di notare che su tali temi si è soffermato anche il Tribunale di Monaco con la decisione dell'11 novembre 2025 – più oltre riportata e più analiticamente commentata – resa nella causa GEMA / OpenAI. Il Tribunale ha in particolare sancito che l'utilizzo di testi di canzoni da parte della piattaforma ha

violato il diritto d'autore memorizzando testi (di canzoni) e visualizzando parti dei testi nell'output.

Il Tribunale ha dunque stabilito che non si applica la limitazione del text-and-data mining ai sensi dell'art. 4 della direttiva 2019/790 in quanto tali testi, essendo contenuti nei parametri del modello e potendo essere restituiti come output, non sono semplicemente "utilizzati" a fini analitici, ma "memorizzati" in senso proprio, cioè incorporati in modo stabile e riutilizzabile.

Orbene, proprio tale fenomeno di "memorizzazione" segna il discrimine rispetto alla mera estrazione di testo e dati che caratterizza il *text and data mining*. Nella specie, la memorizzazione si verifica quando, durante l'addestramento, i modelli linguistici non solo estraggono informazioni dal set di dati, ma mostrano anche una loro completa incorporazione nei parametri specificati dopo l'addestramento stesso, tanto che i testi delle canzoni sono anche riprodotti negli output. Quando, all'esito dell'addestramento, l'opera non è soltanto analizzata ma resta incorporata in modo stabile e riutilizzabile nei parametri del modello, sì da poter essere restituita — in tutto o in parte — negli output, non si verte più nel campo dell'estrazione di testo e dati, bensì in quello della riproduzione riservata al titolare. Ne consegue, come affermato dal Landgericht München I nel caso GEMA, che per i modelli che memorizzano la licenza è necessaria a prescindere dall'esercizio di qualsiasi riserva: l'eccezione di *text and data mining*, funzionalmente vincolata all'analisi, non copre né l'incorporazione permanente nel modello né la successiva comunicazione al pubblico dell'output. In questa prospettiva l'opt-out non è neppure il vero fronte del contendere per tali modelli, poiché la tutela opera a monte del limite: eccezione circoscritta alla fase di analisi,

¹²⁸ Sul tema *NY Times v. Microsoft e OpenAI*, nonché *Disney Enterprises Inc. v. Midjourney Inc.*, più oltre commentati.

¹²⁹ L'art. 17, par. 4, della Direttiva (UE) 2019/790 prevede infatti che «[q]ualora non sia concessa alcuna autorizzazione, i prestatori di servizi di condivisione di

contenuti online sono responsabili per atti non autorizzati di comunicazione al pubblico, compresa la messa a disposizione del pubblico, di opere e altri materiali protetti dal diritto d'autore».

¹³⁰ In argomento, ancora TASSONE e MANNI, *Commento all'art. 25 L. IA.*, cit., 383 ss.

pieno diritto esclusivo sulla memorizzazione e sull'output.

E in proposito il Tribunale ha distinto tra riproduzioni meramente preparatorie – necessarie per l'analisi dei dati – e riproduzioni che danno luogo a un'incorporazione stabile dell'opera nel modello.

Nel primo caso, la legge ammette la riproduzione temporanea in quanto funzionale alla successiva analisi. Nel secondo, invece, la riproduzione permanente inciderebbe sui diritti di sfruttamento economico dell'autore. Pertanto, il Landesgericht München I, ha condannato le convenute OpenAI USA e OpenAI Ireland a cessare la riproduzione, in tutto o in parte, senza il consenso dell'attrice GEMA, dei testi delle canzoni allegati nel fascicolo K1 all'interno di modelli linguistici (Large Language Models) e/o di far eseguire tali atti da terzi; nonché a rendere pubblicamente accessibili e/o riprodurre, in tutto o in parte e/o sotto forma di elaborazioni, senza il consenso dell'attrice, i testi delle canzoni allegati nel fascicolo K1 negli output ("Outputs") di un chatbot, e/o di far eseguire tali atti da terzi (fascicolo K2).

Le convenute sono state, altresì, condannate a fornire all'attrice le informazioni sull'estensione delle condotte descritte e quali introiti esse ne abbiano ricavato, indicando il numero e l'entità delle condotte, nonché i ricavi per ciascuna delle convenute.

È stato allora disposto che le convenute devono risarcire all'attrice ogni danno che le è già derivato e/o le deriverà dalle condotte descritte al punto 1, secondo quanto risulterà dalle informazioni rese ai sensi del punto 2, con valutazione rispetto alla quale un valido criterio di riferimento può essere costituito dal noto *three-step test* di cui all'art. 9(2) della Convenzione di Berna, nonché all'art. 5 della Direttiva 2001/29/CE, il quale prevede che il ricorso alle eccezioni e limitazioni al diritto d'autore è ammesso alle seguenti condizioni: i) caso speciale (l'uso rientra in un'eccezione chiaramente definita e di natura circoscritta), ii) assenza di conflitto con il normale sfruttamento dell'opera (l'eccezione non deve compromettere il mercato attuale o potenziale dell'opera), e iii)

assenza di pregiudizio ingiustificato agli interessi legittimi del titolare (l'impatto sui diritti dell'autore deve rimanere proporzionato e non irragionevole).

Il tutto senza dimenticare che, in ogni caso, «[l]e eccezioni e le limitazioni non possono essere applicate in modo da arrecare pregiudizio agli interessi legittimi dei titolari dei diritti o da essere in contrasto con la normale utilizzazione economica delle loro opere o materiali protetti», sicché è «possibile che la portata di alcune eccezioni o limitazioni debba essere ulteriormente limitata nel caso di taluni nuovi utilizzi di opere e materiali protetti» (considerando 44).

Un ancoraggio euro-unionale, di rango vincolante, si rinviene nella sentenza della Corte di giustizia VCAST c. RTI (C-265/16, 29 novembre 2017). Pur resa in tema di copia privata, essa enuncia principi trasversali all'intera materia delle eccezioni: che le stesse vanno interpretate restrittivamente (punto 32); che l'eccezione riproduttiva non priva il titolare del diritto di autorizzare o vietare l'accesso alle opere di cui si intenda realizzare copia (punti 34 e 39); che un servizio dotato di «doppia funzionalità» - riproduzione e, insieme, messa a disposizione delle opere (punti 37-38) - realizza una comunicazione al pubblico, da intendersi in senso lato ai sensi dell'art. 3 e del considerando 23 (punto 40), soggetta ad autorizzazione; e che, quando la trasmissione avvenga con un mezzo tecnico diverso da quello originario, essa configura una comunicazione al pubblico distinta, che esige un autonomo consenso dei titolari (punti 43 e 48-49). La Corte esclude altresì che operino, in materia di diritto d'autore, le esenzioni proprie della direttiva sul commercio elettronico (punti 24-25), e valorizza la natura commerciale dell'attività e l'intervento attivo del prestatore (punto 30 e dispositivo).

Trasposti al text and data mining per l'addestramento generativo, tali principi confermano che l'eccezione copre, al più, la riproduzione funzionale all'analisi, e in alcun modo l'incorporazione stabile nel modello né la diffusione dell'output - atto di comunicazione al pubblico per il quale, a maggior ragione ove il

prestatore assuma un ruolo attivo e persegua scopi commerciali, resta necessario il consenso del titolare.

Sul piano comparatistico, e pur trattandosi di iniziative ancora in itinere e non di diritto vigente, si segnala che in Francia ha preso corpo un movimento legislativo avente ad oggetto l'opt-in. La *proposition de loi «relative à l'instauration d'une présomption d'exploitation des contenus culturels par les fournisseurs d'intelligence artificielle»*, depositata al Senato il 12 dicembre 2025 su iniziativa dei senatori Laure Darcos, Agnès Evren e Pierre Ouzoulis, introduce una presunzione di sfruttamento che rovescia l'onere probatorio: in presenza di un utilizzo verosimile di contenuti protetti, spetterebbe al fornitore del sistema di IA dimostrare di non aver impiegato opere protette senza autorizzazione. Il testo, adottato in prima lettura e trasmesso all'Assemblée nationale, dà attuazione al rapporto d'informazione del Senato del 9 luglio 2025, «*Création et IA: de la prédation au partage de la valeur*».

Nel corso dell'esame all'Assemblée nationale la prospettiva si è fatta più netta: il relatore ha espressamente perorato il ritorno a un sistema di autorizzazione preventiva — l'opt-in in senso proprio — fondato su licenze e su un'equa remunerazione negoziata, sulla scia delle richieste degli organismi di gestione collettiva e degli editori. La stessa esposizione dei motivi muove da un rilievo che investe direttamente l'opt-out: la facoltà di riserva riconosciuta dal diritto unionale non è precisata dai testi europei ed è, in pratica, di assai difficile attuazione — critica che si salda con il nodo della machine-readability emerso nel caso Kneschke e con l'onere probatorio ivi posto a carico del titolare. Il dibattito dottrinale si dispone così lungo due poli: da un lato la tesi dell'opt-out, imperniata sulla riserva; dall'altro la tesi dell'opt-in, che ancora ogni utilizzazione al consenso preventivo del titolare e riconduce al fornitore l'onere di dimostrare la liceità dell'impiego. Dunque, la direzione del dibattito conferma che il baricentro della tutela tende a spostarsi verso il consenso preventivo.

Nella medesima direzione si colloca, in sede istituzionale, la risoluzione del Parlamento europeo del 10 marzo 2026 su diritto d'autore e intelligenza artificiale generativa — opportunità e sfide (2025/2058(INI), P10_TA(2026)0066), adottata a larga maggioranza su relazione di Axel Voss. Pur trattandosi di atto di iniziativa, e dunque privo di efficacia vincolante, la risoluzione sollecita lo sviluppo di un mercato delle licenze per l'addestramento dell'IA e sottolinea l'esigenza di garantire che i titolari dei diritti mantengano un controllo effettivo sull'utilizzo delle proprie opere nell'addestramento dei modelli e negli altri usi correlati all'IA. Ed infatti, laddove non sia stata esercitata l'opzione di esclusione, la Commissione europea è incoraggiata a facilitare lo sviluppo di meccanismi di licenza settoriali che consentano agli sviluppatori di IA di accedere a dati di addestramento di alta qualità, garantendo nel contempo un equo compenso per i creatori e i titolari dei diritti. Viene suggerito, in particolare, che i fornitori di IA erogino un compenso equo, proporzionato e non discriminatorio alla stampa e ai mezzi di informazione i cui servizi sono sfruttati. Infine, la Commissione viene invitata a valutare se i meccanismi di equa remunerazione debbano affrontare anche gli utilizzi passati di opere protette dal diritto d'autore nell'addestramento dell'IA laddove non sia ancora stato possibile istituire un mercato delle licenze.

Dunque, così delineato il punto di equilibrio trovato dai giudici tedesco e unionale — il quale è approfondito nel commento a questa ed altre decisioni (e procedure in corso) nella seconda parte del presente lavoro —, è ora il momento di riprendere il discorso su quello disegnato dal Codice di Buone Pratiche (di cui sopra si diceva), da esaminare nel più ampio contesto dei Codici di condotta previsti dall'AI Act.

11. Codice di Buone Pratiche e Codici di Condotta per l'IA

Come appena anticipato, uno dei tratti di disciplina dell'AI Act che vale la pena qui di richiamare è quello che attiene al Codice di Buone Pratiche per l'IA (CPAI), al quale i fornitori

di modelli di intelligenza artificiale generativa devono conformarsi.

Al riguardo, esso dispone che qualora i fornitori (firmatari del Codice, ossia “signatories” nel testo di cui in nota) utilizzino web crawler o li facciano utilizzare per loro conto, per estrarre o altrimenti compilare dati ai fini del text and data mining, per l'addestramento dei loro modelli di intelligenza artificiale generici, si devono impegnare a a) impiegare web crawler che leggano e seguano le istruzioni espresse in conformità al Robot Exclusion Protocol (robots.txt), e b) identificare e rispettare altri protocolli leggibili da una macchina appropriati per esprimere diritti / riserve ai sensi dell'art. 4(3) della direttiva (UE) 2019/790¹³¹. E secondo la normativa domestica tali regole sono accompagnate dalla sanzione penale per chi «atter) riproduce o estrae testo o dati da opere o altri materiali disponibili in rete o in banche di dati in violazione degli articoli 70-ter e 70-quater, anche attraverso sistemi di intelligenza artificiale» (art. 171 L. n. 633/1941 come modificato dalla L. n. 132/2025).

Al riguardo, una recente applicazione pratica di quanto esposto la si rinviene nella pronuncia del Tribunale di Amburgo (Landgericht Hamburg, Zivilkammer 10, Az. 310 O 227/23, Verkündet am 27 settembre 2024) – pure di seguito riportata e commentata – la quale, affrontando proprio il tema della liceità, o meno, della raccolta di opere protette dal diritto d'autore per creare dataset per l'addestramento di sistemi di

IA, ha giudicato legittima siffatta attività posta in essere, nel caso di specie, da una organizzazione no profit (la decisione si fonda sul § 60d UrhG che recepisce i principi ex art. 3 Direttiva 2019/790).

E, ancora, la già citata pronuncia resa dal Tribunale di Monaco di Baviera (Landgericht München I, 42. Zivilkammer, Urteil vom 11. November 2025, Az. 42 O 14139/24) – pure riportata e commentata nel presente lavoro, come già si è detto – tratta dell'impiego di testi di canzoni protette dal diritto d'autore ai fini dell'addestramento e per il funzionamento di modelli di GenAI per poi negare la possibilità di ricorrere, nella specie, all'eccezione sul text and data mining.

Per giunta, se è qui chiaro il rilievo del Codice di Buone Pratiche – al quale si dedicherà qualche altra riflessione più oltre –, qualche parola va spesa sui codici di condotta.

In proposito, è stato notato in altra sede che un'adeguata regolamentazione dei sistemi di IA risulta ostacolata dalla costante evoluzione dei sistemi stessi, che tende a renderla “vecchia” già nel momento in cui viene emanata: tanto che persino un fondamentale presupposto normativo che presiede all'applicazione della disciplina – quale è la definizione di IA – si è rivelato subito assai problematica e a rischio di obsolescenza per lo stesso legislatore unionale, oltre che per gli esperti del settore¹³².

Al riguardo, l'art. 95 del Reg. IA, collocato nel capo X intitolato «codici di condotta e

¹³¹ Il Codice di Buone Pratiche in effetti vuole che «[i]n order to help ensure that Signatories will identify and comply with, including through state-of-the-art technologies, machine-readable reservations of rights expressed pursuant to Article 4(3) of Directive (EU) 2019/790 if they use web-crawlers or have such web-crawlers used on their behalf to scrape or otherwise compile data for the purpose of text and data mining as defined in Article 2(2) of Directive (EU) 2019/790 and the training of their general-purpose AI models, Signatories commit: a) to employ web-crawlers that read and follow instructions expressed in accordance with the Robot Exclusion Protocol (robots.txt), as specified in the Internet Engineering Task Force (IETF) Request for Comments No. 9309, and any subsequent version of this Protocol for which the IETF demonstrates that it is technically feasible and implementable by AI providers and content providers,

including rightsholders, and b) to identify and comply with other appropriate machine-readable protocols to express rights reservations pursuant to Article 4(3) of Directive (EU) 2019/790, for example through asset-based or location-based metadata, that have either have been adopted by international or European standardisation organisations, or are state-of-the-art, including technically implementable, and widely adopted by rightsholders, considering different cultural sectors, and generally agreed through an inclusive process based on bona fide discussions to be facilitated at EU level with the involvement of rightsholders, AI providers and other relevant stakeholders as a more immediate solution, while anticipating the development of standards».

¹³² Sul tema B. TASSONE e S. MANNI, *Codici di condotta, intelligenza artificiale e forme di regolazione “mista”*, in *Dir. internet*, 2026, 633 ss.

orientamenti», al par. 1 dispone che «[l']ufficio per l'IA e gli Stati membri incoraggiano e agevolano l'elaborazione di codici di condotta, compresi i relativi meccanismi di governance, intesi a promuovere l'applicazione volontaria ai sistemi di IA, diversi dai sistemi di IA ad alto rischio, di alcuni o di tutti i requisiti di cui al capo III, sezione 2, tenendo conto delle soluzioni tecniche disponibili e delle migliori pratiche del settore che consentono l'applicazione di tali requisiti».

Ancorché non ci si possa qui soffermare sull'esegesi della norma, è evidente che essa divisa un meccanismo di estensione "volontaria" ai sistemi classificati a rischio basso e minimo dei requisiti previsti per i sistemi ad alto rischio¹³³: proprio per tale ragione il legislatore unionale ne auspica l'applicabilità – tramite i Codici di condotta – anche ai sistemi di classificazione diversa.

Peraltro, i Codici di condotta possono anche prevedere l'applicazione di requisiti specifici e ulteriori per tutti i sistemi di IA, indipendentemente dalla relativa classificazione di rischio, basati su obiettivi chiari e il cui perseguimento sia misurabile tramite indicatori chiave di prestazione (c.d. *Key Performance Indicators*, o *KPIs*), come precisa il par. 2 della norma.

Inoltre, con una scelta sistematica non troppo felice e una formulazione non certo cristallina la parte finale del successivo par. 3 – nel suo complesso dedicato ai soggetti coinvolti nell'elaborazione dei codici di condotta – prevede che essi «possono riguardare uno o più sistemi di IA tenendo conto della similarità della finalità prevista dei sistemi pertinenti».

In proposito e sotto il profilo soggettivo, l'elaborazione dei Codici di condotta deve essere «agevolata» dagli Stati membri e dal citato Ufficio per l'IA, mentre essa è materialmente di

competenza di una pluralità di soggetti, ossia dei singoli fornitori, dei *deployer*, di organizzazioni che li rappresentano o di entrambi: i Codici di condotta sono dunque elaborati da parte degli stessi soggetti che li devono poi applicare, con la precisazione per cui, durante la loro redazione, si deve tener conto degli interessi delle PMI e delle *startup*¹³⁴.

Per giunta, l'AI Act prevede qui la «partecipazione dei portatori di interessi e delle relative organizzazioni rappresentative, incluse le organizzazioni di società civile e mondo accademico» – come sempre precisa il par. 3 –, similmente a quanto stabiliscono varie altre sue norme, nel senso che i «portatori di interessi» sono spesso e più in generale coinvolti nella elaborazione delle regole¹³⁵.

Di poi, sotto il profilo oggettivo, i Codici di condotta possono riguardare l'applicazione dei requisiti previsti dai paragrafi 1 e 2 dell'art. 95 dell'AI Act a uno o più sistemi di intelligenza artificiale, purché accomunati da una finalità d'uso analoga o da caratteristiche funzionali simili, essendo questo il significato più coerente da attribuire alla citata proposizione di chiusura del par. 3¹³⁶.

Per giunta, anche se l'art. 95 non contiene richiami espressi ad altre disposizioni, l'attento esame del Regolamento consente di rinvenire varie altre regole che li riguardano.

Intanto, pure i codici di condotta dovranno essere applicati all'interno degli spazi di sperimentazione normativa (c.d. *regulatory sandboxes*) per l'IA ex art. 58.

Inoltre, l'art. 95 è collocato all'interno del capo dedicato anche agli «orientamenti» e qui si può aggiungere che ciò accade proprio perché i codici devono essere «ispirati» secondo il Considerando 27 dai sette principi etici dell'AI Hleg (cioè il Gruppo indipendente di esperti ad alto livello sull'intelligenza artificiale)¹³⁷.

¹³³ Sul punto, G. M. MARSICO, *L'approccio basato sul rischio*, in CASSANO e TRIPODI, *Il regolamento europeo sull'intelligenza artificiale*, cit., 377 ss.

¹³⁴ Art. 95 par. 4.

¹³⁵ Per l'inventario di tali regole sempre TASSONE e MANNI, *Codici di condotta*, cit., 633 ss.

¹³⁶ Il Considerando 165 aggiunge che i requisiti impiegati andranno «adattati in funzione della finalità prevista dei

sistemi e del minor rischio connesso» e dovranno tener conto «delle soluzioni tecniche disponibili e delle migliori pratiche del settore».

¹³⁷ Ossia «*intervento e sorveglianza umani, robustezza tecnica e sicurezza, vita privata e governance dei dati, trasparenza, diversità, non discriminazione ed equità, benessere sociale e ambientale e responsabilità*», i quali devono costituire il riferimento valoriale e metodologico

Ancora, non solo il par. 2 (così come il Considerando 165) prevede una valutazione del conseguimento degli obiettivi attraverso i citati indicatori chiave di prestazione, ma sempre una lettura complessiva del Reg. AI fa emergere che la Commissione ne deve stimare l'efficacia entro il 2 agosto 2028 e successivamente con cadenza triennale, rientrando i Codici di condotta tra le materie per le quali l'AI Act prevede la periodica valutazione e il riesame a mente dell'art. 112.

In ultimo, è da segnalare che anche l'art. 101 del Reg. IA contiene un riferimento ai «codici di condotta», ma che esso deriva da una impropria – quanto atecnica – traduzione italiana del termine riportato dal Regolamento stesso in una delle lingue originali usate in sede di redazione, come altre volte accaduto (valga per tutti l'eclatante caso della Direttiva 93/13) e come mette in evidenza l'area del diritto comparato che si occupa dei problemi di “traduttologia”, come si è più approfonditamente rilevato in altra sede e il quale aiuta paradossalmente a mettere in evidenza la diversità dei due istituti¹³⁸.

In effetti, a differenza dei Codici di condotta i Codici di buone pratiche sono volti a favorire l'attuazione, da parte dei fornitori di IA per finalità generali, delle norme contenute nel Reg. IA con particolare riguardo a quelle riguardanti gli obblighi di trasparenza, responsabilità e mitigazione del rischio per i fornitori di tali modelli, anche in considerazione degli approcci internazionali.

posto alla base dell'elaborazione degli stessi¹³⁷: il che mette in esponente il legame dei codici di condotta con le finalità generali del Reg. AI, posto che secondo il Considerando 165 «[l]o sviluppo di sistemi di IA diversi dai sistemi di IA ad alto rischio in conformità dei requisiti del presente regolamento può portare a una più ampia adozione nell'Unione di un'IA etica e affidabile»

¹³⁸ Su tale profilo sempre TASSONE e MANNI, *Codici di condotta*, cit., 633 ss. In proposito, l'art. 101, rubricato «[s]anzioni pecuniarie per i fornitori di modelli di IA per finalità generali», tratta degli «impegni assunti [...] nei pertinenti codici di condotta in conformità dell'articolo 56», ma questo è relativo ai «[c]odici di buone pratiche», secondo quanto già rende evidente la sua rubrica, ossia i codici la cui elaborazione è agevolata e incoraggiata dall'Ufficio per l'IA «al fine di contribuire alla corretta applicazione del presente regolamento». Ma non si tarda qui a scoprire che nella versione inglese l'art. 56 parla di «*codes of practice*» esattamente come fa l'art. 101 Reg. IA

Inoltre, così come i Codici di condotta, essi non sono vincolanti, ma a differenza di questi ultimi i fornitori che non intendono aderire ad essi dovranno dimostrare «mezzi alternativi adeguati di conformità ai fini di approvazione da parte della Commissione» (art. 53 par. 4 Reg. IA): secondo un meccanismo che appare adiacente a quello del *comply or explain* utilizzato dal legislatore unionale in altri contesti normativi e sufficiente a far ricadere i Codici di buone pratiche in generale e il Codice di buone pratiche di recente approvato dalla Commissione europea il 10 luglio 2025 (prima del quale se ne poteva discorrere al plurale, come testé fatto nel commentare le norme che li contemplano) nel novero degli strumenti di co-regolazione, invece che di auto-regolazione¹³⁹.

Ma lasciando da parte qui un ulteriore approfondimento sul Codice di buone pratiche steso a seguito del processo iterativo, inclusivo e trasparente di elaborazione, svoltosi durante sessioni ospitate dall'Ufficio europeo per l'Intelligenza Artificiale e alle quali hanno partecipato anche numerosissimi portatori di interessi rilevanti e il Comitato per l'IA, è perlomeno discusso che un'analoga soluzione ricostruttiva, cioè quella che li intende come strumenti di co-regolazione invece che di auto-regolazione, valga anche per i Codici di condotta¹⁴⁰: il che ci conduce alle considerazioni che seguono.

della stessa, il quale discorre di «commitments [...] made in relevant codes of practice in accordance with article 56».

¹³⁹ Il principio “*comply or explain*”, nato nel diritto anglosassone, è stato introdotto nel diritto euro-unitario dall'art. 20, par. 1, lett. b), Dir. 2013/34/UE, il quale prevede che: «*quando un'impresa, a norma del diritto nazionale, si discosta dal codice di governo societario di cui alla lettera a), punti i) o ii), essa rende noto da quali parti del codice di governo societario si discosta e i motivi di tale scelta; qualora l'impresa abbia deciso di non fare riferimento ad alcuna disposizione di un codice di governo societario di cui alla lettera a), punti i) o ii), ne spiega i relativi motivi*». Astratto dall'ambito societario esso implica che la scelta di discostarsi da codici disciplinari o di condotta, la cui adesione è volontaria, sia debitamente spiegata dall'impresa che prende questa decisione.

¹⁴⁰ Benché ambedue strumenti ad applicazione – più o meno volontaria – volti a rendere l'IA più sicura e

12. Co-regolazione, autoregolazione e altre riflessioni sul ruolo della autonomia privata

Un ulteriore tema che i Codici di condotta pongono, come si è appena detto, è quello della loro “qualificazione” rispetto alle varie forme che caratterizzano il rapporto fra la regolazione pubblica e quella privata anche in seno all’ordinamento euro-unitario.

Del resto, lo stesso Reg. IA ricorre a tre diverse modalità regolatorie, prevedendo (i) una parte di regole c.d. *top-down*, proprie dell’*hard law*, per proibire o regolamentare l’uso di determinati sistemi di intelligenza artificiale, (ii) una parte di misure “co-regolatorie” (*co-regulation*) e (iii) una parte di misure “auto-regolatorie” (*self-regulation*)¹⁴¹.

Al riguardo, una prima lettura delle disposizioni inerenti ai Codici di condotta porta a qualificarli come appartenenti a quest’ultima categoria, dacché nel Regolamento è quasi “insistito” il riferimento – fin dal *nomen iuris* dell’istituto e come chiarisce la rubrica stessa dell’articolo che li disciplina – al fatto che la loro adozione è del tutto «volontaria»¹⁴².

D’altronde, non è e non sarebbe certo la prima volta che il diritto euro-unitario adotta strumenti di *soft law*, come pure dimostra in vari altri settori l’esperienza dei Codici di comportamento, secondo quanto rammenta la dottrina specialistica che si è occupata del tema¹⁴³.

affidabile, i Codici di condotta e i Codici di buone pratiche si distinguono sotto molteplici profili, *in primis* quello dei soggetti chiamati ad applicarli. Difatti, mentre i Codici di condotta sono destinati ad essere applicati dai fornitori e *deployer* di sistemi di IA diversi da quelli ad alto rischio (per l’applicazione volontaria dei requisiti previsti per questi ultimi) e, in generale, ove si tratti di applicare requisiti ulteriori (ai sensi dell’art. 95, par. 2), i Codici di buone pratiche sono destinati ai modelli di IA per finalità generali, allo scopo di contribuire alla corretta attuazione degli obblighi per essi previsti dal regolamento, in specie, l’attuazione degli obblighi vincolanti ex art. 53 e 56. Già solo tale diversità di obiettivi e destinatari risulta sufficiente a distinguere i codici *de quibus*. Tuttavia, si deve ritenere che il principale *discrimen* tra i due risieda nell’art. 56 par. 6, laddove è previsto che la Commissione possa «approvare, mediante atto di esecuzione, un codice di buone pratiche e conferire ad esso una validità generale all’interno dell’unione». Sicché, da “codici di buone pratiche”, ove l’uso del plurale è da intendersi come

In proposito, i Codici di comportamento – del cui *genus* i Codici di condotta di cui all’art. 95 Reg. IA sono una *species* o ai quali sono omologabili, stante il carattere sinonimico delle parole “comportamento” e “condotta” – sono non di rado impiegati dal legislatore come parte integrante di un approccio di tipo “co-regolatorio”¹⁴⁴.

Di contro, rispetto all’elaborazione dei Codici di condotta ex art. 95 Reg. IA la partecipazione di soggetti istituzionali si risolve come si è detto in una mera «agevolazione» ed essa non vale certo ad infondere loro una forza normativa c.d. “dall’alto” atta ad attrarli nel fenomeno co-regolatorio in senso proprio diversamente da quanto afferma parte della dottrina facendo leva sulla semplice partecipazione dei portatori di interessi alla loro stesura e sul collegamento di tale partecipazione con la sfera pubblica assicurato dalla procedimentalizzazione della stessa¹⁴⁵.

Infine e – in un’ottica funzionale – soprattutto, non sono in via espressa previste sanzioni o conseguenze giuridiche in caso di mancata applicazione degli stessi, né incentivi per la loro adozione, sicché li si potrebbe qualificare come forme di regolazione completamente auto-imposte e funzionali a promuovere un comportamento conforme ai principi delineati dal Regolamento, sebbene in assenza di alcun

riferito alla categoria in generale, si giunge a parlare di un singolo codice, quello approvato dalla Commissione lo scorso 10 luglio 2025. Per ulteriori riflessioni sui codici di buone pratiche, P. D’ORAZIO, ART. 56, in MANTELERO, RESTA, RICCIO, *Intelligenza artificiale*, cit., 519 ss., nonché, più di recente, TASSONE e MANNI, *Codici di condotta*, cit., 633 ss.

¹⁴¹ U. PAGALLO, *LLMs meet the AI Act: Who’s the sorcerer’s apprentice?*, in *Cambridge Forum on AI: Law and Governance*, 2025, 1, e1, 1-13, disponibile presso www.cambridge.org.

¹⁴² L’art. 95 Reg. IA è rubricato «[a]pplicazione volontaria dei codici di condotta» ed essi sono come tali richiamati nel resto del Regolamento al netto dell’errore di traduzione di cui si è detto nel paragrafo 11.

¹⁴³ P. INTURRI, *Intelligenza artificiale e soft law. Il ruolo dei codici di comportamento nell’Artificial Intelligence Act*, in *Nuove Autonomie*, 2025, 481 ss.

¹⁴⁴ *Ibidem*, 482.

¹⁴⁵ P. D’ORAZIO, ART. 95, in MANTELERO, RESTA, RICCIO, *Intelligenza artificiale*, cit., 830-831.

obbligo normativo cogente, il che ne potrebbe pregiudicare la diffusione e l'efficacia.

Anzi, i descritti caratteri dei Codici di condotta fanno sì che essi possano persino diventare un'arma a doppio taglio, nel senso che le clausole negli stessi contenute, pur appearing formalmente allineate ai requisiti previsti dalla sezione 2 del capo III o agli altri requisiti fissati dal Regolamento, potrebbero risultare eccessivamente generiche o vaghe, ammantando di una parvenza di eticità pratiche che, nella sostanza, non soddisfano i criteri di affidabilità e responsabilità realmente propri dell'AI Act¹⁴⁶.

Ciò rischierebbe dunque di rendere più complessa l'attività di *governance* e controllo sui sistemi di IA, che si presenterebbero come maggiormente conformi ai valori etici di quanto non lo siano nella realtà, il che non deve essere sottovalutato rispetto ad alcun "tipo" di sistema di intelligenza artificiale secondo la tassonomia dello stesso AI Act, dacché anche sistemi classificati come "a basso rischio" possono produrre effetti significativi sul contesto sociale e umano in cui vengono integrati¹⁴⁷.

Inoltre, pur a fronte del coinvolgimento del Consiglio per l'IA, dell'Ufficio per l'IA e degli Stati membri nel processo di elaborazione dei Codici di condotta, non può essere trascurato il fatto che anche in tale contesto specie le c.d. *big tech* possono avere un peso determinante e far sì che

tali strumenti riflettano in misura prevalente gli interessi del settore privato, risultando fortemente orientati a favore degli operatori economici, piuttosto che configurarsi come reali strumenti di garanzia a tutela degli interessi pubblici o comunque superindividuali.

In ultimo, non va sottaciuto che il carattere volontario dei Codici di condotta vuole essere, nell'ottica del legislatore euro-unitario, uno dei punti di forza dell'istituto, secondo quanto altresì emerge dal confronto fra la versione originaria della Proposta di Regolamento e il successivo emendamento del 2022¹⁴⁸.

In tale ottica, la scelta di rendere del tutto volontaria – e potenzialmente parziale – l'applicazione dei requisiti è da intendersi come una strategia volta a favorirne la diffusione su larga scala e nella medesima direzione va il lessico impiegato dal Regolamento, che spesso usa in modo non casuale i citati termini «agevolare» e «incoraggiare».

D'altronde, non è un caso che ad oggi si registrano: una reazione negativa di talune imprese rispetto al Codice di buone pratiche ufficialmente approvato dalla Commissione¹⁴⁹; una certa resistenza in generale nell'adattarsi alle "imposizioni" del Reg. IA¹⁵⁰; e la richiesta di più tempo per potervi procedere oltre che di una semplificazione della normativa espressa dalla EU AI Champions Initiative con la lettera aperta all'UE "stop the clock"¹⁵¹.

¹⁴⁶ U. PAGALLO, *On Twelve Shades of Green: Assessing the Levels of Environmental Protection in the Artificial Intelligence Act*, in *Minds and Machines*, 2025, 12, disponibile presso www.springer.com.

¹⁴⁷ PAGALLO, *LLMs meet the AI Act*, cit., 7.

¹⁴⁸ In particolare, P. BELOTTI, *I codici di condotta di cui alla Proposta di Regolamento UE sull'Intelligenza Artificiale, ipotesi per un'applicazione dello Human Rights-Based Approach*, in *Rivista di Scienze Giuridiche, Scienze Cognitive ed Intelligenza Artificiale*, 2022, 51 ss., disponibile presso www.i-lex.unibo.it. Nella specie, in tale passaggio si registrano interventi che sembrano riflettere un cambio di prospettiva, la quale consente l'applicazione selettiva di alcuni requisiti previsti per i sistemi di IA ad alto rischio senza imporre l'adozione integrale delle relative disposizioni, per cui è «come se il legislatore europeo abbia deciso di imporre "meno" in modo da facilitare l'adozione di prassi virtuose ad una più ampia parte del tessuto economico»

¹⁴⁹ Benché molti *tech giants* lo abbiano firmato (*inter alia*, Amazon, Anthropic, Google, IBM, Microsoft e OpenAI), tra i grandi assenti spicca Meta.

¹⁵⁰ Un'indagine condotta da Deloitte sull'impatto del Regolamento sull'intelligenza artificiale nel settembre 2024 ha attestato che il 48,6% delle aziende intervistate non aveva ancora avviato concrete iniziative di adeguamento o implementazione delle disposizioni del Regolamento. Tale dato, letto in correlazione con la convinzione, del 52,3% degli intervistati, che le loro capacità di innovazione in tema di IA sarebbero state ristrette dal Regolamento segnala, più che semplice impreparazione, un atteggiamento di manifesta ritrosia – se non di aperto rifiuto – verso la normativa europea sull'IA. Così DELOITTE, si veda *AI Act Survey 2024 – Survey on the Impact of the AI Act*, September 2024, 6, disponibile presso www.deloittelegal.de.

¹⁵¹ Premesso che la *EU AI Champions Initiative* è un movimento rappresentato da oltre 110 organizzazioni con lo scopo di accelerare l'adozione dell'intelligenza

In definitiva, è ancora presto per dire se e fino a che punto l'approccio del tutto *liberal* del legislatore euro-unitario condurrà al conseguimento degli obiettivi che lo stesso si è prefissato di raggiungere, mentre le controindicazioni sopra schematizzate non paiono di poco conto.

Del resto, una indagine empirica in altra sede compiuta dimostra che fra i vari – principali – *player* mondiali che hanno adottato, con contenuti assai diversi dei codici di condotta (che indichiamo con la lettera minuscola per le ragioni fra un attimo esplicitate), nessuno li ha espressamente adottati come Codici di condotta ai sensi dell'art. 95 del Reg. IA¹⁵².

Ciò sembra in effetti un segnale di chiara diffidenza rispetto al descritto approccio del legislatore europeo: che da un lato «agevola» l'adozione dei codici e non poche parole spende sui loro contenuti, ma dall'altro mai dice – espressamente – quali sono le conseguenze giuridiche, positive o negative, che derivano dalla loro adozione, sebbene già fossero disponibili tecniche ulteriori rispetto a quelle proprie della co-regolazione o della auto-regolazione per dare rilievo ai codici di condotta. Qui, allora, emerge ancora una volta il ruolo dell'interprete, essendo dato di configurare una ipotesi ricostruttiva non solo solida, ma idonea a

valorizzare i *pro* di entrambe le forme di regolazione.

In specie, si è ipotizzato che qui si abbia un modello “misto” di co-regolazione e auto-regolazione, caratterizzato dal fatto che il legislatore definisce (e in tal senso co-regola) a monte il piano della fattispecie, senza prevedere un *autonomo effetto* a valle, come quello che solitamente deriva dalla “giuridificazione” pubblica della regola di origine privata¹⁵³: il che ci conduce immediatamente al tema delle sanzioni e della responsabilità civile¹⁵⁴.

In effetti, l'esame delle disposizioni in tema di sanzioni e alle altre misure di esecuzione del Reg. IA – che meriterebbe altro spazio – svela subito che i codici di condotta non sono mai menzionati. Anzi, a veder bene, una volta che la versione italiana dell'art. 101 è stata emendata dall'errore di traduzione sopra evidenziato, essa svela che a venire citati sono solo e soltanto i «*codici di buone pratiche*», il che potrebbe intendersi come una chiara presa di posizione del legislatore euro-unitario circa il fatto che i codici di condotta *non* rilevano punto rispetto all'applicazione degli artt. 99-101¹⁵⁵.

Di contro, proprio l'ipotesi ricostruttiva testé avanzata, ossia quella della regolazione “mista”, può indurre l'interprete a chiedersi se non sia più corretta un'opposta conclusione.

artificiale e rafforzare la competitività dell'Europa in materia, attraverso la lettera aperta “Stop the Clock” alcuni attori del settore IA hanno formalmente richiesto all'Unione Europea un rinvio delle scadenze previste, sostenendo che l'industria europea non fosse ancora pronta a conformarsi pienamente ai nuovi requisiti. I firmatari hanno evidenziato come, nel regolare l'IA, il tradizionale equilibrio europeo tra promozione dell'innovazione e regolamentazione si sia ormai sbilanciato eccessivamente a favore di quest'ultima, con il rischio di compromettere le ambizioni strategiche dell'UE nel settore, ostacolando lo sviluppo delle imprese europee e limitando la loro capacità di progettare e implementare sistemi di IA competitivi a livello globale. Per ulteriori riferimenti TASSONE - MANNI, *AI ACT: Stop the clock alle prossime scadenze*, in *Diritto di internet – news*, 1° agosto, 2025, disponibile all'indirizzo www.dirittodiinternet.it.

¹⁵² Al riguardo si vedano ancora TASSONE e MANNI, *Codici di condotta*, cit., 633 ss., i quali muovono dal Codice di condotta internazionale elaborato dai Paesi del G7 in seno all'*Hiroshima Process* conclusosi nell'ottobre 2023, per

analizzare il “Codice di Condotta per i Servizi di Intelligenza Artificiale di Microsoft” (che stabilisce linee guida per l'uso responsabile delle tecnologie IA offerte da *Microsoft* e mira a garantire che le applicazioni sviluppate siano sicure, trasparenti e rispettose dei diritti umani), il “*Supplier Code of Conduct* di *OpenAI*”, che fa invece emergere un approccio differente, meno compatibile con l'applicazione volontaria di requisiti specifici non vincolanti, nonché il “trust center” di Anthropic¹⁵², presso il quale sono resi disponibili, in un unico “centro” virtuale appunto, i documenti di *compliance*, gli standard ai quali Anthropic si conforma, e altri documenti utili, tra cui codici di condotta per i *supplier* e per gli utenti, nonché report e dettagli sui controlli applicati.

¹⁵³ TASSONE e MANNI, *Codici di condotta*, cit., 633 ss.

¹⁵⁴ Evidenziano tale condivisibile aspetto SEPE e URBANI, *I codici di condotta*, cit., 721 ss.

¹⁵⁵ Su tali norme, per tutti, F. PELUSO e C. MATTEONI, *Procedure sanzionatorie*, in CASSANO e TRIPODI, *Il regolamento europeo sull'intelligenza artificiale*, cit., 889 ss.

Al riguardo, il par. 1 dell'art. 99 vuole che, entro la cornice del Regolamento, gli Stati «stabiliscono le norme relative alle sanzioni e alle altre misure di esecuzione [...] applicabili in caso di violazione del presente regolamento da parte degli operatori, e adottano tutte le misure necessarie per garantirne un'attuazione corretta ed efficace, tenendo conto degli orientamenti emanati dalla Commissione a norma dell'articolo 96»: sicché quel rilievo espresso che il Regolamento non dà ai codici di condotta potrebbe ben emergere in sede, appunto, di attuazione dello stesso da parte degli Stati membri.

Del resto, nella direzione per cui non si tratta di una facoltà ma un obbligo degli Stati membri o addirittura in quella per cui *già* il Regolamento impone all'interprete di tenerne conto va il richiamo agli «orientamenti» di cui all'art. 96: che assieme al precedente art. 95 compone il capo rubricato «codici di condotta e orientamenti» e, soprattutto, il quale vuole che ai principi espressi dagli stessi «orientamenti» i codici di condotta si ispirino. Sarebbe dunque assai singolare che un principio etico rilevi ai fini della irrogazione delle sanzioni e delle altre misure di attuazione del Regolamento solo nella sua dimensione astratta stabilita dalla norma e non anche in quella concreta alla quale il singolo operatore ha deciso di auto-vincolarsi.

E ciò per non considerare che sempre il par. 1 dell'art. 99 richiama il principio di proporzionalità – che davvero assume un ruolo cardine nell'intero ordinamento euro-unitario quando si discorre della irrogazione delle più varie sanzioni (da quelle proprie della Direttiva Enforcement a quelle del DSA, tanto per fare degli esempi) – e la sua applicazione potrebbe ben indurre l'interprete a considerare in tale contesto la previa adozione di un codice di condotta.

Per giunta, il par. 7 dell'art. 99 chiarisce che «[n]el decidere se infliggere una sanzione amministrativa pecuniaria e nel determinarne l'importo in ogni singolo caso, si tiene conto di tutte le circostanze pertinenti della situazione specifica e, se del caso, si tiene in considerazione quanto segue: [...] e) eventuali altri fattori

aggravanti o attenuanti applicabili alle circostanze del caso, ad esempio i benefici finanziari conseguiti o le perdite evitate, direttamente o indirettamente, quale conseguenza della violazione». E poiché è evidente che l'elencazione delle «circostanze pertinenti della situazione specifica» non ha carattere tassativo, così come non lo hanno i «fattori [...] attenuanti applicabili alle circostanze del caso» costituenti una delle stesse esemplificazioni, vi è qui una doppia possibilità di dare rilievo ai Codici di condotta.

Di poi, premesso che la clausola generale relativa a «tutte le circostanze pertinenti della situazione specifica», a conferma del suo rilievo, si ritrova pure nell'art. 100 relativo alle «sanzioni amministrative pecuniarie inflitte a istituzioni, organi e organismi dell'Unione», il par. 1 dell'art. 101 dedicato alle «sanzioni pecuniarie per i fornitori di modelli di IA per finalità generali», nella sua parte finale, dispone similmente che «[n]el determinare l'importo della sanzione pecuniaria o della penalità di mora, si tengono in considerazione la natura, la gravità e la durata della violazione, tenendo debitamente conto dei principi di proporzionalità e di adeguatezza» e che «[l]a Commissione tiene conto anche degli impegni assunti in conformità dell'articolo 93, paragrafo 3, o assunti nei pertinenti codici di condotta in conformità dell'articolo 56».

Ebbene, di là dal fatto che, di nuovo, il principio di proporzionalità potrebbe di suo valere a dare rilievo ai codici di condotta e che qui la norma si riferisce in verità e come si è detto ai «codici di buone pratiche», solo i medesimi vengono qui menzionati per la semplice ragione che il loro contenuto minimo è, secondo la funzione ad essi assegnata dal Reg. AI, quello di dare attuazione agli obblighi di cui agli artt. 53 e 55 rivolti esclusivamente ai «fornitori di modelli di IA per finalità generali» (a rischio sistemico e no) di cui l'art. 101 si occupa.

Inoltre, secondo la ricostruzione proposta in altra sede gli strumenti di co-regolazione sono caratterizzati dalla previsione di un effetto tipico, che appunto l'art. 101 indica nella rilevanza in sede di irrogazione delle sanzioni; sicché la previsione di cui all'art. 101 non si deve ritenere

per così dire *ad excludendum* rispetto al rilievo dei codici di condotta, senza contare che un ulteriore sviluppo della riflessione porta a considerare altresì la responsabilità civile¹⁵⁶.

In effetti, così come vale per le sanzioni e le misure di attuazione, anche per la responsabilità civile non è dato di tracciare un quadro di massima in poche battute.

Limitando all'osso le considerazioni da svolgere facendo tesoro di quelle esposte altrove, un primo e generale rilievo dei Codici di condotta è ovvio: se un dato soggetto si è vincolato all'osservanza di determinate regole di comportamento che egli stesso ha scritto, contribuito a scrivere o fatte proprie, non è poi revocabile in dubbio che esse rientrano fra quelle rilevanti ai fini della individuazione degli standard di comportamento cui lo stesso si deve conformare ai sensi dell'art. 2043 c.c. e delle altre norme per le quali la colpa rileva¹⁵⁷.

In secondo luogo, poiché lo scopo dichiarato del Reg. IA è quello di "innalzare l'asticella" e far sì che mediante i Codici di condotta gli obblighi connessi all'uso di sistemi di intelligenza artificiale ad alto rischio siano estesi *anche* a sistemi di livello inferiore (secondo la ideale piramide cui sopra si accennava) seppur con gli adattamenti del caso, sembra potersi affermare che da ciò discende da un lato la possibilità di applicare il regime che il Reg. AI prevede per i primi, il quale riguarda il tema delle sanzioni e non della responsabilità civile; ma – dall'altro lato – un indice per rendere operativo uno standard di diligenza più stringente, se si vuole ricalcando la dinamica che porta ad un irrigidimento delle maglie quando si passa dalla colpa semplice alla colpa professionale secondo l'art. 1176 c.c.

Infine, proprio tali considerazioni sul rilievo dei Codici di condotta offrono il destro per una riflessione a compasso più allargato circa il rapporto fra l'AI Act e le norme di responsabilità nel loro complesso, della quale si possono qui delineare due soli – quanto pregnanti – aspetti. Intanto, se il carattere volontario delle regole di condotta al cui rispetto l'operatore si vincola in

forza dei codici di condotta incide sulla applicazione delle regole di responsabilità civile e in specie ai fini della definizione della colpa, una analoga conclusione va raggiunta rispetto a tutte le altre regole di condotta contenute nell'AI Act: che si devono postulare consonanti rispetto alle regole per così dire derivate e contenute negli stessi codici di condotta al primo "riconducibili" secondo lo schema, ancora una volta, della regolazione "mista". E – si badi – una tale conclusione è rilevante anche perché mentre un illecito penale è sempre un illecito (anche) civile secondo l'art. 185 c.p., analogo automatismo non vi è per gli illeciti di carattere amministrativo.

Del resto, la finalità di prevenire i danni è espressamente menzionata dai Considerando 9, 27, 29 e 155, mentre la stessa definizione di «incidente grave» di cui al Punto 49 dell'art. 3 vuole che «un incidente o malfunzionamento di un sistema di IA [sia quello] che, direttamente o indirettamente, causa una delle conseguenze seguenti: a) il decesso di una persona o gravi danni alla salute di una persona; [...] d) gravi danni alle cose o all'ambiente».

Di poi, un passo più in là, si colloca l'idea per cui la stessa qualificazione delle attività "ad alto rischio" secondo il Reg. AI non può non incidere sull'applicazione dell'art. 2050 c.c., che proprio sul livello di pericolosità di per sé legato al suo svolgimento basa, integralmente, la propria *ratio* (al punto che secondo la Relazione al Codice Civile lo stesso art. 2054 dettato in tema di circolazione di veicoli costituisce una mera epifania dell'art. 2050 proprio per il rischio di incidenti che reca con sé).

E in tal modo, peraltro, si dà piena consistenza ad un'anima inespressa del peculiare sistema costruito dal Regolamento rispetto ai Codici di condotta, cioè a dire quella che lo vede come applicazione del principio di sussidiarietà: il che, di necessità, richiama il ruolo dell'Agcom, al quale vale la pena di dedicare alcune considerazioni prima di venire alle battute finali.

13. Il ruolo dell'Agcom

¹⁵⁶ Si vedano sempre TASSONE e MANNI, *Codici di condotta*, cit., 633 ss.

¹⁵⁷ *Ibidem*.

All'Agcom compete un ruolo di primo ordine nella materia della tutela del diritto d'autore essendosi in tal senso più volte espresso lo stesso Legislatore nel corso degli anni, come emerge dalla più volte citata Legge 22 aprile 1941, n. 633, dal D.Lgs. 9 aprile 2003, n. 70, dal D.Lgs. 15 marzo 2017, n. 35 e dal D.Lgs. 8 novembre 2021, n. 208, oltre che da altri recenti interventi normativi¹⁵⁸.

Inoltre, il Regolamento del Parlamento Europeo relativo a un mercato unico dei servizi digitali - regolamento sui servizi digitali, o "Digital Services Act" (DSA) – vuole che "gli Stati membri designano una o più autorità competenti incaricate della vigilanza dei fornitori di servizi intermediari e dell'esecuzione del presente regolamento («autorità competenti»)" (art. 49)¹⁵⁹.

In proposito, proprio l'Autorità per le garanzie nelle comunicazioni è stata designata quale Coordinatore dei Servizi Digitali (art. 15 D.L. 15 settembre 2023, n. 123 - Misure urgenti di contrasto al disagio giovanile, alla povertà educativa e alla criminalità minorile, nonché per la sicurezza dei minori in ambito digitale -

convertito in legge, con modificazioni, dall'art. 1, l. 13 novembre 2023, n. 159).

Del resto, il ruolo che essa sta sempre più assumendo viene ben delineato da alcune riflessioni che sono state svolte sul rilievo sistematico – oltre che operativo – della già citata legge n. 93/2023, recante "[d]isposizioni per la prevenzione e la repressione della diffusione illecita di contenuti tutelati dal diritto d'autore mediante le reti di comunicazione elettronica"¹⁶⁰: la quale ha lo scopo di rafforzare la tutela del diritto d'autore *online* soprattutto (ma non solo) mediante l'attribuzione di nuovi poteri all'Autorità per le Garanzie nelle Comunicazioni (di seguito "AGCOM" o "Autorità")¹⁶¹.

Tale legge, infatti, incide non poco sulle competenze dell'Agcom a fronte di una storia decisamente travagliata proprio rispetto al settore che qui interessa.

In proposito, la Legge n. 248/2000 aveva attribuito all'AGCOM specifiche funzioni di vigilanza e poteri ispettivi, da esercitare in coordinamento con la SIAE, novellando la L.d.a. mediante l'introduzione degli artt. 182-*bis* e 182-*ter*, mentre con gli artt. da 14 a 16 il D.Lgs.

¹⁵⁸ Da ultimo, senza pretese di completezza si registrano il D.Lgs. 8 novembre 2021, n. 177 (Attuazione della direttiva (UE) 2019/790 del Parlamento europeo e del Consiglio, del 17 aprile 2019, sul diritto d'autore e sui diritti connessi nel mercato unico digitale) secondo cui "decorsi due anni dalla data di entrata in vigore del presente decreto, l'Autorità per le garanzie nelle comunicazioni trasmette alle Camere una relazione, integrata da verifica d'impatto della regolazione, sull'applicazione di propria competenza delle disposizioni che vi sono contenute, con particolare riferimento ai criteri e alle modalità di determinazione dell'equo compenso per gli editori di pubblicazioni giornalistiche, di cui all'articolo 43-bis della legge 22 aprile 1941, n. 633, e alla procedura di determinazione dei compensi per gli autori, artisti, interpreti e esecutori, previo confronto con gli organismi di gestione collettiva o entità di gestione indipendenti di cui al decreto legislativo 15 marzo 2017 n. 35" (art. 2); nonché la Legge 14 luglio 2023, n. 93 (Disposizioni per la prevenzione e la repressione della diffusione illecita di contenuti tutelati dal diritto d'autore mediante le reti di comunicazione elettronica) secondo cui, tra l'altro, l'Agcom "con proprio provvedimento, ordina ai prestatori di servizi, compresi i prestatori di accesso alla rete, di disabilitare l'accesso a contenuti diffusi abusivamente mediante il blocco della risoluzione DNS dei nomi di dominio e il blocco

dell'instradamento del traffico di rete verso gli indirizzi IP prevalentemente destinati ad attività illecite" (comma I). E ancora "Con il provvedimento di cui al comma 1, l'Autorità ordina anche il blocco di ogni altro futuro nome di dominio, sottodominio, o indirizzo IP, a chiunque riconducibili, comprese le variazioni del nome o della semplice declinazione o estensione (cosiddetto top level domain), che consenta l'accesso ai medesimi contenuti diffusi abusivamente e a contenuti della stessa natura" (comma II).

¹⁵⁹ Per un commento a tale complesso atto normativo si vedano i contributi editi in G. CASSANO e B. TASSONE (a cura di), *L'impatto del Digital Services Act sulla legislazione italiana – Speciale ragionato*, numero monografico di *Dir. internet*, Pisa, 2023.

¹⁶⁰ B. TASSONE, *Primo commento alla L. 14 luglio 2023, n. 93 in materia di tutela del diritto d'autore online e alla Delibera 189/23/CONS dell'AGCOM che modifica il Regolamento sul diritto d'autore online*, in *Dir. ind.*, 2023, 415 ss.

¹⁶¹ Per una recente ricostruzione unitaria del diritto delle comunicazioni ove anche, ovviamente, la attenta individuazione dei ruoli e funzioni dell'AGCOM nel suo complesso, V. Franceschelli, *Diritto delle comunicazioni*, 2022.

n. 70/2003 - chirurgicamente analizzato sul punto nella lunga Delibera 680/13/CONS con cui l'Autorità ha approvato il citato Reg. AGCOM – affidava all'autorità amministrativa “avente funzioni di vigilanza” il potere di “esigere, anche in via d'urgenza, che il prestatore, nell'esercizio delle proprie attività come ivi definite, impedisca o ponga fine alle violazioni commesse, agendo immediatamente per rimuovere le informazioni illecite o per disabilitarne l'accesso”¹⁶².

Di poi, nel 2010 il D.Lgs. n. 44/2010 (meglio noto come “Decreto Romani”) ha inserito l'art. 32-*bis* nel D.Lgs. 177/2005 - ossia nel Testo unico della radiotelevisione, ora sostituito dal Testo Unico dei servizi di media audiovisivi e radiofonici di cui al D.Lgs. 208/2021 - il quale è stato pure considerato idoneo a munire di base giuridica i poteri che l'Autorità ha esercitato al fine di emanare il citato Reg. AGCOM.

Così, a seguito di tre diverse consultazioni pubbliche l'AGCOM lo ha adottato nel 2013 e sulla sua legittimità si è subito sviluppato un ampio dibattito¹⁶³: se non è mai stata in discussione la competenza dell'AGCOM di dettare disposizioni regolamentari per garantire il rispetto del diritto d'autore da parte dei

fornitori dei servizi di media audiovisivi e radiofonici, maggiori perplessità aveva creato la “auto-attribuzione” del potere di farlo per le comunicazioni elettroniche, in via ermeneutica, sulla scorta del citato art. 182-*bis* L.d.a. e, ancora, del D. Lgs. 259/2003, ossia il Codice delle comunicazioni elettroniche, nonché - più in generale - sulla scorta della L. n. 481/1995 (recante “[n]orme per la concorrenza e la regolazione dei servizi di pubblica utilità”) e della L. n. 249/1997 (che, come detto, ha istituito l'Autorità)¹⁶⁴.

Tuttavia, con la nota decisione del 3 dicembre 2015, n. 247, la Corte costituzionale ha dichiarato inammissibili le questioni di legittimità costituzionale sollevate con riguardo al Reg. AGCOM e la giurisprudenza amministrativa si è poi allineata a tale posizione¹⁶⁵.

Inoltre, un riconoscimento almeno indiretto e parziale della legittimità del Reg. AGCOM è derivato dall'art. 2 della L. n. 167/2017 che, nel dare attuazione all'art. 8 della Dir. Infosoc, nonché agli artt. 3 e 9 della Dir. Enforcement, ha munito l'AGCOM del potere di adottare uno specifico regolamento attuativo ai fini della disciplina del procedimento cautelare

¹⁶² Le disposizioni richiamate nel testo, al pari dell'art. 17 dello stesso decreto, sono state abrogate dal D.Lgs. 25 marzo 2024, n. 50. L'abrogazione segue alla soppressione degli artt. 12-15 della direttiva 2000/31/CE disposta dall'art. 89 del Reg. (UE) 2022/2065 (DSA), i cui artt. 4, 5, 6 e 8 ne hanno assunto la funzione. La possibilità che un'autorità giudiziaria o amministrativa, secondo l'ordinamento dello Stato membro, esiga dal prestatore di impedire o porre fine a una violazione è oggi fatta salva dagli artt. 4, par. 3, 5, par. 2, e 6, par. 4, del DSA. Gli ordini di contrastare contenuti illegali sono invece disciplinati dall'art. 9.

¹⁶³ In senso contrario alla legittimità del Reg. AGCOM, fra i tanti, G. De Minico, *Libertà e copyright nella Costituzione e nel diritto dell'Unione*, in *Ass. it. cost.*, 2014, 1, e M. Renna, *Le questioni di legittimità del regolamento dell'Agcom sulla tutela del diritto d'autore* on line, in *Ann. it. dir. aut.*, 2014, 111 ss. Di contro e in senso favorevole, sempre senza pretese di completezza, A. Stazi, *Il regolamento di cui alla delibera n. 680/13/CONS dell'Agcom per la tutela del diritto d'autore in rete*, in *Riv. dir. ind.*, 2014, 2, 13 ss., nonché Id., *La tutela del diritto d'autore in rete: bilanciamento degli interessi, opzioni regolatorie europee e “modello italiano”*, in *Dir. inf.*, 2015,

89 ss. In chiave comparatistica, G. Giannone Codiglione, *“Copyright” e diritto d'autore in rete: “bilanciamento interno” e “conflitto esterno” tra diritti fondamentali*, in *Riv. crit. dir. priv.*, 2014, 421 ss.

¹⁶⁴ A tale tema è stata prontamente dedicata una parte dei contributi monografici pubblicati nella Parte I del XXII Volume di *Ann. it. dir. aut.* (Milano, 2013), ad alcuni dei quali si farà in seguito riferimento.

¹⁶⁵ Secondo la decisione della Consulta (fra le altre in *Foro it.*, 2016, I, 405) “[s]ono inammissibili, per contraddittorietà, ambiguità e oscurità della motivazione, le questioni di legittimità degli artt. 5, 1° comma, 14, 3° comma, 15, 2° comma e 16, 3° comma, d.leg. 9 aprile 2003, n. 70 e 32 bis, 3° comma, d.leg. 31 luglio 2005, n. 177 (testo unico dei servizi di media audiovisivi e radiofonici), che attribuiscono all'autorità per le garanzie nelle comunicazioni il potere di adottare disposizioni regolamentari volte al rispetto della proprietà intellettuale degli operatori sulla rete di comunicazione elettronica, con cui, tra l'altro, intervenire anche in via d'urgenza sulle attività di trasporto e memorizzazione di informazioni”. Sulla successiva giurisprudenza amministrativa, Di Giorgi - Mula, *La prova nei procedimenti di fronte all'AGCOM*, cit., 281 ss.

concernente i provvedimenti emanati a tutela del diritto d'autore¹⁶⁶.

Premesso che infine il citato D.Lgs. n. 177/2021 ha confermato che in materia l'AGCOM è ormai un punto di riferimento e che alcune delle competenze da esso attribuite all'Autorità hanno anche valenza sanzionatoria, sul piano generale la L. n. 93/2023 è intervenuta in argomento sotto due profili, che qui vale la pena di ricordare assai brevemente¹⁶⁷.

In primo luogo, i dibattiti inerenti alla legittimazione della AGCOM a emanare disposizioni regolamentari a tutela del diritto d'autore nel settore delle comunicazioni elettroniche non possono che essere definitivamente sopiti in ragione sia delle competenze attribuite all'AGCOM, sia della disposizione che (seppur con refusi riscontrabili anche nella versione pubblicata sulla Gazzetta Ufficiale) le impone di modificare nel suo complesso il Reg. AGCOM (già modificato con le Delibere n. 490/18/CONS, n. 295/20/CONS e n.

233/21/CONS) la cui legittimità è dunque data ulteriormente per presupposta¹⁶⁸.

In secondo luogo, ha previsto misure – quelle di cui all'art. 2 – da “calare” nel tessuto normativo facente capo al Reg. AGCOM del quale la L. n. 93/2023 richiede appunto l'adeguamento¹⁶⁹.

E ciò, a veder bene, è indicativo del ruolo che l'AGCOM è venuta ad assumere – e che viene testimoniato dagli ulteriori sviluppi (fra luce e ombre) inerenti al *Piracy Shield* e alle stesse modifiche alla legge di cui si tratta, anche se sull'uno e sugli altri non ci si può qui soffermare oltre, occorrendo ora spiegare qualche prima riflessione conclusiva¹⁷⁰.

14. Riflessioni conclusive attorno al bilanciamento di interessi

L'ingiunzione dinamica amministrativa prevista dalla Legge n. 93/2023 comporta un passaggio dal *notice and take-down* al *notice and stay down* che rende complessa l'applicazione del principio appena evocato¹⁷¹: anche perché fin

¹⁶⁶ In base alla disposizione di cui nel testo l'AGCOM “su istanza dei titolari dei diritti, può ordinare in via cautelare ai prestatori di servizi della società dell'informazione di porre fine immediatamente alle violazioni del diritto d'autore e dei diritti connessi, qualora le violazioni medesime risultino manifeste sulla base di un sommario apprezzamento dei fatti e sussista la minaccia di un pregiudizio imminente, e irreparabile per i titolari dei diritti”. Aggiunge il comma 2 che “[l]'Autorità disciplina con proprio regolamento le modalità con le quali il provvedimento cautelare di cui al comma 1 è adottato e comunicato ai soggetti interessati, nonché i soggetti legittimati a proporre reclamo avverso il provvedimento medesimo, i termini entro quali il reclamo deve essere proposto e la procedura attraverso la quale è adottata la decisione definitiva dell'Autorità”.

¹⁶⁷ Ciò accade, in specie, ove vengano violati gli obblighi di trasparenza di cui ai commi 1, 2 e 3 dell'art. 110-*quater* (imposti ai “soggetti ai quali sono stati concessi in licenza o trasferiti i diritti e i loro aventi causa hanno l'obbligo di fornire agli autori e artisti interpreti e esecutori, anche tramite gli organismi di gestione collettiva e le entità di gestione indipendenti di cui al decreto legislativo 15 marzo 2017 n. 35”) secondo quanto dispone il comma 4 (perché “[s]ull'adempimento degli obblighi di comunicazione e di informazione cui ai commi 1, 2 e 3, vigila l'Autorità per le garanzie nelle comunicazioni che, in caso di violazione di tali obblighi, applica una sanzione amministrativa pecuniaria a carico del soggetto inadempiente fino all'1 per cento del fatturato realizzato

nell'ultimo esercizio chiuso anteriormente alla notifica della contestazione”).

¹⁶⁸ Secondo il comma 1 dell'art. 6 della legge in commento “[e]ntro sessanta giorni dalla data di entrata in vigore della presente legge, l'Autorità provvede, nel rispetto delle disposizioni della legge 7 agosto 1990, n. 241, a modificare il regolamento in materia di tutela del diritto d'autore sulle reti di comunicazione elettronica, di cui alla deliberazione della tata di 10 unità, di cui 1 unità di livello di-medesima Autorità n. 680/13/CONS del 12 dicembre 2013, al fine di adeguarlo alle disposizioni di cui alla presente legge”.

¹⁶⁹ Questo, in particolare, prevede due tipi di azioni complementari. In primo luogo, il sostegno allo sviluppo del mercato dei contenuti mediante la predisposizione di misure dirette ad incentivare l'offerta legale di opere digitali e l'educazione del pubblico alla corretta fruizione delle stesse. In secondo luogo, l'uso di procedure volte all'accertamento e alla cessazione delle violazioni del diritto d'autore e dei diritti connessi, comunque realizzate, poste in essere sulle reti di comunicazione elettronica.

¹⁷⁰ Sulla interpretazione delle norme di cui nel testo, sempre TASSONE, *Primo commento alla L. 14 luglio 2023*, cit., 415 ss.

¹⁷¹ Sul rilievo del principio di proporzionalità in materia di IP, per tutti, R. Romano, *Il giusto equilibrio fra diritti di proprietà intellettuale e altri diritti fondamentali*, in *Ann. it. dir. aut.*, 2021, 88 ss.

quando si discorre di diritti su eventi sportivi è certo che si sta trattando di posizioni “minori” secondo la stessa L.d.a.¹⁷².

Quid iuris, ad esempio, ove l’ulteriore sito sul quale ha luogo la violazione contenga molte altre pagine con contenuti del tutto legittimi?

In proposito, la citata Delibera AGCOM 189/23/CONS adottata a seguito della Legge n. 93/2023 opportunamente richiede che «gli indirizzi IP identificativi dei ‘Main server’ e dei ‘Delivery Server’ dei siti in questione e i nomi di dominio a loro associati», abbiano «il carattere dell’univocità», cioè quello per cui «i dati tecnici comunicati ai destinatari dal titolare devono risultare univocamente destinati alla lesione dei diritti di proprietà intellettuale del titolare istante», mentre più “morbida” appare la previsione contenuta nella Raccomandazione UE contro la pirateria *online* degli eventi sportivi e di altri eventi in diretta del 4 maggio 2023, che fortemente incoraggia gli Stati Membri all’uso della ingiunzione dinamica¹⁷³.

Al riguardo, il fatto che l’ordine possa essere limitato ad un sottodominio contribuisce di certo a circoscriverne l’impatto anche perché, di contro, occorre evitare condotte opportunistiche che facciano leva su “contenuti-scudo” per impedirne la pronuncia.

Tuttavia, non è detto che tale possibilità di delimitazione del perimetro dell’ordine

scongiuri a sufficienza il rischio di *overblocking* ove un sito o una sua sezione abbiano contenuti “misti” oppure ove uno stesso indirizzo IP bloccato ospiti più siti assai diversi fra loro.

In proposito, il Regolamento AGCOM del 2013 (in materia di tutela del diritto d’autore sulle reti di comunicazione elettronica e ai fini della gestione della procedura di notifica delle violazioni del diritto d’autore meglio nota come *notice and take down*) non ha accolto nonostante le varie novelle che lo hanno interessato il suggerimento emerso durante la consultazione che ha portato alla citata Delibera AGCOM (e caldeggiato dalla citata Raccomandazione) di limitare temporalmente l’ordine in presenza di eventi *live*, anche perché ciò avrebbe obbligato il titolare dei diritti a reiterare la richiesta ogni volta che tramite lo stesso sito fosse trasmesso un ulteriore evento in violazione dei suoi diritti, il che peraltro avrebbe favorito non poco i contraffattori¹⁷⁴: ma non è escluso che si possano trovare altre soluzioni che comunque rendono efficace un blocco temporaneo¹⁷⁵.

Per il resto, è comunque evidente che in certi contesti - come quello degli eventi sportivi - il richiamato principio del “doppio binario” è alquanto teorico, nel senso che la giurisprudenza civile non ha i mezzi per adottare (in tempi sufficientemente brevi) provvedimenti

¹⁷² Nel senso che quelli sulle manifestazioni sportive sono diritti che rientrano nel novero dei diritti connessi o comunque dei diritti *sui generis*, la cui copertura costituzionale è certamente inferiore, per tutti L.C. Ubertazzi, Commentario breve, cit., 1796 ss.

¹⁷³ Secondo il Punto 16 “Nel caso in cui l’ingiunzione assuma la forma di una misura di blocco, è opportuno assicurarsi che sia rivolta ai servizi pirata identificati dai siti internet dedicati principalmente a fornire accesso a ritrasmissioni non autorizzate o ad altri tipi di contenuti non autorizzati”.

¹⁷⁴ Secondo il Punto 19 “Gli Stati membri sono incoraggiati a prevedere la possibilità di chiedere un’ingiunzione nei confronti di un determinato prestatore di servizi intermediari, che possa essere estesa per consentire il blocco dei servizi pirata che effettuano ritrasmissioni non autorizzate di eventi sportivi in diretta, anche qualora tali servizi pirata non siano stati identificati al momento della richiesta di ingiunzione, purché si tratti dello stesso evento sportivo, in linea con le rispettive norme procedurali nazionali”.

¹⁷⁵ Una di esse potrebbe essere quella che lo parametrizza al campionato cui l’evento appartiene, così come non è escluso che l’Autorità si potrà muovere in tal senso - magari a seguito delle pronunce giurisdizionali sui suoi provvedimenti volti ad applicare l’ingiunzione dinamica amministrativa - visto che il “più contiene il meno” e che la giurisprudenza di altri Paesi già è andata in detta direzione, come emerge da *Football Association Premier League Limited v. British Telecommunications Plc & Others*, 2017, EWHC 480 (Ch), 13 marzo 2017, *Football Association Premier League Limited v. British Telecommunications Plc & Others*, 2017, EWHC 1877 (Ch) 5 luglio 2017 e *Football Association Premier League Limited v. British Telecommunications Plc & Others*, 2018, EWHC 1828 (Ch) 18 luglio 2018 (le successive rese sulla “proroga” dell’ordine reso con la prima). Sulla Giurisprudenza inglese, N. MALOVIC, *The Evolution of Copyright Website Blocking In the UK: Live Blocking Orders*, in *European Intellectual Property Review*, 2018, 810 ss.

idonei a fronteggiare le risorse della pirateria.

Al riguardo, viene però da domandarsi se il contributo dato dal formante giudiziale nei modi indicati non possa essere “ricambiato” in via di interpretazione sistematica, cioè se l’elaborazione delle regole che sovrintendono all’ingiunzione dinamica civile non possa giovare anche di quelle usate per la ingiunzione dinamica amministrativa.

Si allude, in specie, alla circostanza per cui un’attenta dottrina processualistica ha patrocinato un’applicazione della prima che considera essenziale la partecipazione al giudizio del contraffattore¹⁷⁶: secondo una soluzione che - senza qui esaminare l’impianto teorico che conduce a tale conclusione - certamente ha un effetto pratico assai limitante. Peraltro, proprio la menzione di tale ultima problematica dimostra che quello del bilanciamento di interessi è uno degli aspetti sui quali si scaricheranno le maggiori tensioni e si può aggiungere che queste verranno sempre più incrementate dalla stessa evoluzione tecnologica, come diviene evidente pensando - ad esempio - a provvedimenti che hanno ad oggetto concerti *live* trasmessi via Facebook oppure canzoni ascoltate in *streaming* tramite *play list* diffuse all’interno di gruppi di ascolto chiusi¹⁷⁷.

Del resto, tutti gli altri “contesti” presi in esame nel presente lavoro richiedono l’individuazione del migliore bilanciamento di interessi possibile, peraltro da intendersi in chiave dinamica e non statica.

Ad esempio, pur se – come già detto – saranno le successive elaborazioni giurisprudenziali e dottrinali a definire con maggiore precisione i criteri da usare per l’attribuzione della tutela autoriale a opere realizzate con l’ausilio dell’IA,

alla luce di quanto si è detto si stima ragionevole che a tale scopo siano tenuti presenti quelli che prendono in esame: (i) il grado di specificità e di intenzionalità delle istruzioni (*prompt*) fornite dall’autore; (ii) l’esistenza di un processo iterativo e selettivo di scelta consapevole tra output alternativi; (iii) gli interventi di modifica, integrazione o rielaborazione successivi alla generazione automatica di un output; (iv) la funzione creativa complessivamente svolta dall’autore nell’ideazione del progetto e nella determinazione del risultato finale; (v) l’eventuale integrazione dell’output generato in un contesto più ampio che a sua volta riflette scelte creative autonome (sequenza, montaggio, contestualizzazione).

E, in proposito, intuitive ragioni di certezza del diritto e parità di trattamento avrebbero forse consigliato una formula legislativa che demandasse ad autorità competenti – quali, ad esempio, l’UIBM o la SIAE, eventualmente di concerto con le autorità individuate ai sensi dell’art. 20 della Legge sull’IA – l’adozione di criteri e linee guida volti a colmare il vuoto semantico insito nel nuovo art. 1 L.D.A., sul quale ci si è sopra soffermati¹⁷⁸.

D’altronde, la scelta di campo del Legislatore italiano – laddove precisa che le opere protette sono il frutto del solo ingegno «umano» – ha luogo in un contesto euro-unitario evidentemente non armonizzato sul punto, perché non vi è alcuno strumento unionale che impone detta soluzione, com’è visto. E se la legge n. 132/2025 ha codificato la distinzione tra opere *AI-assisted* e *AI-generated* consacrando l’ingresso solo delle prime nelle Opere a pieno titolo, un giudice ceco ha negato con fermezza la tutelabilità autoriale e qualsiasi forma di diritto esclusivo sull’output in capo all’utente¹⁷⁹:

¹⁷⁶ M. STELLA, *Access Provider e tutela inibitoria*, in *Riv. dir. proc.*, 2023, 829 ss.

¹⁷⁷ Sulle interazioni fra metaverso e proprietà intellettuale in generale, C. GALLI, *Metaverso o metaversi tra innovazione tecnica e innovazione giuridica*, in *Dir. ind.*, 2023, 29 ss.

¹⁷⁸ Particolarmente virtuoso risulta l’esempio dello United States Copyright Office che, in data 29 gennaio 2025, ha pubblicato la parte 2 del suo *Report on Copyright and Artificial Intelligence* dal titolo «*Copyrightability*», con cui

ha affrontato la possibilità di conferire tutela autoriale agli output creati mediante GenAI, categorizzando diversi possibili tipi di interazione uomo-macchina e determinando, per ogni categoria individuata, se e in che termini l’output potesse essere considerato opera (sul tema UNITED STATES COPYRIGHT OFFICE, *Copyright and Artificial Intelligence Part 2: Copyrightability*, 2025, disponibile presso www.copyright.ai.gov).

¹⁷⁹ ČESKÁ REPUBLIKA ROZSUDEK JMÉNEM REPUBLIKY Č. J. 10 C 13/2023- 16, POINTS I-III.

mentre l'Ucraina – appena fuori dai confini dell'Unione – con la legge n. 2974-IX del 20 marzo 2023 ha emendato la disciplina del diritto d'autore per inserire una tutela *sui generis* per gli oggetti generati da computer che, pur riconosciuti come non originali, non risultano assimilabili ad altre opere¹⁸⁰.

Né, come si è anticipato, la conclusione è necessitata dall'elaborazione dei Giudici di Lussemburgo sulla figura dell'autore, seppur – giustamente – “antropocentrica”, così come dall'*acquis* comunitario in tema di brevetti.

Per giunta, qui va notato che limitazioni come quelle riscontrabili nel sistema italiano potrebbero portare in breve tempo alla creazione di molte opere oggettivamente analoghe a quelle protette grazie all'AI prive di ogni tutela e si potrebbe magari ritenere che una tale evenienza sia da salutare con favore. Basta però riflettere sugli effetti che una tale scelta potrebbe produrre nel medio o lungo periodo per smorzare facili ottimismo.

Invero, la creazione a basso costo di opere anch'esse a basso costo perché non protette potrebbe in poco tempo occupare spazi attualmente di pertinenza del mercato delle opere protette, andando in una prospettiva aggregata a discapito degli autori, perché molti utenti preferirebbero usare opere nate già in pubblico dominio rispetto a opere sulle quali i diritti patrimoniali si estinguono 70 anni dopo la morte dell'autore.

Inoltre, il proliferare di tali opere – magari stimolato dal fatto che esse sono gratuite – andrebbe a creare una massa sempre più rilevante di “anteriorità distruttive” la quale spiegherebbe i propri effetti nei rapporti e contenziosi fra i titolari dei diritti su nuove opere

protette, anche qui andando in definitiva a pregiudicare gli interessi degli autori.

Infine, dal punto di vista della tutela degli investimenti e della stessa creazione di applicativi sempre migliori, è in qualche modo paradossale, sempre con riferimento al “modello italiano”, che quanto più un *software* di IA utilizza processi interni non pienamente conoscibili né prevedibili, sì da essere maggiormente autonomo nella creazione, tanto più è probabile che esso non generi opere protette.

E che in tale contesto occorra una visione di insieme lo dimostra la prima Enciclica di Papa Leone XIV «*Magnifica Humanitas, sulla custodia della persona umana nel tempo dell'intelligenza artificiale*», la quale – fra le molte altre cose – si ispira alla Dottrina Sociale della Chiesa e dunque a valori come «la dignità inalienabile della persona, il bene comune, la destinazione universale dei beni, la sussidiarietà, la solidarietà e la giustizia sociale» non solo per addivenire ad una migliore regolamentazione dell'IA, ma anche per far fronte alla «concentrazione del potere economico nelle mani di pochi»¹⁸¹.

In ultimo, un passo più in là si colloca ovviamente un nodo di fondo toccato solo di scorcio nel presente lavoro e inerente a quali soggetti, in una diversa prospettiva, sarebbero poi titolari dei relativi diritti, ma con riserva di trattarne altrove, qui – piuttosto – va portata l'attenzione su un'ultima considerazione.

Le precedenti pagine dimostrano che l'IA – probabilmente da declinare al plurale – è un fenomeno di suo assai complesso e che occorre studiarlo in chiave multi-disciplinare, sicché invece che soffermarsi su questioni di massima, è più utile pensare ad un approccio

¹⁸⁰ Закон України № 2811-IX від 1.12.2022 Про авторське право і суміжні права (зведена версія від 15.04.2023 р.) Стаття 33 (Legge dell'Ucraina n. 2811-IX del 1° dicembre 2022 “Sul diritto d'autore e i diritti connessi” (versione consolidata al 15 aprile 2023, emendata con L. N. 2974-IX del 20 marzo 2023, articolo 33), traduzione non ufficiale di OPRYSK, disponibile in *Social Science Research Network*. In specie, il diritto *sui generis* è riconosciuto ai soggetti che detengono i diritti economici o dispongono di poteri di licenza relativi al programma informatico – autori

dei programmi per computer, loro eredi, persone alle quali gli autori o i loro eredi abbiano trasferito (alienato) i diritti economici sul programma informatico –, ovvero utilizzatori legittimi del programma stesso. Il titolare gode del diritto esclusivo di autorizzare terzi all'utilizzo dell'opera su base contrattuale, per l'intera durata di validità del diritto *sui generis*.

¹⁸¹ Per un primissimo commento, S. Di NOLA e A.M. Russo, *Magnifica Humanitas, il rapporto fra uomo e tech*, Milano Finanza, 25 maggio 2026, 78 ss.

metodologico peculiare rispetto ai temi e problemi che il medesimo tocca.

In sostanza, appare meglio domandarsi quali spazi abbia l'interprete rispetto a singoli momenti di disciplina ed esaminare gli stessi in modo per così dire "verticale" per poi espandere l'analisi, sempre per così dire in modo "orizzontale", agli altri plessi normativi e della umana conoscenza cui il singolo aspetto è correlato, recuperando solo alla fine il legame con riflessioni di più ampio respiro: secondo la linea che, auspicabilmente in modo efficace, si è scelto anche qui di seguire.

15. Postfazione (Rethinking Copyright: La proprietà intellettuale alla prova dei grandi modelli di IA)

1. Asimmetrie della conoscenza e derive feudali nel nuovo mercato della ricchezza immateriale – Osservava, qualche anno fa Jay Kesan, Professor of Law alla Illinois University che, per il giurista, il rapporto fra diritto e tecnologia somiglia ad una battuta di caccia, dal sapore entusiasmante ma anche a tratti inquietante e, comunque, dall'esito niente affatto scontato.

L'incertezza per il *legal hunter* (il giurista) – nonostante l'attrezzatura e gli sforzi profusi – si connette all'inanità di ogni sforzo volto a catturare una velocissima *scientific tiger*, la quale (a tutto concedere) sembrerebbe lasciarsi afferrare solo *by the tail*¹⁸². La metafora lancia un chiaro messaggio.

Le sollecitazioni sollevate dalla modernità impongono di ripensare con coraggio a categorie e limiti normativi esistenti, attualmente percepiti dall'interprete come dotati di un coefficiente di adeguamento che in realtà non

posseggono¹⁸³. La veloce corsa della tigre tecnologica restituisce allo specchio l'immagine dell'inadeguatezza dello strumentario del giurista e degli accaniti sforzi euristici volti a lasciare inalterato ciò che la tradizione normativa consegna, in applicazione (forse) dell'idea che "ciò che è vecchio è anche bello" e può continuare a funzionare, salvo qui e lì qualche necessario adattamento.

L'integrazione della tecnologia nei processi sociali ed economici ha, piaccia o meno, profondamente trasformato la funzione del diritto e i modelli di regolazione tradizionalmente offerti. Dinanzi alla rapidità di movimento della tecnologia (la tigre) – a meno di non accontentarsi, appunto, di prenderla per la coda – il giurista dovrebbe adottare un atteggiamento anticipatore e lungimirante, volto ad evidenziare in che modo la legge possa evolvere, cambiare o essere interpretata per abbracciare le nuove sfide sollevate dalla modernità¹⁸⁴. Come dire che se è vero che esigenza della regola è fare ordine e creare le condizioni per il futuro funzionamento del sistema, a partire da condotte oggettivamente analizzabili, è anche vero che pensare di introdurre regolarità e prevedibilità in uno spazio ormai governato da una frenetica evoluzione non solo compromette il successo della battuta di caccia ma segnala, all'opposto, che limitarsi ad afferrare la tigre per la coda rischia solo di provocarne la sua feroce reazione. L'inarrestabile fenomeno della documentazione in rete di tutte le attività della persona, nel ridefinire insomma le condizioni dell'esperienza umana e del suo rapporto con la tecnologia, impone un cambio nell'appello al diritto e alla dimensione regolatoria ad esso sottesa¹⁸⁵.

¹⁸² J. P. Kesan, *A tiger by the tail? The law contends with science and technology in America*, in *Scienza e diritto nel prisma del diritto comparato*, a cura di Comandè e Ponzanelli, Giappichelli, 2004, p. 387 ss.

¹⁸³ V. Zeno Zencovich, *Artificial intelligence, natural stupidity and other legal idiocies*, in *Media law*, 3, 2024, consultabile <https://www.medialaws.eu/artificial-intelligence-natural-stupidity-and-other-legal-idiocies-by-vincenzo-zeno-zencovich/>

¹⁸⁴ Mark A. Lemley, *How generative ai turns copyright upside down*, in *Science & Technology Law Review*, 2024, p. 24 ss.

¹⁸⁵ S. Rodotà, *La vita e le regole*, 2006, p. 15 ss. Alla visione dell'unitarietà della persona, proprio alla luce delle sfide poste dalla rivoluzione tecnologica, si è ormai sostituita la considerazione della molteplicità di interessi di cui essa è portatrice. Il rilievo di questi interessi si gioca non da un punto di vista del loro astratto riconoscimento, ma sul piano di una valutazione della loro meritevolezza di tutela che sottende la "selezione" degli interessi ritenuti rilevanti

Esigenza che si apprezzi allorché l'indagine trasla dal piano delle mere architetture regolatorie a quello dell'effettività dei diritti, incrociando i pericoli di una deriva dei valori democratici. Qui, in discussione non è soltanto levare l'angosciato memento su di una tecnologia che altera gli equilibri economici e di mercato e appare, per converso, poco attenta ai diritti della persona¹⁸⁶. Il passaggio all'immateriale non si limita a riconfigurare lo scenario economico, ma finisce per inficiare la neutralità del mercato, adulterandone le logiche di accesso.

Quanto vale maggiormente ribadire sono allora gli effetti di una asimmetria della conoscenza che indebolisce le regole di funzionamento della democrazia perché finisce con il falsare quel rapporto delineato da Locke fra libertà e proprietà, dove questa si fa baluardo spaziale e giuridico entro cui la libertà può esplicarsi al riparo dall'arbitrio¹⁸⁷. Il che spiega, in un contesto socio-economico in cui i dati sono *the new oil* del futuro, l'esigenza di guardare con attenzione alle regole che sovrintendono alla appropriabilità della ricchezza onde arginare forme di feudalesimo tecnologico¹⁸⁸. La concentrazione dell'informazione nelle mani di pochi operatori, aprendo ad un "capitalismo dell'informazione", falsa infatti le regole del gioco democratico perché è in grado di vincolare le stesse scelte del potere politico¹⁸⁹. Si tratta di uno scenario, quello appena delineato, che interpella in profondità il diritto privato e le regole tradizionalmente preposte al trasferimento della ricchezza. Chi "possiede" i dati, invero, esercita oggi un potere che le

vecchie regole sulla proprietà non riescono a contenere.

L'avvertenza ulteriore è una sola, rotola a valle delle questioni più rilevanti sollevate dalla tecnologia, e sembra iscriversi nel solco della fulminante battuta formulata nel lontano 1941 da Louis D. Brandeis, membro progressista della Suprema Corte degli Stati Uniti il quale osservava che "Possiamo avere la democrazia o la ricchezza concentrata nelle mani di pochi, ma non possiamo avere entrambe le cose".

2. Dalla fiera einaudiana all'ordine algoritmico – Nel descrivere la vita economica di un borgo di campagna Einaudi icasticamente restituiva l'immagine di un luogo dove "i compratori desiderano di acquistare a buon mercato ed i venditori di vendere a caro prezzo. Spinti da motivi opposti essi si affrettano verso lo stesso luogo, verso la fiera, il mercato", evidenziandosi la necessità di "regole alle quali si deve ubbidire" e uniformarsi onde garantire il corretto funzionamento del sistema¹⁹⁰. Nel perimetro di questa materialità, così efficacemente descritta, il mercato si offriva storicamente come uno spazio fenomenico governato da un *nomos* eteronomo. Quale apparato di regole, pur espressione di precise traiettorie politiche, quel mercato manteneva tuttavia una distinzione formale rispetto alla sostanza dei beni scambiati. Si tratta di una separazione che l'avvento dell'immateriale, determinando una mutazione antropologica e strutturale delle categorie economiche, definitivamente scardina.

L'informazione, infatti, abdicando alla sua funzione ancillare rispetto alla produzione, lungi

e, in quanto tale, è espressione di una opzione di natura eminentemente politica.

¹⁸⁶ Come ben dimostra (da ultimo), la disciplina europea sull'utilizzo dei dati sanitari, EHDS.

¹⁸⁷ J. Locke, *Il secondo trattato sul governo*, trad. A. Gialluca, ed. BUR, Milano, 1998.

¹⁸⁸ Henry Moulds, *Private property in John Locke's state of nature*, in *The American Journal of Economics and Sociology*, Vol. 23, No. 2, Apr. 1964, pp. 179-188; A. Lustig, *Natural law, property, and justice: the general justification of property in John Locke*, in *The Journal of Religious Ethics*, Vol. 19, No. 1, Spring, 1991, p. 119-149; K. Olivecrona, *Appropriation in the state of nature: Locke on the origin of property*, in *Journal of the History of Ideas*,

Vol. 35, No. 2, Apr. – Jun. 1974, p. 211-230; David C. Snyder, *Locke on natural law and property rights*, in *Canadian Journal of Philosophy*, Vol. 16, No. 4, Dec., 1986, pag. 723-750.

¹⁸⁹ A. Las Casas, *Capitalismo dell'informazione e circolazione della ricchezza. Modelli giuridici statunitensi*, Napoli, 2024, p. 20 ss.; G. Guzzardi, *Il paradigma identitario nella società digitale*, in *Persona e Mercato*, 2023, p. 524 ss.; K. Cuckier, V. Mayer-Schoenberg, *The rise of big data. How it is changing the way we think about the world*, in *Foreign Affairs*, 2013, p. 28 ss.

¹⁹⁰ L. Einaudi, *Lezioni di politica sociale* (1944), Torino, 1975, pp. 11-12.

dal porsi quale mero riflesso informativo della realtà si trasforma – in quanto processata da algoritmi e sistemi di calcolo predittivo – in vera e propria materia prima, espressione di una nuova ontologia del valore. In questa prospettiva le dinamiche estrattive e le logiche di appropriazione non si limitano più a regolare lo scambio, ma appaiono plasmare *ex ante* lo stesso comportamento dei soggetti.

Ed infatti, la fusione progressiva fra reale e informazionale, consente all'impresa di agire quale vero e proprio "collettore di informazioni", rendendo difficile stabilire (come si vedrà) proprietà, criteri di distribuzione del valore e responsabilità. Vieppiù ove si consideri che i dispositivi di attribuzione della ricchezza – abbandonata la dimensione dello scambio – operano attraverso un istantaneo impulso digitale, che formalizza lo scambio di dati in cambio di servizi. L'estrazione di valore da flussi informativi frammentati e la loro ricomposizione in modelli di produzione e consumo orientati al profitto rende così evidente l'inadeguatezza delle regole classiche del possesso materiale. Né di solo questo si tratta.

A cambiare, a ben vedere, è lo stesso luogo dello scambio che non è più quel mercato vividamente descritto da Einaudi ma sono architetture digitali (le piattaforme) le quali non sono semplici "luoghi" dove si scambia il prodotto. Esse piuttosto creano e controllano le relazioni sociali ed economiche delle persone e il mercato stesso in cui quel prodotto circola presentandosi, in questo modo, sotto forma di un vero e proprio ordinamento "paragiuridico" che, si sovrappone a quello statale e stabilisce chi accede alla ricchezza e a quali condizioni. Laddove ciò che appare chiaro è, per un verso, il fatto che il valore dell'informazione non risiede più nella circolazione libera della stessa, ma è direttamente collegato ad una attività di "cattura e recinzione" della stessa e, per l'altro, che il controllo dei flussi informativi si trasforma in un potente strumento di dominio economico in

grado di scavalcare le tutele contrattuali classiche e svuotare di significato l'autonomia individuale¹⁹¹.

Quale ecosistema progettato per far incontrare gruppi diversi di interessi, la piattaforma assume così le funzioni di un vero e proprio intermediario capace, non solo, di "aspirare" tutti i dati da ogni interazione che avviene al suo interno ma anche di agire "libera" da regole che non siano quelle del codice tecnologico e dei termini di servizio. Il che, spostato che sia il discorso sotto il profilo assiologico, dà conto di una profonda frattura tra la fenomenologia economica e la sua traduzione normativa.

Mentre, infatti dal punto di vista economico, l'appropriazione del dato è garantita da prassi di segretezza e da una tecnologia computazionale volte ad assicurarne l'esclusività, la risposta dell'ordinamento giuridico mette piuttosto in rilevante (come si avvertiva) l'inefficienza dei vecchi paradigmi proprietari e la difficoltà di governare adeguatamente questi processi con regole nate per disciplinare rapporti aventi ad oggetti beni materialmente intesi, ma relazioni sociali ed economiche tra milioni di individui. Un buon esempio, in tale direzione, viene dalle questioni sollevate dall'AI in rapporto alle regole sul diritto d'autore¹⁹².

Il tentativo, sia nel contesto nordamericano che in quello europeo, di "recintare" l'immateriale attraverso il rinvio alle norme in tema di proprietà intellettuale, segreti industriali, tutela delle banche dati etc., finisce con il ricalcare un riflesso pavloviano di tipo conservativo il quale, mirando prevalentemente a stabilizzare la ricchezza, non riesce a risolvere in maniera soddisfacente la questione legata alla tutela dei soggetti che pure quella informazione hanno originariamente generato. Il tutto con quelle ricadute sul piano dello spazio di libertà e democrazia cui si è fatto cenno. Il diritto privato insomma (per tornare alla metafora della battuta di caccia) sembra inseguire la tigre senza riuscire ad afferrarla e, lungi da assurgere ad

¹⁹¹ A. Las Casas, *op. cit.*, p. 180 ss.; H. Zech, *Data as a Tradeable Commodity*, in A. De Franceschi (a cura di), *European Contract Law and the Digital Single Market. The implications of the Digital Revolution*, 2016, p. 51ss.

¹⁹² F. Banterle, *Data ownership in the data economy: a European dilemma*, in T.-E. Synodinou, P. Jougleux, C. Markou, T. Prastitou (a cura di), *EU Internet Law in the digital era*, Heidelberg, 2020, 199-225.

interlocutore privilegiato per garantire un'equa distribuzione del valore, finisce piuttosto con il vestire i panni del notaio chiamato a prendere atto di un'asimmetria di potere generata da un utente, il cui consenso ratifica di fatto la propria spoliatura economica.

Né le cose cambiano avuto riguardo alla tutela dei dati personali che circolano in rete e al loro sfruttamento economico. Lo spostamento del baricentro da un paradigma di libertà eterodiretta ad un ecosistema a estrazione predatoria, determinando una sussunzione integrale dell'esistenza, trasfigura la spontaneità del vissuto fenomenologico in un *asset* economico rigidamente predeterminato. In tale orizzonte, la soggettività viene progressivamente svuotata della propria eccedenza ontologica per essere ricondotta ai parametri della calcolabilità universale, ove l'agire umano non è più l'origine di una scelta autonoma, bensì il prodotto derivato di una costante e pervasiva capitalizzazione dell'immateriale.

3. Dati-fatto vs Dati-informazione: lo scardinamento delle categorie del copyright nell'era della riproduzione sintetica: il laboratorio della *litigation* – Un esempio utile ad illustrare le ricadute viene da quella crescente *litigation* che, nel tentativo di individuare la risposta alla domanda che tutti noi si conosce, sta progressivamente componendo (al di qua e al di là dell'oceano) il futuro del diritto d'autore nell'era dell'AI generativa. Il punto, si tratti di controversie legate all'uso non autorizzato di dati, immagini o musica si flette sulla possibilità di utilizzare (e in che misura) le opere protette da *copyright*, tenuto conto che – come noto – l'addestramento dell'AI si compie anche attraverso attività di *scraping*, uso di dati pubblici

e accordi di licenza¹⁹³. Laddove in discussione è lo scardinamento operato dalle piattaforme tecnologiche della stessa distinzione fra “dati-fatto” (liberamente circolanti) e “dati-informazione” (coperti dalle regole sul diritto di autore) su cui si è costruita la disciplina in tema di proprietà intellettuale. La scomposizione del dato in mero “vettore” statistico, senza distinguere fra mero dato e profilo creativo, e la sua sussunzione in una struttura algoritmica, rischia di espropriare l'identità espressiva dell'autore. Fenomeno che, tuttavia, non sempre prende la scena, trattandosi in concreto di verificare 1) il caso in cui queste informazioni siano state legittimamente raccolte e poi però rielaborate autonomamente e 2) quale sia il meccanismo appropriativo che è a fondamento del loro utilizzo.

All'interno di tale cornice di riferimento ciò che emerge è la centralità del contributo di una giurisprudenza che si afferma quale *ambiente* ideale per l'emergere di appropriate risposte in grado di tener conto – ci piaccia o no – del carattere estremamente dinamico di una realtà¹⁹⁴ la quale, già in passato, ha sfidato la comprensione e applicazione delle regole in tema di diritto d'autore. In tal senso, milita la considerazione delle provocazioni sollevate, nel tempo, dalla comparsa di registratori, fotocopiatrici e, da ultimo, internet avevano già messo “in manutenzione”, per dir così, non poche regole in tema di diritto d'autore. Tanto, in un contesto nel quale tutti, 1) autori dell'opera 2) sviluppatori 3) *public at large*, in fondo – parafrasando il lavoro di Paolo Sorrentino – hanno ragione, a sottolineare la legittimità delle rivendicazioni che prendono la scena in una situazione di confusione generata da strumenti innovativi¹⁹⁵.

¹⁹³ Y. Benhamou, *Licensing Big Data: a holistic analysis based on a three-step approach*, 2020, consultabile: https://www.researchgate.net/publication/339140374_Licensing_Big_Data_a_holistic_analysis_based_on_a_three-step_approach; R. Caso, *Il (declino del) diritto d'autore nell'era digitale: dalle misure tecnologiche di protezione al Digital Rights Management (Nota a Cass., sez. III, 7 aprile 2004)*, in *Foro it.*, 2004, II, c. 610; ID, *Digital Rights Management. Il commercio delle informazioni*

digitali tra contratto e diritto d'autore, Padova, 2004 (ristampa digitale, Trento, 2006, consultabile all'indirizzo <http://www.jus.unitn.it/users/caso/pubblicazioni/drm/home>).

¹⁹⁴ S. Rodotà, *Diritto, scienza, tecnologia: modelli e scelte di regolamentazione*, in *Scienza e diritto nel prisma del diritto comparato*, cit., p. 394.

¹⁹⁵ P. Sorrentino, *Hanno tutti ragione*, Feltrinelli, 2010.

Nella causa New York Times (NYT) vs OpenAI (e Microsoft, in quanto investitore chiave) attualmente in corso presso la Corte Distrettuale degli Stati Uniti per il Distretto Sud di New York¹⁹⁶ il punto centrale del contenzioso è, non a caso, la presunta violazione del *copyright* da parte di OpenAI, accusata di aver utilizzato milioni di articoli del quotidiano, protetti da diritto d'autore, per addestrare i suoi modelli generativi (come ChatGPT). E di averlo fatto senza autorizzazione né licenza. Il NYT ha sostenuto che l'addestramento è, nella specie, illegittimo perché non solo viola il loro *copyright*, ma consente a ChatGPT di generare *output* che di fatto sostituiscono i contenuti originali, sia 1) talvolta riproducendoli quasi alla lettera, sia 2) minando il sistema di abbonamento (paywall) e il traffico verso il loro sito. Il NYT ha, quindi, chiesto un risarcimento miliardario e la distruzione di tutti i modelli di AI addestrati con i loro contenuti, in quanto essi finiscono con l'operare una capitalizzazione parassitaria del valore intellettuale altrui. Dal canto suo OpenAI si è difesa appellandosi alla dottrina americana del *fair use*¹⁹⁷ così sostenendo che l'uso dei testi per l'addestramento dei modelli, finalizzato a creare qualcosa di nuovo e trasformativo, rientra pur sempre nei limiti normativamente fissati ed è comunque lecito.

Al momento, l'*order* del 4 aprile 2025 giunge come una vittoria parziale per il NYT, perché ha garantito che le rivendicazioni più importanti e centrali della causa — violazione diretta del *copyright*, *Contributory Copyright Infringement Claims* e *Trademark Dilution Claims* — procedano verso la fase di *discovery* e, dunque, verso un potenziale giudizio. Il cuore della

battaglia legale tra i media e l'AI è, dunque, confermato. Il cuore del *claim* investe la questione se l'addestramento e l'output dell'AI sui contenuti giornalistici violino o meno le norme in tema di *copyright* e se tale violazione possa essere giustificata dalla *fair use doctrine*.

La causa ha finito così con il trasformarsi in un'importante “guerra di trincea”, la cui decisione, quando arriverà, finirà con l'avere implicazioni enormi per il futuro dell'AI generativa e per il mercato dei contenuti online a livello globale.

Un ulteriore esempio delle tensioni cui vanno incontro le regole tradizionali si ricava da Thomson Reuters Enterprise Centre GmbH and West Publishing Corp., vs ROSS Intelligence Inc. (District Court of Delaware – febbraio 2025) relativo all'uso di un sistema di AI non generativo.

In questione, nel caso di specie, era l'addestramento del sistema di ROSS sulla base degli *Headnotes di Westlaw*. Operazione effettuata da una società terza cui ROSS aveva commissionato l'elaborazione di una serie di contenuti da offrire agli avvocati per facilitarne l'attività di ricerca. La difesa di ROSS si è basata sulla considerazione che quei sommari dei casi analizzati erano stati trasformati *into numerical data* e, dunque, in qualcosa di diverso dal dato grezzo primariamente utilizzato. In un'ottica lockiana, *Westlaw* ha insomma rivendicato la proprietà come frutto del proprio lavoro di sistematizzazione (*Key Number System*). ROSS, di contro, ha basato la sua linea difensiva sull'importanza di una conoscenza giuridica che, una volta digitalizzata, dovrebbe essere libera di fluire per alimentare nuove forme di intelligenza.

¹⁹⁶ The New York Times Company v. Microsoft Corporation (1:23-cv-11195) District Court, S.D. New York.

¹⁹⁷ L. Bently e B. Sherman, *Intellectual Property Law*, 4th ed., Oxford, 2014. I pilastri della eccezione di cui alla Sect. 107 del *Copyright Act* statunitense del 1976 sono rappresentati da 1) Finalità e carattere dell'uso che se trasformativo non va incontro ad alcuna limitazione; 2) natura dell'opera protetta; 3) quantità e consistenza del materiale utilizzato; 4) effetto dell'uso sul mercato potenziale. Il contenzioso *New York Times v. Microsoft/OpenAI* sposta il dibattito su un piano ontologico. La sfida è definire se l'addestramento dei

modelli linguistici (LLM) costituisca un uso trasformativo o una sussunzione parassitaria. B. Ely Marks, *Copyright Protection, Privacy Rights, and the Fair Use Doctrine: The Post-Salinger Decade Reconsidered*, in *New York University Law Review*, 1997, p. 1377 ss.; A. Bell and G. Parchomovsky, *The Dual-Grant Theory of Fair Use*, in *The University of Chicago Law Review*, vol. 83, 3, 2016, p. 1051 ss.; ID, *Considering the Role of Fairness in Copyright Fair Use*, *Institute for Intellectual Property & Information Law (IPIL) Symposia*, Vol. 61, 2, 2023, consultabile all'indirizzo <https://houstonlawreview.org/article/92123-considering-the-role-of-fairness-in-copyright-fair-use>

Il ragionamento dei giudici, nel sanzionare ROSS, è interessante e si aggancia alla considerazione che questi dati erano stati utilizzati *to create a legal research tool to compete with Westlaw* e, dunque, erano privi di carattere trasformativo. In discussione, non è l'uso del mero dato (possibile *ex fair use doctrine*) ma il fatto che “computer programs differ from books, films, and many other literary works in that such programs almost always serve functional purposes”¹⁹⁸. Ma i giudici, sono cauti e nel precisare che “Ross’s use is not transformative because it does not have a further purpose or different character from Thomson Reuters’s”, riconoscono che la soluzione offerta potrebbe non vale per l’AI generativa.

La causa è in appello (U.S. Court of Appeals for the Third Circuit) e ROSS sta chiedendo ai giudici di verificare se gli *headnotes* abbiano il carattere dell’originalità e possano essere comunque utilizzati per addestrare modelli di AI. La pendenza del giudizio di secondo grado testimonia, però, l’irrisolta tensione tra la tutela della proprietà intellettuale come baluardo lockiano e la spinta verso una liberalizzazione dei flussi informativi necessari all’evoluzione algoritmica. L’esito di tale contesa sarà quindi determinante per stabilire se il sistema normativo sarà in grado di governare la sussunzione dell’esperienza o se dovrà soccombere di fronte alla logica dell’accumulazione predatoria. Un dato comune alle due vicende giudiziarie comunque emerge e ha a che vedere con il fatto che in entrambe le circostanze i giudici hanno guardato alla 1) quantità del materiale utilizzato 2) alla natura

commerciale dell’uso e 3) alla maggiore o minore originalità del dato.

Diversa l’impostazione seguita in Europa, dove le cause sono state portate avanti ex artt. 3 e 4 della direttiva sul Digital Service Market (Direttiva 2019/790), la cui interpretazione è cruciale per l’esito di queste controversie. Il DSA, come noto, subordina a precise condizioni il *text and data mining* (ossia l’utilizzo in grandi quantità di dati, come testi, suoni, immagini e altro)¹⁹⁹. Esso infatti è lecito a) se compiuto da organismi di ricerca e da istituti di tutela del patrimonio culturale per scopi di ricerca scientifica; b) ai sensi dell’articolo 4, se ha ad oggetto dati (opere o altri materiali) contenuti in reti o in banche dati cui si ha legittimamente accesso ai fini dell’estrazione. In questo secondo caso, l’estrazione è tuttavia consentita esclusivamente se i titolari dei diritti non abbiano espresso una chiara riserva d’uso degli stessi, esercitando il cosiddetto *opt-out*. Rimane nondimeno in discussione la necessità di capire quanto parte dell’informazione sia stata utilizzata e se (e quale) l’apporto creativo in fase di trasformazione.

Il *claim* che ha contrapposto la *Gesellschaft für musikalische Aufführungs- und mechanische Vervielfältigungsrechte* (GEMA) da un lato e la *Suno Inc.*, dall’altro, (e che in un futuro non remoto, promette di coinvolgere organismi di gestione collettiva, come SIAE, SACEM o PRS for Music) è interessante perché investe il settore musicale e sembra iscriversi nel solco della sentenza resa nel caso *Robert Kneschke vs. LAION eV.*²⁰⁰.

La questione nel caso GEMA traeva alimento dall’uso dei testi di canzoni coperte da *copyright*.

¹⁹⁸ La dottrina del *fair use* consente dunque l’uso di opere coperte da diritto d’autore a condizione che siano rispettati i criteri di cui al 17 U.S. Code § 107 - Limitations on exclusive rights: Fair use. In giurisprudenza, altresì, *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 24 (2021). I giudici si concentrano su due profili 1) se l’uso ha carattere commerciale e 2) se è di natura trasformativa, in ciò richiamandosi a *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 529–31 (2023).

¹⁹⁹ M. W. Carroll, *Copyright and the Progress of Science: Why Text and Data Mining is Lawful*, 53 U.C. Davis L. Rev. 893 (2019).

²⁰⁰ Cfr. *Robert Kneschke v. LAION eV.*, Rif.: 310 O 227/23 che ha fornito lo schema argomentativo su cui si sono poi mossi i giudici del Tribunale di Monaco nel caso GEMA. La decisione riguardava il fotografo Robert Kneschke le cui foto erano state inserite in un dataset per l’addestramento dell’IA nonostante il suo “opt-out”. In dottrina F. Sibilla, G. Puzio e G. Ruma, *ChatGPT e diritto d’autore: la sentenza tedesca che ridefinisce i confini dell’AI generativa* [Landgericht München I, Urteil vom 11.11.2025, Az. 42 O 14139/24 (GEMA v. OpenAI, in *Diritto di Internet. Intelligenza artificiale. Digital Copyright e Data Protection*, 2025, passim).

I giudici, nel condannare OpenAI effettuano un ragionamento a partire dalla segmentazione del processo di acquisizione dei dati che si regge sulla osservazione di ciò che accade nella fase, per dir così, dinamica dell'uso delle informazioni, vale a dire quella in cui il set di dati reso leggibile è poi arricchito da metadati utilizzati poi per addestrare l'AI²⁰¹. In discussione non è l'analisi automatizzata dei dati di addestramento. Il punto è, piuttosto, l'attività di memorizzazione dei dati, perché è questa che consente la riproduzione tramite *prompt* e, dunque, incide sull'uso delle informazioni. Su queste basi per i giudici si sono create le condizioni per quella *Zugänglichmachung* di cui al § 19a *Urheberrechtsgesetz* (UrhG)²⁰².

Per i giudici tedeschi non v'è spazio per l'eccezione di cui del § 44b UrhG, perché questa opera al solo livello di compilazione del set di dati di addestramento (Text and Data Mining) senza che possa trovare spazio alcuno l'incorporazione e la memorizzazione permanente delle opere protette nel modello di intelligenza artificiale²⁰³. GEMA, da parte sua, aveva esercitato il diritto di *opt-out* nel maggio 2022, modificando l'accordo di adesione dei suoi autori per vietare l'uso dei contenuti a fini di addestramento dell'AI. Tra i tanti spunti che questa decisione offre, sicuramente interessante è il ragionamento in punto di consenso del *Rechtsinhaber*. Per i giudici l'interferenza di OpenAI con i diritti di sfruttamento di GEMA non è giustificata da alcun consenso dei titolari dei diritti, ritenendosi che l'addestramento dei

modelli non sia da considerarsi una forma d'uso consuetudinaria e prevedibile che un titolare dei diritti debba anticipare (o aspettarsi).

La decisione sembra segnare la fine di una sorta di impunità algoritmica. A differenza di quelli americani, i giudici in Europa (almeno i giudici tedeschi) sembrano inclini a riconoscere all'autore il diritto di opporsi all'uso. Diversamente da quanto accade per la giurisprudenza americana (*ROSS Intelligence*) che si dibatte nel tentativo di misurare quanto l'algoritmo "distorca" il mercato, la posizione europea (sintetizzata dal caso *GEMA*) sposta l'asse sulla legittimità stessa del prelievo dell'informazione. In questa direzione il diritto europeo, atteso che il mercato non è più solo uno spazio di scambio tra beni, ma un ecosistema di estrazione di senso, pare incline a suggerire che un argine alla predazione può soltanto essere individuato nella creazione di infrastrutture giuridiche che impongano alla tecnica di negoziare con la creatività.

4. L'ontologia estrattiva e l'illusione del consenso: dirigismo burocratico e limiti della tutela nella "società algoritmica" – Né le cose sono più semplici sul piano dell'uso dei dati personali, atteso che la barriera eretta a tutela della persona e della sua riservatezza non riesce a dar effettivamente risposta ad esigenze di controllo. L'ontologia estrattiva che è a base dell'architettura dei processi digitali è qui particolarmente evidente, specie ove si considerino le dinamiche che ne sono a

²⁰¹ R. Caso e G. Pascuzzi, *Il diritto d'autore nell'era digitale*, in G. Pascuzzi (a cura di), *Il diritto dell'era digitale*, Bologna, 2016, 145-195. Nella prima fase, un set di dati di training viene creato e convertito in un formato leggibile dalle macchine. Successivamente, nella seconda fase, il set di dati viene analizzato e arricchito con metadati. Il modello di intelligenza artificiale viene addestrato utilizzando questo set di dati. La terza fase, infine, prevede l'utilizzo del modello addestrato tramite *prompt* e la ricezione di output.

²⁰² § 19a *Urheberrechtsgesetz* (UrhG) recita "Il diritto di messa a disposizione del pubblico è il diritto di rendere l'opera accessibile al pubblico, con o senza filo, in modo tale che ciascuno possa avervi accesso dal luogo e nel momento scelti individualmente". La *Zugänglichmachung* descrive l'atto di inserire un'opera in un ambiente

(solitamente telematico) in modo che il pubblico possa fruirne "on demand".

²⁰³ § 44b Text und Data Mining

(1) Text und Data Mining ist die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen.

(2) Zulässig sind Vervielfältigungen von rechtmäßig zugänglichen Werken für das Text und Data Mining. Die Vervielfältigungen sind zu löschen, wenn sie für das Text und Data Mining nicht mehr erforderlich sind.

(3) Nutzungen nach Absatz 2 Satz 1 sind nur zulässig, wenn der Rechtsinhaber sich diese nicht vorbehalten hat. Ein Nutzungsvorbehalt bei online zugänglichen Werken ist nur dann wirksam, wenn er in maschinenlesbarer Form erfolgt.

fondamento; dinamiche legate ad una raccolta massiva di dati quale presupposto indefettibile per l'erogazione di servizi apparentemente gratuiti. In tale scenario, la fruizione di contenuti e funzionalità si configura come un'utilità che — secondo una prospettiva pragmatica — parrebbe legittimare l'acquisizione dei flussi informativi quale controprestazione implicita. Il godimento di funzionalità digitali si pone come l'elemento prestazionale che legittima il drenaggio dei flussi informativi, operando una surrettizia conversione del dato in controprestazione contrattuale.

Certo, la Direttiva (UE) 2019/770²⁰⁴ nel tentativo di ricondurre il fenomeno entro le rassicuranti categorie dello scambio, delinea una struttura sinallagmatica *sui generis* che pare intesa a "neutralizzare" la narrazione di una deriva puramente espropriativa. La contraddizione è qui, per un verso, nell'esigenza di garantire un nesso di proporzionalità tra la ricchezza immateriale ceduta e la remunerazione effettivamente conseguita dall'utente e, per l'altro, nell'opacità dei meccanismi per assicurare un reale equilibrio economico. Tensione cui solo in apparenza tenta di porre rimedio una disciplina che impone requisiti di trasparenza, volontarietà e consapevolezza dei termini di siffatta circolazione del valore.

Né sfugge la diversa risposta — proprio in punto di dati personali — articolata sulle due sponde dell'Atlantico; una diversità che rende manifesta la profonda divergenza nelle tecniche regolatorie e nelle *rationes* disciplinari. Ancorché radicate in una medesima tradizione giuridica, l'esperienza europea e quella statunitense si lasciano apprezzare per l'emergere di visioni inconciliabili circa il limite tra la libertà contrattuale e la tutela della sovranità individuale dinanzi alla potenza tecnica.

Il discrimine essenziale risiede nella polarizzazione tra il pragmatismo del "marketplace discourse" d'oltreoceano — volto a massimizzare il flusso informativo in chiave utilitaristica — e la resistenza dogmatica europea, orientata piuttosto a privilegiare una "rights talk vision", dove la normazione funge da presidio dei diritti della personalità e delle libertà fondamentali. L'architettura regolatoria dell'Unione Europea in materia di Intelligenza Artificiale non si esaurisce in un mero esercizio di normazione tecnica. Questa proiezione dei valori costitutivi dell'ordinamento eurounitario emerge nell'adozione di una precisa strategia geopolitica che, lungi dal risolversi nell'esercizio di una normazione tecnica, mira a definire le coordinate per uno *human-centric approach*. La subordinazione dello sviluppo tecnologico alla tutela dei diritti fondamentali sembra trovare così riflessi concreti in un solido apparato normativo che dà respiro al c.d. *Brussels Effect*²⁰⁵. Laddove obiettivo sistematico è la saturazione dello spazio regolatorio, volta ad impedire la creazione di "zone franche" estrattive e, piuttosto, a riaffermare che l'innovazione non può mai prescindere dalla salvaguardia della dignità della persona.

In questa direzione non v'è dubbio che il principio di 1) autonomia del dato ex art. 8 CDFUE; 2) del primato dello stato di diritto (Criteri di Copenaghen), al fine di impedire che l'efficienza algoritmica diventi una forma di potere extragiuridico capace di erodere le basi stesse della convivenza democratica; 3) la decisa presa di distanza da quel *presumptive entitlement* statunitense (che presuppone un diritto implicito all'estrazione dati) di cui al GDPR che intende trasformare il trattamento dei dati da "libertà di esercizio" per le piattaforme a "eccezione regolata", si iscriva nel solco di una vera e propria strategia di de-mercificazione

²⁰⁴ Cfr. Direttiva (UE) 2019/770 del Parlamento Europeo e del Consiglio del 20 maggio 2019 relativa a determinati aspetti dei contratti di fornitura di contenuto digitale e di servizi digitali, recepita con D. Lgs. 4 novembre 2021, n. 173. In dottrina, per tutti, S. Sica, V. D'Antonio, *La commodification dei dati personali nella data driven society*, in P. Stanzione (a cura di), *I poteri privati delle*

piattaforme e le nuove frontiere della privacy, Torino, 2022, p. 129 ss.

²⁰⁵ Il "Modello Bruxelles" si presenta in questo senso come un tentativo di influenzare lo sviluppo tecnologico, rendendo la conformità ai valori etici europei un requisito per accedere al vasto mercato interno. E ciò appare chiaro proprio guardando, *inter alia*, al Reg 2016/679 (Liceità del trattamento, art. 6; diritti dell'interessato art. 13-21).

della persona, onde restituire al soggetto il potere di veto sul proprio vissuto digitale²⁰⁶.

E, tuttavia, l'impianto regolatorio europeo, pur elevando il consenso a dogma di liceità (art. 6 del GDPR), sembra oggi prigioniero di una contraddizione ontologica. Nella prassi della *platform economy* il consenso si trasfigura in un dispositivo di assoggettamento dove la "meritevolezza" del trattamento è mero simulacro formale che occulta un processo di patrimonializzazione forzata dell'identità, come pure dimostra il recente intervento in tema di dati sanitari di cui all'EHDS²⁰⁷.

Si tratta di una frizione di cui è specchio proprio la Dir. 2019/770, dove l'inadeguatezza del consenso si palesa non solo come asimmetria informativa, ma come vera e propria coercizione strutturale espressa dal tentativo del legislatore di applicare le tutele del consumatore al titolare del dato. La protezione consumeristica si fa così, in ultima istanza, espressione di un rimedio palliativo. Timido tentativo di offrire garanzie formali rispetto ad un'espropriazione di valore che il consenso, ormai svuotato di ogni reale potere di veto, non può più interdire, ma solo silenziosamente assecondare. Come dire che, a dispetto della fedeltà proclamata allo *human approach*, in concreto le cose vanno diversamente perché non si riesce né nell'intento di offrire una effettiva tutela alla persona, né nell'obiettivo di "controllare" la tecnologia.

L'aspirazione a definire uno standard di tutela uniforme si traduce, piuttosto, nell'ipertrofia di un apparato burocratico pur inteso ad adeguare lo sviluppo scientifico e tecnico in nome di principi etici di ordine superiore. In questo senso, il Regolamento (UE) 2024/1689 (AI Act) si pone quale espressione di un dirigismo normativo che, nel pretendere di regolare l'imprevedibilità del progresso, rischia di soffocare la radice epistemologica. L'effetto che

sembra prodursi evoca per analogia storica, come non ha mancato di sottolineare attenta dottrina, la parabola delle scoperte astronomiche di Galileo Galilei, i cui esiti furono silenziati sotto la minaccia della scomunica e dell'esilio poiché rei di aver sfidato la dogmatica teologica del tempo²⁰⁸. Ciò che tragicamente manca nella *ratio* dell'AI Act è, in ultima istanza, la comprensione dell'epistemologia tecnica. Il dogma della regolamentazione universale finisce per scontrarsi con l'essenza stessa della scoperta scientifica che, per sua natura, mal sopporta le catene di una burocrazia che pretende di normare l'ignoto prima ancora che esso sia compiutamente esplorato.

Sotto il profilo strettamente economico, l'AI Act si configura così come una tipica Barriera Tecnica al Commercio (TBT), finalizzata a proteggere un'industria digitale europea strutturalmente non competitiva che, in questo modo, tenta di acquisire una posizione di forza nel quadro dei negoziati bilaterali con i blocchi statunitense e cinese. In tale prospettiva il legislatore europeo appare ignorare la peculiarità di un'attività (quella tecnologica) che si pone in ontologica opposizione alla tendenza a regolamentare ogni aspetto del reale. Né le cose cambiano avuto riguardo all'esperienza italiana e al tentativo di cui alla Legge 23 settembre 2025, n. 132 (recante disposizioni e deleghe in materia di intelligenza artificiale) di promuovere un uso etico, trasparente e antropocentrico dell'AI, rafforzando (*inter alia*) gli obblighi di informazione.

Diversa l'impostazione seguita in USA come pure emerge dall'American Artificial Intelligence Initiative del 2019. Il mosaico è qui flessibile, orientato al mercato e quindi con lo sguardo alla promozione dell'innovazione, competitività globale e sicurezza nazionale. Sul piano giuridico il modello del *notice and choice* si fa fonte privilegiata di un sistema che sembrerebbe

²⁰⁶ G. Bincoletto, *La privacy by design. Un'analisi comparata nell'era digitale*, Ariccia, 2019.

²⁰⁷ Sul punto sia consentito il rinvio a V. Zambrano, *EHDS e soggetti vulnerabili: tra esigenze di condivisione dei dati sanitari e tutela della persona*, in *Accademia. Rivista dell'Associazione italiana civilisti*, 2025, p. 1096 ss.; M. Arisi e P. Guarda, *Blockchain and eHealth: seeking*

compliance with the General Data Protection Regulation, in *Biolaw Journal – Rivista di Biodiritto*, 2020, n. 2, 477-496; A. Bernes, *La protezione dei dati personali nell'attività di ricerca scientifica*, in *Nuove leggi civ.*, 2020, 175.

²⁰⁸ V. Zeno Zencovich, *Artificial intelligence, natural stupidity and other legal idiocies*, cit.

rinvia a forme esclusive di appartenenza costrette a fare i conti con la costruzione di situazioni giuridiche esemplificate dal dualismo *right to privacy vs right to publicity*.

In antitesi al modello europeo, la prospettiva d'oltreoceano delinea insomma un modello dove 1) il mercato libero è inteso come volano di innovazione; 2) si riconosce la libertà per sviluppatori e utenti di orientare gli strumenti tecnici secondo necessità, nel rispetto di standard di sicurezza di buon senso (*common-sense*); 3) si fissa il solo divieto di strumentalizzare l'intelligenza artificiale per finalità di controllo sociale o di coercizione tra Stati²⁰⁹. Né la dimensione proprietaria riesce a dar conto pienamente degli effetti di tale circolazione con riguardo alla ricchezza di ritorno. Le risposte più recenti (come il Blueprint for an AI Bill of Rights del 2022 e l'Ordine Esecutivo 14110 del 2023) hanno colmato parzialmente questo vuoto, introducendo principi etici (come il diritto di rinuncia ai sistemi automatizzati) e imponendo requisiti di sicurezza a livello governativo. Si tratta, tuttavia, in gran parte di linee guida e direttive piuttosto che normative vincolanti come accade in una Europa piuttosto interessata a porre un "freno" etico preventivo per prevenire i danni. Laddove gli USA sembrano inclini ad affrontare i problemi etici *ex post*, utilizzando strumenti flessibili e settoriali.

5. Conclusioni: La faglia atlantica e l'aporia della regolazione tecnica – Cosa emerge all'esito del discorso? A dispetto delle differenze di

impostazione nel rapporto fra soggetto e raccoglitore del dato siamo dinanzi – piaccia o meno – ad un medesimo processo di accumulazione di "capitale informativo" governato più che dalla logica commutativa dello scambio dalla inevitabilità dell'accesso alla "digital enclosure", con la relativa esposizione alla sorveglianza che tutto ciò comporta unitamente alla capacità degli attori del "capitalismo informazionale" (S. Zuboff) di estrarre valore dalle risorse rese disponibili dall'innovazione tecnologica²¹⁰.

Certo l'UE sembra voler correre ai ripari. E qui sul piano normativo gioca un solido impianto regolatorio che si compone (oltre a quanto detto) dei vari Regolamenti sul governo dei dati (Reg. 2022/868 sui limiti alla negoziazione in relazione alla cessione dei dati; art. 10-15 sulla assistenza alla manifestazione del consenso; Reg. sui mercati digitali (Reg. 2022/1925) che vuole regolare i rapporti fra utenti commerciali, piattaforme e utenti finali – con divieto di utilizzo per alcuni impieghi dei dati raccolti (artt. 5 e 6). In ultima analisi, la divergenza tra le due sponde dell'Oceano palesa una frammentazione assiologica profonda all'interno del medesimo alveo della civiltà occidentale. Se l'Europa, arroccata nel suo "Modello Bruxelles", sembra aver optato per un diritto difensivo che utilizza la regolamentazione quale barriera identitaria e mercantile, gli Stati Uniti riaffermano una visione promozionale e competitiva, in cui la norma resta ancillare rispetto alla libertà di mercato e alla dominanza tecnologica. Né si tratta di una sfasatura meramente burocratica.

²⁰⁹ In questa direzione, cfr. il documento inviato da OpenAI il 13 marzo 2025, *all'Office of Science and Technology Policy*. Prendendo spunto dalle parole di Trump ("It is the policy of the United States to sustain and enhance America's global dominance in order to promote human flourishing, economic competitiveness, and national security" – President Donald J. Trump, *Executive Order 14179, January 23, 2025*), nel documento si evidenzia la necessità di un "holistic approach that enables voluntary partnership between the federal government and the private sector, and neutralizes potential PRC benefit from American AI companies having to comply with overly burdensome state laws". Il tutto nella prospettiva della promozione di 1) "A free market promoting free and fair competition that drives innovation; 2) Freedom for

developers and users to work with and direct our tools as they see fit, in exchange for following clear, common-sense technical standards that help keep AI safe for everyone, and being held accountable when they don't; 3) Preventing government use of AI tools to amass power and control their citizens, or to threaten or coerce other states", in <https://cdn.openai.com/global-affairs/openai-us-economicblueprint-feb-2025-edu-update.pdf>.

²¹⁰ A. Las Casas, *op. cit.*, p. 180 ss.; S. Zuboff *The Age for Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power* (Profile Books, 2019); N. Srnicek, *Capitalismo digitale. Google, Facebook, Amazon e la nuova economia del web*, LUISS University Press, Roma, 2017, pp. 13.

Essa, piuttosto, riflette l'incompiutezza di una sintesi tra la tutela dei diritti fondamentali e l'imprevedibilità dell'innovazione. Mentre il legislatore europeo insegue l'illusione di una "tecnica governata" per decreto — rischiando quel segnalato anacronismo galileiano di una verità imposta per via dogmatica — la prospettiva d'oltreoceano scommette sulla capacità del mercato di autoregolarsi attorno a standard di sicurezza minimi. Si assiste così ad una polarizzazione sistematica che mette a nudo la vera complessità del presente: vale a dire l'impossibilità di ricondurre a un'unità coerente la velocità della scoperta scientifica e la staticità della norma giuridica. L'Occidente si trova così diviso tra un "museo dei diritti" e una "frontiera senza limiti", dimostrando che (a meno di accontentarsi di prendere la tigre per la coda) la regolazione delle tecnologie non è solo una sfida tecnica, ma una drammatica scelta di civiltà tra la sicurezza e il rischio del progresso.

16. Postfazione (La persona nell'ecosistema digitale: tutela dei dati e nuove forme della proprietà intellettuale)

La tutela giuridica della persona, tradizionalmente ordinata entro la categoria dei «diritti della personalità», è stata chiamata — al pari di gran parte degli istituti del diritto pubblico e privato contemporaneo — a misurarsi con il paradigma tecnologico del terzo millennio. Negli ultimi anni, tuttavia, la trasformazione ha assunto caratteri ulteriori. Le tecnologie digitali non offrono più soltanto nuovi mezzi per l'esercizio di diritti già conosciuti: intervengono sulla conformazione delle relazioni sociali ed economiche, modificano le condizioni di accesso ai beni e ai servizi, distribuiscono capacità di controllo e incidono, per tale via, sulle possibilità concrete attraverso le quali la persona svolge liberamente la propria personalità.

Ai diritti della personalità vengono tradizionalmente riferiti i caratteri dell'assolutezza, della non patrimonialità, dell'intrasmissibilità, dell'irrinunciabilità e dell'imprescrittibilità. Anche l'indisponibilità, frequentemente richiamata, deve essere intesa

senza irrigidimenti, alla luce delle singole situazioni protette e delle forme di consenso ammesse dall'ordinamento. Tali qualificazioni esprimono, nondimeno, la funzione essenziale della categoria: proteggere gli attributi costitutivi dell'individuo, nella sfera materiale dell'integrità fisica come in quella morale e immateriale.

Sul piano storico, la categoria — non sistematizzata in forma unitaria nell'esperienza francese post-rivoluzionaria del *Code civil* napoleonico — ha rinvenuto un decisivo terreno di elaborazione nella dottrina tedesca di fine Ottocento, diffondendosi successivamente negli ordinamenti europei attraverso itinerari non coincidenti.

Nell'esperienza tedesca, la riflessione dottrinale sulla personalità, sviluppata poi dalla giurisprudenza costituzionale del secondo dopoguerra, ha condotto al riconoscimento di un «diritto generale della personalità» (*Allgemeines Persönlichkeitsrecht*), capace di ricondurre a una comune matrice le differenti manifestazioni dell'individualità. Altri ordinamenti hanno privilegiato una ricostruzione pluralistica, fondata sul riconoscimento di singoli diritti posti a protezione di specifici attributi della persona. La contrapposizione non descrive, peraltro, sistemi rigidamente separati: clausole generali, previsioni tipiche e interventi giurisprudenziali hanno prodotto aperture e contaminazioni reciproche.

Una tendenza originariamente pluralistica può registrarsi nell'esperienza italiana dei Codici e delle leggi speciali, nei quali trovano puntuale tutela il nome e lo pseudonimo, l'immagine e l'integrità fisica, mentre ulteriori situazioni — dall'onore alla reputazione — sono state progressivamente definite anche attraverso l'elaborazione giurisprudenziale. Questa trama normativa deve però essere letta alla luce dei principi costituzionali e, soprattutto, dell'art. 2 Cost., che ha conferito al sistema una maggiore elasticità, rendendolo capace di reagire all'evoluzione sociale, economica e politica senza dipendere dalla preventiva tipizzazione legislativa delle singole situazioni protette.

La rilettura costituzionalmente orientata dei diritti della personalità ha così consentito di riconoscere figure non espressamente contemplate dal legislatore. È il caso del diritto all'identità personale, intesa non soltanto quale identità fisica del soggetto, riconducibile ai suoi connotati identificativi, ma quale proiezione ideale della persona nel contesto delle relazioni sociali e in rapporto alla rappresentazione che di essa viene fornita all'esterno. La distanza tra ciò che la persona è, ciò che comunica di sé e ciò che altri possono ricostruire sul suo conto avrebbe assunto, con la digitalizzazione, un rilievo difficilmente prevedibile entro le coordinate originarie della categoria.

Su questa trama dogmatica irrompe prepotente il diffondersi delle tecnologie digitali, che incidono sull'essenza, come sul concreto declinarsi, dei diritti della personalità. Internet non costituisce ormai soltanto un mezzo di interconnessione, ricerca e diffusione delle informazioni. È infrastruttura per lo studio e il lavoro, per l'accesso ai servizi e l'esercizio dell'iniziativa economica, per la partecipazione politica e il mantenimento delle relazioni personali. La disponibilità della tecnologia diviene, pertanto, condizione di esercizio di libertà che conservano una propria consistenza al di fuori della rete, ma che nella rete trovano uno dei principali ambienti di attuazione.

La dimensione digitale non si sovrappone a quella analogica come uno spazio separato: la attraversa e la trasforma. L'attività, più o meno consapevole, compiuta online, così come i dati che altri (con o senza il consenso dell'interessato) immettono nei circuiti informativi, lasciano un'impronta che, progressivamente stratificandosi, concorre a costruire la personalità digitale – *recte*, l'identità digitale – di ciascun individuo. A tale rappresentazione non contribuiscono soltanto le informazioni volontariamente comunicate: la loro aggregazione e l'elaborazione algoritmica permettono di inferire abitudini, inclinazioni e preferenze che il soggetto potrebbe non avere mai manifestato e delle quali potrebbe persino ignorare l'esistenza.

La nozione opera, peraltro, su due piani distinti, ancorché comunicanti. L'identità digitale designa la proiezione della persona nell'ambiente tecnologico, risultante dalle sue manifestazioni, relazioni e tracce informative; in un'accezione più stringente, indica i dati e gli strumenti che consentono di identificare l'utente, abilitarne l'accesso ai servizi e imputargli atti giuridicamente rilevanti. Ne discendono problemi diversi ma connessi: il controllo sulla rappresentazione della persona e sulla circolazione delle informazioni che la riguardano; l'autenticazione, l'anonimato, l'accesso e la certezza dell'imputazione degli atti compiuti in rete.

Il pervasivo impiego delle tecnologie dell'informazione e della comunicazione ha così imposto agli interpreti problemi nuovi di efficacia ed effettività delle tutele. Si è cominciato, in tal senso, a discorrere di «diritti digitali», intendendo per essi le forme attraverso le quali la personalità umana si esplica nell'ambiente tecnologico e le risposte offerte dall'ordinamento ai rischi che quell'ambiente produce o amplifica.

In prima approssimazione, alla locuzione sono state ricondotte due categorie di situazioni giuridiche, distinguendo tra «diritti nativi analogici» e «diritti nativi digitali». I primi corrispondono alle nuove declinazioni assunte da diritti già conosciuti nel confronto con la tecnologia: la riservatezza, l'identità personale, la libertà di manifestazione del pensiero e il diritto d'autore, una volta esercitati online, incontrano problemi in parte sconosciuti alla dimensione puramente analogica. I secondi presuppongono, nel loro stesso esistere, un'infrastruttura tecnologica: è il caso dell'accesso alla rete e delle situazioni connesse alla *net neutrality*, al *digital divide*, alla cittadinanza digitale o al diritto alla disconnessione.

La bipartizione conserva una sua utilità descrittiva, ma non esaurisce il campo. Accanto alla trasformazione dei diritti preesistenti e all'emersione di situazioni che presuppongono la rete, hanno assunto rilievo le garanzie rese necessarie dalla comparsa di nuovi poteri:

raccogliere e correlare dati, ordinare le informazioni, selezionare i contenuti visibili, profilare gli utenti, assumere decisioni automatizzate, consentire o impedire l'accesso agli spazi nei quali si svolge una parte crescente della vita sociale. Le tre aree non compongono una tassonomia chiusa. L'accesso è insieme condizione di esercizio di altri diritti e garanzia contro l'esclusione infrastrutturale; la protezione dei dati costituisce tanto un'evoluzione della riservatezza quanto un limite ai poteri informativi; la libertà di espressione dipende ormai anche dalle regole negoziali e tecniche che governano la visibilità dei contenuti.

Il diritto di accesso alla rete non può, pertanto, essere ridotto alla materiale disponibilità di una connessione. Esso richiama le condizioni economiche, territoriali e culturali che rendono effettiva la partecipazione all'ambiente digitale, nonché la qualità degli strumenti mediante i quali quella partecipazione si realizza. La neutralità della rete investe, a sua volta, il rapporto tra operatori delle telecomunicazioni, fornitori di servizi e contenuti, concorrenza e libertà degli utenti. Un accesso formalmente aperto perderebbe gran parte della propria consistenza ove la circolazione delle informazioni fosse sottoposta a discriminazioni opache, rallentamenti selettivi o corsie preferenziali determinate dalla capacità economica dei soggetti coinvolti.

All'esclusione infrastrutturale si affianca il potere di chi organizza gli spazi digitali. Le piattaforme regolano, attraverso condizioni generali e strumenti di autonomia privata, l'accesso a luoghi essenziali della comunicazione; stabiliscono le regole di permanenza, organizzano la visibilità dei contenuti, ne dispongono la rimozione e applicano misure suscettibili di incidere sulla posizione degli utenti. Il contratto non costituisce, in questo quadro, un semplice involucro formale, ma uno dei luoghi nei quali si distribuiscono poteri capaci di condizionare l'identità, la libertà di espressione e la partecipazione alla vita sociale. Assumono allora rilievo la trasparenza delle condizioni applicate,

la proporzionalità delle misure, l'effettività degli strumenti di contestazione e i rimedi contro l'esercizio arbitrario del potere privato.

Il progressivo offuscamento dei confini tra pubblico e privato spiega l'emersione del «costituzionalismo digitale», inteso quale tentativo di verificare come la limitazione del potere e la protezione dei diritti possano operare in un ambiente non governato esclusivamente dallo Stato. La locuzione non implica la meccanica trasposizione delle categorie costituzionali in uno spazio immateriale, né esaurisce le questioni sollevate dal diritto delle tecnologie. La produzione delle regole dipende ormai dall'interazione tra legislazione, giurisprudenza, autorità indipendenti, *standard* tecnici, autoregolamentazione e discipline negoziali. La prospettiva costituzionale si intreccia, perciò, con problemi propriamente privatistici: autonomia negoziale e suoi limiti, responsabilità degli intermediari, tutela dei dati nelle relazioni economiche, posizione dell'utente e predisposizione di rimedi effettivi.

L'eterogeneità delle esperienze ordinamentali richiede una comparazione che non si arresti all'accostamento delle disposizioni, ma distingua il ruolo della legislazione, delle corti, delle autorità amministrative, della dottrina, delle prassi negoziali e degli *standard* tecnici. Le differenze possono dipendere dalla struttura dell'ordinamento, dal grado di consolidamento della tutela o dal momento nel quale essa viene osservata. Le soluzioni, inoltre, circolano e vengono adattate a concezioni non coincidenti della persona e del rapporto tra libertà e potere. La vicenda della riservatezza offre un primo esempio. Nel formante dottrinale e giurisprudenziale statunitense, il *right to privacy* ha trovato una celebre elaborazione nel saggio di Warren e Brandeis, che riprese la formula *cooleyana* del *right to be let alone*, sviluppandosi poi quale protezione da intrusioni indebite e, su un diverso versante, quale spazio di autonomia nelle decisioni attinenti alla sfera più intima della persona. Nello spazio giuridico europeo, il diritto al rispetto della vita privata e familiare, consacrato dall'art. 8 CEDU, si è

incontrato con una disciplina progressivamente autonoma del trattamento dei dati personali. La tutela non consiste più soltanto nell'esclusione delle ingerenze, ma attribuisce all'interessato poteri di conoscenza, intervento e controllo sulla circolazione delle informazioni che lo riguardano.

Entro questa traiettoria, un passaggio decisivo è rappresentato dal formante giurisprudenziale tedesco.

La sentenza del *Bundesverfassungsgericht* sul censimento del 1983 ha ricavato dalla dignità umana e dal libero sviluppo della personalità il diritto all'autodeterminazione informativa, inteso quale potere dell'individuo di decidere, in linea di principio, sulla divulgazione e sull'impiego delle informazioni appartenenti alla propria sfera personale. La costruzione ha esercitato una significativa influenza sulla cultura europea della protezione dei dati, senza tradursi in un dominio assoluto dell'interessato sull'informazione: il controllo individuale deve confrontarsi con interessi collettivi, obblighi normativi e finalità pubbliche legittime.

La tutela dell'identità e dei dati personali si intreccia con la libertà di espressione. Il diritto del singolo a non essere indefinitamente rappresentato attraverso informazioni risalenti può entrare in conflitto con l'interesse pubblico alla conoscenza e con la memoria collettiva. L'anonimato può proteggere da repressioni e discriminazioni, ma anche rendere più difficile l'individuazione del responsabile di condotte illecite. La moderazione dei contenuti si colloca, a sua volta, nella zona di confine tra tutela degli utenti, libertà di manifestazione del pensiero e autonomia privata delle piattaforme. La soluzione dipende dalle regole applicate, dalle ragioni e dalla proporzionalità della misura, nonché dalla presenza di procedure che ne consentano la contestazione.

La rete modifica, insieme al controllo delle informazioni personali, le condizioni di circolazione dei contenuti protetti. La riproducibilità delle opere e la dimensione transnazionale delle comunicazioni hanno sottoposto a pressione una disciplina costruita su diritti territoriali, supporti materiali e forme

di diffusione più facilmente controllabili. Le convenzioni internazionali, le legislazioni nazionali e gli interventi dell'Unione europea non producono una *reductio ad unitatem*. Il punto di equilibrio tra tutela degli autori e dei titolari, accesso alla conoscenza, innovazione e libertà di circolazione dipende dalla configurazione delle eccezioni, dalla posizione attribuita agli intermediari e dai rimedi disponibili contro la violazione.

La cittadinanza digitale porta il discorso oltre il rapporto individuale con il dato. Essa riguarda l'accesso ai servizi, la partecipazione alla vita pubblica, l'identificazione nei rapporti con l'amministrazione e la capacità di agire consapevolmente negli ambienti online. Il confronto sulle tecnologie di sorveglianza mostra, tuttavia, come strumenti almeno in parte simili possano inserirsi in differenti combinazioni di interessi e formanti. Nell'esperienza cinese assume particolare rilievo l'integrazione tra apparati pubblici, raccolta dei dati e infrastrutture tecnologiche; in quella statunitense, ai poteri connessi alla sicurezza si affianca la capacità informativa delle grandi imprese private; nell'ordinamento dell'Unione europea, legislazione, autorità e giurisprudenza tentano di sottoporre la raccolta e l'impiego delle informazioni ai principi di necessità, proporzionalità e protezione dei diritti fondamentali. Non si tratta di modelli monolitici: le tensioni interne e le reciproche contaminazioni impediscono di far coincidere ciascuna esperienza con una sola logica di controllo.

Eppure, la riflessione sui diritti digitali non esaurisce la proiezione della persona verso la contemporaneità. Ciò che, nell'introduzione alla precedente edizione, poteva apparire come futuro prossimo venturo è ormai oggetto di un dibattito giuridico concreto. Le neurotecnologie rendono progressivamente accessibili informazioni relative all'attività cerebrale e prospettano possibilità di intervento sulla sfera cognitiva, emotiva e volitiva. Il problema non riguarda più soltanto ciò che la persona comunica o ciò che può essere inferito dai suoi comportamenti, ma la possibilità di osservare e,

in prospettiva, condizionare processi appartenenti alla sua sfera interiore.

Le esperienze spagnola e cilena rendono visibili due differenti strategie regolatorie. La *Carta de derechos digitales* spagnola affida a un atto di *soft law* l'individuazione di principi destinati a orientare la disciplina e tende prevalentemente a ricondurre le nuove esigenze entro diritti già conosciuti. Il Cile ha percorso la via dell'*hard law*, intervenendo sul testo costituzionale e affidando alla legislazione la regolazione dell'accesso e del monitoraggio dell'attività cerebrale. La contrapposizione non esaurisce la complessità delle due esperienze, ma mostra i limiti di entrambe le tecniche: la prima è priva di coerenza; la seconda rischia di moltiplicare le situazioni soggettive prima di avere chiarito se e in quale misura le garanzie esistenti siano inadeguate.

La questione consiste allora nello stabilire se dignità, integrità mentale, libertà di pensiero, riservatezza e protezione dei dati possano offrire una trama normativa sufficiente o se la peculiarità dei dati neurali richieda garanzie ulteriori. Il riconoscimento di nuovi diritti non può derivare dalla sola novità della tecnologia: deve fondarsi sull'individuazione di interessi non adeguatamente protetti e di rischi che le categorie vigenti non siano in grado di governare. Al tempo stesso, la riconducibilità di una lesione a un diritto già conosciuto non assicura che gli strumenti predisposti per la sua tutela siano adeguati alla capacità delle neurotecnologie di accedere direttamente ai processi mentali.

In definitiva, i diritti digitali non identificano un catalogo stabile di situazioni soggettive nate o trasformate dalla tecnologia. Esprimono la ricerca delle condizioni giuridiche necessarie affinché la persona possa conservare libertà, dignità e autodeterminazione all'interno di ambienti regolati da norme pubbliche, poteri privati, infrastrutture tecniche e sistemi algoritmici.

La traiettoria conduce dall'*habeas corpus* dello Stato di diritto all'*habeas data* della società dell'informazione, sino all'*habeas mentem* della società post-informatica. Non una successione

lineare, nella quale ciascuna formula sostituisce la precedente, ma la compresenza di livelli di protezione che l'innovazione rende sempre più strettamente connessi. Corpo, informazioni e mente non rappresentano tappe ormai superate: sono dimensioni della persona che il diritto è chiamato a proteggere nella loro crescente interdipendenza.

SEZIONE 2

**RASSEGNA DI GIURISPRUDENZA STRANIERA
RAGIONATA E COMMENTATA**

**IL CONTENZIOSO ANTHROPIC: USO TRASFORMATIVO,
PIRATERIA E GUARDRAILS**

**CONCORD MUSIC GROUP, INC. ET AL. V. ANTHROPIC
PBC E BARTZ V. ANTHROPIC PBC (N.D. CAL., No.
5:24-cv-03811 E No. 4:24-cv-05417)**

di Sara Manni e Leonardo Ortenzi*

I. Premessa

Il contenzioso che ha visto Anthropic PBC convenuta in giudizio dinanzi alle corti federali statunitensi si colloca all'interno della più ampia stagione giurisprudenziale sviluppatasi negli ultimi anni intorno all'impiego di opere protette nei processi di sviluppo e funzionamento dei modelli di intelligenza artificiale generativa. Rispetto, tuttavia, ai procedimenti aventi ad oggetto opere figurative, fotografiche o contenuti giornalistici, le controversie qui in esame presentano un tratto peculiare che giova sin d'ora segnalare: la questione non si esaurisce nell'uso delle opere durante la fase di addestramento del modello (in avanti, opere *input*), ma investe altresì, e in misura crescente, la capacità del sistema *Claude* di riprodurre, in fase di *output* e su richiesta degli utenti, contenuti sostanzialmente coincidenti con opere protette.

Le vicende qui in commento riguardano due distinti filoni processuali.

Il primo, *Bartz v. Anthropic PBC* (N.D. Cal., No. 3:24-cv-05417-WHA), è stato promosso da tre autori statunitensi - Andrea Bartz, Charles Graeber e Kirk Wallace Johnson - i quali hanno contestato alla società convenuta di aver riprodotto e conservato, senza autorizzazione, milioni di opere letterarie protette, in parte acquisite mediante il ricorso a *shadow libraries* (ovverosia, archivi di opere non autorizzati e, sostanzialmente, pirata).

Il secondo, *Concord Music Group, Inc. et al. v. Anthropic PBC* (originariamente M.D. Tenn., No. 3:23-cv-01092, trasferito al Northern District of California con il nuovo numero di registro 5:24-

cv-03811), è stato instaurato da alcuni dei principali editori musicali statunitensi - fra cui Universal Music Corp., Concord Music Group, Inc. e ABKCO Music, Inc. - i quali hanno lamentato l'utilizzo non autorizzato di testi musicali (*i.e.*, nell'azione, le «*lyrics*») di loro titolarità tanto nella fase di addestramento, quanto nella successiva generazione di output testuali da parte del sistema *Claude*.

II. Bartz v. Anthropic PBC

Prima ancora del procedimento promosso dagli editori musicali, il dibattito statunitense sull'intelligenza artificiale generativa si è sviluppato in misura significativa intorno alla questione dell'utilizzo di opere protette per finalità di addestramento dei *Large Language Models*. L'azione *Bartz* si inserisce in tale filone, ed è stata depositata nell'agosto 2024 presso la *United States District Court for the Northern District of California* da tre autori statunitensi, in proprio e a nome di una classe potenziale di titolari di diritti d'autore.

Le doglianze attoree muovevano dalle disposizioni del *Copyright Act of 1976*, e in particolare dalla Section 106 (17 U.S.C. § 106), che attribuisce al titolare del *copyright* il diritto esclusivo di riprodurre l'opera e di sfruttarla economicamente.

Secondo gli attori, Anthropic avrebbe riprodotto, acquisito e conservato senza alcuna autorizzazione un numero estremamente elevato di opere letterarie protette - oltre sette milioni di volumi, in base alle risultanze istruttorie - parzialmente ottenute attraverso archivi digitali operanti in violazione della normativa sul diritto d'autore, quali *Books3*, *Library Genesis* (LibGen) e *Pirate Library Mirror* (PiLiMi), al fine di impiegarle nell'addestramento del modello di intelligenza artificiale denominato *Claude*. A tali download si affiancava l'acquisto di milioni di volumi cartacei, destinati alla digitalizzazione previa distruzione dell'originale, condotta oggetto

* Premesso che l'opera è frutto del lavoro congiunto dei due autori, a Leonardo Ortenzi vanno attribuiti i Paragrafi I e IV e a Sara Manni i Paragrafi II e III.

della seconda fattispecie esaminata dall'ordine di cui si dirà.

La convenuta, dal canto suo, sosteneva che il *training* del modello costituisca una forma di utilizzazione trasformativa riconducibile alla *doctrine of fair use* disciplinata dalla Section 107 del *Copyright Act* (17 U.S.C. § 107).

In tale prospettiva, i *Large Language Models* non opererebbero alla stregua di archivi di opere memorizzate, bensì apprenderebbero relazioni statistiche e strutture linguistiche astratte attraverso processi computazionali, sì da non determinare alcuna riproduzione espressiva delle singole opere²¹¹.

La difesa fondata sul *fair use* richiama in modo pressoché obbligato il precedente *Authors Guild v. Google, Inc.* (804 F.3d 202, 2d Cir. 2015), relativo al progetto *Google Books*, nel quale la Corte d'Appello del Secondo Circuito aveva riconosciuto la natura trasformativa della scansione massiva di opere letterarie, laddove finalizzata all'indicizzazione e alla ricerca testuale²¹². Tale precedente costituisce, com'è noto, uno dei principali riferimenti dottrinali e giurisprudenziali per le società operanti nel settore dell'IA generativa, pur essendovi differenze non trascurabili rispetto allo scenario tecnologico attuale²¹³.

Ebbene, in data 23 giugno 2025 il giudice W. Alsup ha pronunciato una decisione su *partial summary judgment* il cui rilievo non risiede solo nel merito delle questioni decise, ma anche nella

sistematizzazione concettuale delle tre fattispecie ivi distinte:

- (i) l'utilizzazione di opere protette per finalità di addestramento del modello;
- (ii) la digitalizzazione, mediante distruzione delle copie cartacee originali, di volumi legittimamente acquistati;
- (iii) il *download* e la conservazione permanente di copie ottenute da archivi pirata.

Quanto al primo profilo, il giudice ha ritenuto che l'utilizzo di opere protette da *copyright* a fini di addestramento abbia natura marcatamente trasformativa e che, pertanto, il *training* del modello mediante opere letterarie legittimamente acquisite rientri nella *doctrine of fair use*.

La pronuncia si caratterizza, sul piano terminologico, per la perentorietà e l'enfasi con cui viene individuato il carattere trasformativo del *training*: i termini utilizzati – che si riportano in lingua originale – sono stati, tra gli altri, «*exceedingly transformative*», «*spectacularly so*», «*quintessentially transformative*» e «*among the most transformative many of us will see in our lifetimes*».

L'iter argomentativo muove dalla constatazione per cui il processo di addestramento, lungi dal comportare la mera duplicazione delle opere, consista – come argomentato dalla convenuta –

²¹¹ In questo senso, M.A. LEMLEY, B. CASEY, *Fair Learning*, in *Texas Law Review*, vol. 99, 2021, pp. 743 ss., secondo i quali l'addestramento dei modelli linguistici presenta caratteristiche profondamente differenti rispetto alle tradizionali forme di sfruttamento economico dell'opera protetta, in quanto il modello non utilizza il contenuto per finalità espressive dirette, bensì per apprendere correlazioni statistiche e strutture linguistiche astratte; in tale prospettiva, il *training* algoritmico costituirebbe una forma altamente trasformativa di utilizzazione dell'opera, maggiormente assimilabile ai processi di apprendimento umano e di *text and data mining* che non alla distribuzione o riproduzione tradizionale dei contenuti protetti, argomentando dunque per il «*fair learning*» in luogo del «*fair use*». Sull'apprendimento statistico dell'IA, in termini di *machine learning* e *deep learning* v. M. TAMPIERI, *L'intelligenza artificiale e le sue evoluzioni. Prospettive civilistiche*, Milano, 2022, 69-75.

²¹² Decisione a sua volta fondata sulla definizione di uso «trasformativo» stabilita in *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578-579. Sul *transformative use* nella dottrina statunitense, v., *inter alia*, per una rassegna giurisprudenziale, J. LIU, *An Empirical Study of Transformative Use in Copyright Law*, in 22 *Stan. Tech. L. Rev.* 163, 2019.

²¹³ In effetti, il precedente *de quo* ha trovato – come si vedrà nei commenti che seguono – largo impiego nelle comparse dei diversi fornitori di IA, di volta in volta citati a giudizio. D'altronde, il medesimo veniva già diffusamente richiamato nello stesso Report sul *training* delle IA dell'Ufficio Copyright statunitense. USCO, *Report on Copyright and Artificial Intelligence. Part 3: Generative AI Training (Pre-publication)*, 9 maggio 2025, *passim*, disponibile in *copyright.gov*.

in una elaborazione computazionale finalizzata all'estrazione di *pattern* linguistici, strutture sintattiche e regolarità statistiche presenti nel *corpus* testuale, con la conseguenza che il modello risultante costituisce rappresentazione astratta e compressa delle regolarità linguistiche del *dataset*²¹⁴.

Analoga conclusione è stata raggiunta con riferimento alla seconda fattispecie.

Il Giudice ha ritenuto che la digitalizzazione di volumi legittimamente acquistati, accompagnata dalla distruzione delle copie cartacee originali, configurasse del pari una utilizzazione *fair*, in quanto sostanzialmente equivalente a una mutazione di formato che non incide sull'economia delle opere e che agevola unicamente la conservazione e la ricerca interna.

Infine, l'ordine ha invece rigettato l'istanza di *summary judgment* con riferimento alla costituzione e al mantenimento di una c.d. «*permanent central library*» alimentata, in parte consistente, da materiali scaricati da archivi pirata.

Tale condotta integrerebbe infatti autonoma e indipendente violazione del diritto esclusivo di riproduzione spettante al titolare del *copyright*, non trovando giustificazione nella *doctrine of fair use* (17 U.S.C. § 107): la circostanza per la quale l'utilizzazione finale del materiale così acquisito possa presentare profili di "trasformatività" (a valle) non elide l'illiceità della condotta di acquisizione (a monte). Sul punto, l'ordine ha disposto la celebrazione del *trial* in ordine alle copie piratate confluite nella *central library* e ai conseguenti danni, attuali o statuari, anche sotto il profilo della *willfulness*.

Il passaggio appare di notevole rilievo sistematico.

L'ordine sembra infatti introdurre, sebbene in forma ancora embrionale, una distinzione destinata, forse, a divenire centrale: la separazione fra effettivo *training use* e *dataset acquisition*. Secondo tale impostazione, la *doctrine of fair use* potrebbe in linea di principio coprire l'utilizzazione trasformativa delle opere nel processo di addestramento, ma non legittimerebbe le eventuali modalità illecite con le quali le medesime opere siano state acquisite²¹⁵.

La pronuncia, nel complesso, appare equilibrata: da un lato, evita di accogliere una concezione meramente intuitiva o allarmistica del fenomeno dell'IA generativa, riconoscendo che il *training* non coincide con una tradizionale attività di distribuzione o sfruttamento commerciale dell'opera; dall'altro, respinge una lettura eccessivamente ampia del *fair use*, precisando che il carattere innovativo della tecnologia non possa tradursi in una sorta di immunità rispetto alle regole del *copyright*.

Nei mesi successivi alla decisione del giugno 2025, il procedimento ha conosciuto due snodi di considerevole importanza, che paiono aver inciso in misura determinante sull'esito complessivo della controversia.

Dapprima, nel luglio 2025, il Giudice ha provveduto a certificare una *class action* limitatamente al profilo della pirateria, definendo come classe tutti i titolari del diritto esclusivo di riproduzione delle opere comprese nei *dataset* LibGen e PiLiMi, purché regolarmente registrate presso lo U.S. Copyright Office.

²¹⁴ Peraltro, dalla «estrazione di *pattern* linguistici, strutture sintattiche e regolarità statistiche presenti nel *corpus* testuale» consegue un certo *modus* espressivo della IA, tipico di ogni modello: è stata argomentata la riconoscibilità ed evidente differenza – graficamente percepibile nel dato statistico riportato – della scrittura "artificiale" rispetto alla scrittura umana. Cfr. Y. ZOU, F. KUEK, K. H. NG, X. CHENG, *Comparative analysis of text readability and writing styles in AI-generated vs. Human-written academic abstracts*, in *Plos One*, 4, 2026, disponibile online all'indirizzo <<https://doi.org/10.1371/journal.pone.0343163>>; F. A. RODRIGUES, N. F. STURM, F. L. PINHEIRO, *A linguistic*

comparison between human- and AI-generated content, in *iScience*, 29, 2026, disponibile online all'indirizzo <<https://doi.org/10.1016/j.isci.2026.114976>>.

²¹⁵ Il meccanismo di operatività delineato è – a ben vedere – il medesimo della Text and Data Mining exception europea. Difatti, l'utilizzo di opere coperte da copyright ai fini di *training* dell'IA è lecito, nell'Unione, ai sensi dell'art. 53, par. 1, lett. c) dell'AI Act e dell'art. 4, dir. (UE) 2019/790, a patto che alle opere si abbia *legittimamente* accesso e che venga rispettata la riserva di diritti eventualmente esercitata dal titolare.

Avverso tale provvedimento Anthropic ha proposto una *petition* ai sensi della Rule 23(f) dinanzi alla Corte d'Appello del Ninth Circuit (n. 25-4843), chiedendo altresì al giudice di prime cure la certificazione dell'appello interlocutorio sulla parziale reiezione del *summary judgement*, nonché la sospensione del giudizio. L'istanza di sospensione è stata respinta dal giudice Alsup con ordine dell'11 agosto 2025, che ha mantenuto ferma la data del procedimento, fissato sin dall'origine al 1° dicembre 2025; la *petition* al Ninth Circuit, invece, non è mai giunta a decisione, essendo sopravvenuto l'accordo transattivo.

In tale contesto - caratterizzato, da un lato, dall'accertamento dell'illiceità della costituzione della *central library* tratta da archivi pirata e, dall'altro, dalla certificazione di una classe potenzialmente comprensiva di milioni di titolari di diritti - le parti hanno avviato un percorso di mediazione, conclusosi con la stipulazione di un *settlement agreement* annunciato nell'agosto 2025 e ratificato in via preliminare il 25 settembre 2025. L'accordo prevede, fra l'altro, l'impegno di Anthropic a corrispondere una somma non inferiore a un miliardo e mezzo di dollari per la definizione delle pretese connesse all'acquisizione dei volumi mediante *shadow libraries*. L'accordo contempla altresì l'obbligo di distruggere le copie piratate e i relativi dataset, con attestazione scritta dell'avvenuta distruzione, e circoscrive l'effetto liberatorio alle pregresse condotte di acquisizione e *training*, senza estendersi a eventuali pretese fondate sugli output dei modelli. Il procedimento di approvazione, peraltro, non risulta ancora concluso: esaurito il *claims process*, l'udienza di *final approval* si è tenuta nella primavera 2026 e la relativa decisione risulta, allo stato, riservata. La scelta della soluzione transattiva – ancorché si tratti di somma apparentemente esorbitante – appare ampiamente comprensibile alla luce delle conseguenze economiche potenzialmente catastrofiche che una pronuncia di merito sfavorevole avrebbe potuto comportare. La normativa statunitense in materia di diritto

d'autore prevede infatti che il titolare possa conseguire *statutory damages* compresi, in via ordinaria, tra 750,00 e 30.000,00 dollari per ciascuna opera violata, elevabili sino a 150.000,00 dollari in presenza di violazione *willful*²¹⁶.

III. Concord Music Group, Inc. et al. v. Anthropic PBC

Parallelamente al filone *Bartz*, Anthropic si è trovata a fronteggiare un distinto contenzioso instaurato da alcuni dei principali editori musicali statunitensi. La citazione è stata depositata il 18 ottobre 2023 presso la *United States District Court for the Middle District of Tennessee* (No. 3:23-cv-01092), dinanzi al *Chief Judge* Waverly D. Crenshaw, Jr.; a seguito di una pronuncia sulla giurisdizione personale e sulla competenza territoriale emessa nel giugno 2024, il procedimento è stato trasferito al *Northern District of California* (No. 5:24-cv-03811) e assegnato al giudice Eumi K. Lee.

Le doglianze attoree, articolate in più capi, si fondano sulla contestata utilizzazione non autorizzata di testi di canzoni e composizioni musicali (*lyrics*) tanto nella fase di *training* del modello, quanto nella successiva fase di generazione degli *output* testuali; a tali condotte si affianca la dedotta rimozione delle c.d. *copyright management information* (CMI), in violazione della Sec. 1202(b) del Digital Millennium Copyright Act (DMCA). L'atto introduttivo elenca circa cinquecento composizioni del repertorio delle attrici (Exhibit A). Gli editori musicali hanno prodotto, già in fase introduttiva, numerosi esempi nei quali *Claude* generava - in risposta a *prompt* mirati - testi sostanzialmente coincidenti con quelli di celebri composizioni del repertorio internazionale, talvolta limitandosi a sostituire poche parole o ad apportare modifiche marginali. Sul piano dei rimedi, oltre alla declaratoria di violazione diretta del *copyright* ai sensi della Section 501 (17 U.S.C. § 501), le attrici hanno invocato anche le dottrine giurisprudenziali del *contributory infringement* e

²¹⁶Cfr. 17 U.S.C. § 504(c).

del *vicarious infringement*, elaborate dalla giurisprudenza statunitense a precisate, sul versante tecnologico, con i casi *Sony Corp. of America v. Universal City Studios, Inc.* (464 U.S. 417, 1984) e *MGM Studios, Inc. v. Grokster, Ltd.* (545 U.S. 913, 2005)²¹⁷.

Il richiamo a tali precedenti, peraltro, va accolto con qualche cautela. È vero che il contenzioso *Concord* richiama, sul piano della struttura argomentativa, le controversie che nei primi anni Duemila videro l'industria musicale contrapposta a piattaforme quali *Napster* e *Grokster*; è altrettanto vero, tuttavia, che il fenomeno dell'IA generativa presenta tratti tecnologici peculiari, che rendono il parallelismo solo parzialmente adeguato. *Claude* non

distribuisce, infatti, file preesistenti caricati da terzi, bensì genera autonomamente l'*output* mediante processi probabilistici fondati sul *training* algoritmico²¹⁸, sicché le categorie tradizionali della responsabilità secondaria – segnatamente quelle plasmate sul modello del *peer-to-peer* – necessitano di una rilettura non scontata.

La convenuta ha articolato la propria strategia difensiva su due piani concorrenti. Sul piano sostanziale, è stata invocata (anche in questo caso così come – si vedrà – in molti altri) la *doctrine of fair use* di cui alla Section 107 del *Copyright Act*, sostenendosi che il *training* del modello costituisca una utilizzazione altamente trasformativa, finalizzata non già alla

²¹⁷ Trattasi delle due figure in cui la giurisprudenza statunitense ha scomposto la *secondary liability* in materia di copyright; entrambe presuppongono un'altrui violazione diretta e nessuna delle due è codificata.

Il *contributory infringement* designa la responsabilità di chi, consapevole dell'altrui attività illecita, la induca, la determini o vi contribuisca in misura rilevante (la formula risale a *Gershwin Publishing Corp. v. Columbia Artists Management, 1971*). Si fonda dunque su due elementi – la conoscenza (effettiva o, in taluni *circuits*, ricavabile dalle circostanze) e il contributo materiale – ed è temperato, sul versante tecnologico, dalla *staple article of commerce doctrine* enucleata in *Sony Corp. of America v. Universal City Studios* (1984), per cui chi immette in commercio un prodotto «*capable of substantial noninfringing uses*» non risponde per la sola idoneità di esso all'uso illecito. A tale limite si sottrae però l'*inducement*, autonoma figura mutuata dal diritto dei brevetti e consacrata in *MGM Studios, Inc. v. Grokster, Ltd.* (2005): chi distribuisce un dispositivo o un servizio con il fine, manifestato da espressioni chiare o da iniziative positive, di promuoverne l'impiego a fini di violazione risponde delle violazioni che ne derivano, ancorché il prodotto sia astrattamente suscettibile di usi leciti.

Il *vicarious infringement* designa invece la responsabilità di chi (i) abbia il diritto e la possibilità di controllare l'attività illecita e (ii) ne tragga un beneficio economico diretto, a prescindere dalla conoscenza (*Shapiro, Bernstein & Co. v. H.L. Green Co.*, 1963; *Fonovisa, Inc. v. Cherry Auction, Inc.*, 1996). La sua matrice è il *respondeat superior*, esteso oltre il rapporto di lavoro a ogni situazione di controllo accompagnata da interesse economico (così, ad es., per il gestore del mercato rispetto agli ambulanti che vi smerciano falsi). La sintesi è in *Grokster*: si risponde a titolo contributorio inducendo o incoraggiando intenzionalmente la violazione altrui, e a titolo *vicarious* lucrando su di essa pur potendola impedire.

Nell'ordinamento italiano nessuna delle due figure è autonomamente tipizzata. Il *contributory infringement* trova il proprio referente nelle regole sul concorso nell'illecito e sulla responsabilità solidale (art. 2055 c.c.), che peraltro postulano di norma almeno la colpa del concorrente; il *vicarious infringement* evoca piuttosto la responsabilità del committente per il fatto del commesso (art. 2049 c.c.), di carattere oggettivo ma circoscritta al rapporto di preposizione, e dunque più angusta del modello statunitense, che prescinde da ogni vincolo di subordinazione. Per le piattaforme – terreno d'elezione del problema nel contenzioso sull'IA – la funzione svolta oltreoceano dalla *secondary liability* è assorbita dal regime speciale degli intermediari (artt. 14-17 d.lgs. 70/2003, oggi nel *Digital Services Act*, reg. UE 2022/2065; art. 17 dir. (UE) 2019/790, attuato con d.lgs. 177/2021) e dalla giurisprudenza della Corte di giustizia in tema di «comunicazione al pubblico» degli *upload* illeciti (cause riunite C-682/18 e C-683/18, *YouTube* e *Cyando*, 2021).

²¹⁸ In questo senso, dal punto di vista dell'industria musicale, C. Goldstuck, *Past Precedent, Future Proof: Toward a New Legal and Commercial Framework for AI-Generated Music* (2025), il quale rileva come *Napster* non costituisca un precedente in senso proprio, bensì un «cautionary tale»: la vera discontinuità della GAI risiederebbe nel fatto che i modelli non distribuiscono opere preesistenti ma ne generano di nuove, sicché – come nel caso di specie – le categorie della responsabilità secondaria di matrice *peer-to-peer* mal si attagliano al fenomeno. L'A. sottolinea altresì l'inversione delle dinamiche di potere rispetto all'era *Napster*: non più *start-up* isolate, ma piattaforme sostenute dai capitali delle grandi imprese tecnologiche (Microsoft/OpenAI, Amazon/Anthropic).

redistribuzione commerciale dei testi musicali, bensì all'apprendimento statistico delle strutture linguistiche e semantiche della lingua naturale²¹⁹.

Sul piano tecnico-fattuale, Anthropic ha rappresentato di aver implementato specifici sistemi di protezione tecnologica - definiti *guardrails* - volti a impedire la riproduzione integrale dei testi musicali in fase di *output*²²⁰. Tali strumenti, comprensivi di filtri automatici sui *prompt* e di meccanismi di riconoscimento dei *lyrics*, sono stati presentati come idonei a contenere in misura significativa il rischio di *infringing output*. È peraltro proprio su questo profilo che il procedimento ha conosciuto un primo significativo sviluppo: in data 30 dicembre 2024, le parti hanno raggiunto una stipulazione parziale, recepita dal giudice Lee con ordine del 2 gennaio 2025, in forza della quale Anthropic si è impegnata a mantenere i *guardrails* già implementati nei propri modelli e a sottoporsi a un meccanismo di *notice* su impulso degli editori musicali. Tale intesa, pur senza definire il merito della controversia, ha avuto l'effetto di limitare considerevolmente il *thema decidendum* della richiesta cautelare ancora pendente.

Con ordine del 25 marzo 2025, il giudice Lee si è pronunciato sulla *preliminary injunction* residua respingendola; il giorno successivo ha invece accolto la *motion to dismiss* proposta dalla convenuta, pervenendo a una duplice statuizione favorevole ad Anthropic.

Quanto alla richiesta cautelare, la Corte ha ritenuto che le attrici non avessero adeguatamente dimostrato la sussistenza dei

requisiti necessari per l'adozione della tutela d'urgenza, e segnatamente del requisito dell'*irreparable harm*. Il giudice ha osservato, in particolare, che il danno lamentato dagli editori musicali appariva di natura prevalentemente economica e, come tale, potenzialmente ristorabile mediante risarcimento monetario nel giudizio di merito; ha inoltre rilevato che la *injunction* richiesta presentava una formulazione eccessivamente ampia (*overbroad*) e di problematica attuazione, estendendosi a centinaia di migliaia di opere, ivi comprese acquisizioni future, senza che la convenuta potesse ragionevolmente accertarne il perimetro. Il rigetto è stato pronunciato *without prejudice*, senza alcuna statuizione sul merito delle contrapposte tesi in punto di *infringement* e di *fair use*, quest'ultimo espressamente qualificato dalla Corte come questione ancora aperta.

Con riferimento, invece, alla *motion to dismiss*, la Corte ha accolto le eccezioni di Anthropic relativamente ai capi concernenti il *contributory infringement*, il *vicarious infringement* e la rimozione delle informazioni sulla gestione del diritto d'autore (CMI) di cui alla Section 1202 del *Copyright Act*, concedendo peraltro la *leave to amend*. Il giudice ha rilevato, sotto il primo profilo, che le attrici non avevano sufficientemente allegato l'esistenza di un atto di violazione diretta da parte di utenti terzi - presupposto strutturale della responsabilità secondaria - né la conoscenza o conoscibilità di tali violazioni da parte di Anthropic; sotto il secondo, ha osservato che la dedotta rimozione

²¹⁹ In senso critico rispetto all'argomento dell'impraticabilità del licensing, v. M. STRATTON, *Market-Based Licensing for Publishers' Works is Feasible*. *Big Tech Agrees.*, 48 Colum. J.L. & Arts (forthcoming 2025), il quale osserva che la stessa Anthropic, negli atti del presente giudizio, ha riconosciuto l'esistenza di un mercato consolidato per la visualizzazione dei testi musicali online (*Surreponse*, Dkt. No. 256, a p. 2), sì che l'ambito dei titolari e delle opere non risulta proibitivo ai fini della licenza per il *training*. Ancor più significativamente, il consulente della convenuta ha ammesso che un mercato composto da molte migliaia o milioni di transazioni di licenza sarebbe, sul piano economico, «unremarkable» (Declaration of Steven R. Peterson, Dkt. No. 210, par. 10) - concessione che l'A. valorizza per escludere tanto

l'impossibilità quanto l'insostenibilità dei costi di transazione, con evidenti riflessi sul quarto fattore del *fair use*.

²²⁰ Un utile termine di paragone è offerto dall'evoluzione di YouTube: cfr. Goldstuck, *op. cit.*, che ricostruisce come il sistema Content ID (2007) sia divenuto non tanto uno strumento difensivo quanto l'infrastruttura tecnica su cui si è edificato un mercato di licenze (oltre 6 miliardi di dollari corrisposti all'industria musicale nel 2024), suggerendo, *mutatis mutandis*, che i *guardrails* possano trascendere la funzione meramente cautelare per divenire presupposto di un modello remunerativo negoziato.

di *copyright management information* non risultava sufficientemente caratterizzata sul piano dell'intenzionalità, anche in considerazione del fatto che la rimozione non risultava sistematica nell'intero processo di addestramento.

Nel complesso, la pronuncia evita di assumere una posizione netta rispetto alla questione di fondo della liceità dell'uso di opere protette per finalità di *training*. Da un lato, la Corte riconosce che il *training* presenta profili di significativa trasformatività e che il rapporto fra IA generativa e *fair use* costituisce questione giuridica ancora aperta; dall'altro, chiarisce che tale circostanza non esaurisce il problema della rigenerazione di contenuti protetti in fase di *output*.

Interessante notare che, nel caso di specie, ad essere controversa non era tanto l'utilizzazione delle opere durante il *training*, quanto la concreta capacità del modello di generare *output* sostanzialmente coincidenti con le opere originarie. La Corte sembra dunque riconoscere, sia pure implicitamente, che la eventuale qualificazione del *training* in termini di *fair use* non costituisce scriminante automatica rispetto all'*infringing output* successivo.

Particolare rilievo assume, infine, la valutazione dei *guardrails* implementati dalla convenuta. La presenza di tali sistemi di controllo tecnologico contribuisce, nell'economia della decisione, al rigetto della *preliminary injunction*: la Corte rileva che Anthropic aveva già adottato

strumenti tecnici idonei a ridurre il rischio di violazioni sistematiche del *copyright*, circostanza che rende meno giustificabile l'adozione immediata di misure cautelari ulteriori²²¹.

Resta peraltro fermo, secondo il giudice, che la presenza di tali sistemi non risolve in via definitiva il problema sotteso alla controversia: qualora *Claude* continuasse a generare *output* sostanzialmente coincidenti con testi musicali protetti, Anthropic potrebbe comunque essere chiamata a rispondere delle conseguenze patrimoniali derivanti dal funzionamento del modello²²².

Il quadro delineato dai provvedimenti di marzo 2025 è peraltro mutato significativamente negli sviluppi successivi. Avvalendosi della *leave to amend*, gli editori hanno depositato, nell'aprile 2025, una domanda emendata; con ordine del 6 ottobre 2025 il giudice Lee ha respinto la nuova *motion to dismiss* della convenuta, ritenendo che la domanda emendata allegasse in modo plausibile tanto la conoscenza, attuale o almeno per *willful blindness*, delle violazioni commesse dagli utenti, quanto il beneficio economico diretto rilevante ai fini della *vicarious liability*, nonché l'elemento soggettivo richiesto per la rimozione delle CMI: i capi in precedenza espunti sono stati così riammessi al giudizio.

L'implementazione e il monitoraggio dei *guardrails* – valorizzati in sede cautelare a favore della convenuta – sono stati assunti quale base inferenziale della conoscenza, in capo ad

²²¹ Ad ogni modo, l'adozione di sistemi tecnici di controllo, pur rappresentando un elemento rilevante nella valutazione della diligenza degli *AI providers*, difficilmente è sufficiente a risolvere in maniera strutturale il conflitto fra modelli generativi e diritto d'autore: in tal senso MISHRA, *Understanding AI Guardrails: Concepts, Models, and Methods*, in *International Journal of Innovative Research*, 2025, disponibile online all'indirizzo <ijirmps.org>, 1-7.

²²² In prospettiva *de iure condendo*, l'incertezza che circonda tanto la qualificazione del *training* quanto la responsabilità da *output* corrobora la tesi – sostenuta, dal punto di vista dell'industria musicale, da C. GOLDSTUCK, *Past Precedent, Future Proof: Toward a New Legal and Commercial Framework for AI-Generated Music* (2025) disponibile online in SSRN- secondo cui la via maestra non

è quella giudiziale, bensì quella della composizione negoziata e di nuovi modelli di *licensing*. L'A. valorizza, in chiave ammonitrice, l'esperienza di *Viacom v. YouTube*: sette anni di causa e oltre cento milioni di dollari di spese legali per la sola Google, e circa duecento milioni complessivi, definiti nel 2014 con una transazione a saldo zero, a fronte di un sistema) Content ID e accordi di licenza (che nel medesimo arco temporale aveva già generato oltre un miliardo di dollari a favore dei titolari. Sotto questo profilo, la prospettiva di una responsabilità patrimoniale di Anthropic per il perdurante funzionamento del modello – lungi dal risolversi necessariamente in un giudizio di merito – potrebbe fungere da incentivo alla conclusione di un accordo di *licensing*, secondo un modello già collaudato nella parabola dello *streaming*.

Anthropic, delle richieste di lyrics formulate dagli utenti. Risultano attualmente pendenti contrapposte istanze di *partial summary judgement* sulla questione del *fair use*. Nel gennaio 2026, infine, alcune delle attrici hanno promosso un secondo e distinto giudizio dinanzi al medesimo distretto (N.D. Cal. No. 5:26-cv-00880), avente ad oggetto l'asserita riproduzione non autorizzata di spartiti e songbook relativi a oltre ventimila composizioni, con domande di *statutory damages* per oltre tre miliardi di dollari e con evocazione in giudizio, accanto alla società, anche dei co-fondatori: rimane da vedere come la Corte coordinerà i procedimenti.

IV. Conclusioni

Le controversie *Bartz* e *Concord* rappresentano, allo stato attuale, due dei più significativi banchi di prova per l'elaborazione di un quadro giuridico coerente in materia di intelligenza artificiale generativa e diritto d'autore. La loro lettura congiunta consente, ad avviso di chi scrive, di apprezzare la progressiva articolazione del dibattito su tre piani distinti, ma fra loro connessi.

Un primo piano riguarda l'utilizzazione trasformativa delle opere nel processo di addestramento. Sotto questo profilo, l'ordine del giudice Alsup nel caso *Bartz* costituisce la prima vera sistematizzazione giurisprudenziale del rapporto fra *fair use* e *training* algoritmico. Il carattere innovativo e computazionale del processo di addestramento viene riconosciuto, ma non trasformato in una generale immunità rispetto alle regole del *copyright*. L'ordine suggerisce, anzi, che la *doctrine of fair use* possa eventualmente coprire l'apprendimento algoritmico nel suo nucleo, senza però legittimare automaticamente la costituzione di archivi digitali di vaste dimensioni ottenuti mediante modalità illecite.

Un secondo piano, ancora in fase di elaborazione, concerne la responsabilità

derivante dagli *output* generati dai modelli. Il caso *Concord* mostra come il vero nodo del futuro contenzioso in materia di IA potrebbe non riguardare soltanto il *training*, ma soprattutto la capacità dei sistemi di rigenerare contenuti idonei a competere economicamente con le opere originarie. La questione assume particolare delicatezza nel caso dei testi musicali, i quali, in ragione della loro compattezza linguistica e dell'elevata riconoscibilità, possono risultare violati anche a fronte di modifiche marginali della struttura espressiva.

Un terzo piano, infine, concerne le modalità di acquisizione dei *dataset* destinati all'addestramento. Sotto questo profilo, l'ordine *Bartz* introduce una distinzione concettuale di considerevole rilievo, suscettibile di applicazione anche al di fuori dell'ambito specifico del *copyright* letterario. La liceità del ricorso a *shadow libraries* o, più in generale, a fonti non autorizzate, viene apertamente negata, e ciò indipendentemente dal carattere trasformativo dell'utilizzazione finale del materiale così acquisito.

Le decisioni qui esaminate mostrerebbero allora come il diritto si trovi progressivamente costretto ad abbandonare concezioni prestabilite della copia e della riproduzione digitale, suscettibili di rapida obsolescenza a fronte della sempre più rapida mutevolezza tecnologica. I *large language models* non operano come archivi statici di opere memorizzate, ma nemmeno possono essere assimilati a soggetti creativi autonomi nel senso tradizionale del termine²²³.

Essi costituiscono, piuttosto, strutture computazionali ibride, capaci di apprendere regolarità linguistiche astratte e, al tempo stesso, di produrre contenuti potenzialmente idonei a incidere direttamente sul mercato delle opere protette.

È pertanto verosimile che il futuro assetto della materia richieda non soltanto interpretazioni

²²³ In punto di soggettività giuridica delle IA si v. TASSONE, *Intelligenza artificiale, soggettività giuridica e personalità elettronica in prospettiva di comparazione*, in C. M. CASCIONE, G. GIANNONE CODIGLIONE, P. PARDOLESI, *Public and*

private in contemporary societies, Roma Tre e-press, 2024, 319 ss.; nonché ID. *Riflessioni su intelligenza artificiale e soggettività giuridica*, in *Riv. Dir. Internet*, 2023.

giurisprudenziali evolutive, ma anche interventi normativi specifici a livello sovranazionale, idonei a disciplinare in maniera organica il rapporto fra IA generativa e proprietà intellettuale.

Difatti, se nell'ordinamento interno il legislatore è intervenuto con l'art. 25 della L. n. 132/2025, nell'ordinamento statunitense lo U.S. Copyright Office – nel report, diffuso in versione di *pre-publication* nel maggio 2025, dedicato al *training* dell'IA generativa – ha reputato prematuro un intervento regolatorio, confidando nello sviluppo di mercati volontari di licenza²²⁴; del pari altri ordinamenti hanno reagito in modo diverso alla trasformazione in atto, mentre proprio la pervasività della trasformazione medesima – che non è soggetta ai confini statali – così come la forte armonizzazione a livello internazionale che caratterizza la materia autoriale imporrebbero l'adozione di principi condivisi.

In tale prospettiva, sembrano destinati ad assumere crescente importanza i sistemi di *licensing* collettivo, gli obblighi di trasparenza relativi ai *dataset* di addestramento, i meccanismi di remunerazione per i titolari dei diritti, nonché gli standard tecnici concernenti il controllo degli *output* generati dai modelli.

A tal proposito, occorrerà altresì osservare se gli obblighi imposti dall'art. 53 AI Act si ripercuoteranno al di fuori dei confini unionali secondo l'ormai noto effetto Bruxelles²²⁵.

Si tratta, in definitiva, di un cantiere ancora aperto, nel quale le decisioni *Bartz* e *Concord* si configurano non già come approdi, ma come prime occasioni in cui la giurisprudenza statunitense è chiamata a tracciare i confini fra apprendimento algoritmico, utilizzazione trasformativa delle opere e riproduzione illecita dei contenuti nell'era dell'intelligenza artificiale generativa.

²²⁴ In tal senso, la parte 3 - sul *training* - del Report in più parti pubblicato dall'Ufficio Copyright statunitense in tema di copyright e IA, ora disponibile in *pre-publication*. USCO, *Copyright and Artificial Intelligence, Part 3: Generative-AI Training (pre-publication)*, 9 maggio 2025, disponibile all'indirizzo <copyright.gov>.

V. I provvedimenti in commento

A) BARTZ V. ANTHROPIC PBC, ORDER ON FAIR USE DEL 23 GIUGNO 2025.

«INTRODUCTION

An artificial intelligence firm downloaded for free millions of copyrighted books in digital form from pirate sites on the internet. The firm also purchased copyrighted books (some overlapping with those acquired from the pirate sites), tore off the bindings, scanned every page, and stored them in digitized, searchable files. All the foregoing was done to amass a central library of “all the books in the world” to retain “forever.” From this central library, the AI firm selected various sets and subsets of digitized books to train various large language models under development to power its AI services. Some of these books were written by plaintiff authors, who now sue for copyright infringement. On summary judgment, the issue is the extent to which any of the uses of the works in question qualify as “fair uses” under Section 107 of the Copyright Act.

STATEMENT

Defendant Anthropic PBC is an AI software firm founded by former OpenAI employees in January 2021. Its core offering is an AI software service called Claude. When a user prompts Claude with text, Claude quickly responds with text — mimicking human reading and writing. Claude can do so because Anthropic trained Claude — or rather trained large language models or LLMs underlying various versions of Claude — using books and other texts selected from a central library Anthropic had assembled. Claude was first released publicly in March 2023. Seven successive versions of Claude have been released since. Users may ask Claude some questions for free. Demanding users and corporate clients pay to use Claude, generating over one billion dollars in annual revenue (Opp. Exh. 18). Plaintiffs Andrea Bartz, Charles Graeber, and Kirk Wallace Johnson are authors of books that Anthropic copied from pirated and purchased sources. Anthropic assembled these copies into a central library of its own, copied further various sets and subsets of those library copies to include in various “data mixes,” and used these mixes to train various LLMs. Anthropic kept the library copies in place as a permanent, general-purpose resource even after deciding it would not use certain copies to train LLMs or would never use them again to do so. All of Anthropic's copying was without plaintiffs' authorization. Author Bartz wrote four novels Anthropic copied and used: *The Lost Night: A Novel*, *The Herd*, *We*

²²⁵ L'effetto per cui le leggi emanate a Bruxelles si ripercuotano ben oltre i confini territoriali dell'Unione europea viene teorizzato da A. BRADFORD, *The Brussels Effect: How the European Union Rules the World*, Oxford, Oxford University Press, 2019.

Were Never Here, and The Spare Room. Author Graeber wrote two non-fiction books likewise at issue: *The Good Nurse: A True Story of Medicine, Madness, and Murder*, and *The Breakthrough: Immunotherapy and the Race to Cure Cancer*. And, Author Johnson penned three non-fiction books also copied and used: *To Be A Friend Is Fatal: The Fight to Save the Iraqis America Left Behind*, *The Feather Thief: Beauty, Obsession, and the Natural History Heist of the Century*, and *The Fishermen and the Dragon: Fear, Greed, and a Fight for Justice on the Gulf Coast*. Plaintiffs Bartz Inc. and MJ + KJ Inc. are corporate entities that Author Bartz and Author Johnson respectively set up to market their works. Between them, these five plaintiffs (“Authors”) own all the copyrights in the above-listed works. From the start, Anthropic “ha[d] many places from which” it could have purchased books, but it preferred to steal them to avoid “legal/practice/business slog,” as cofounder and chief executive officer Dario Amodei put it (see Opp. Exh. 27). So, in January or February 2021, another Anthropic cofounder, Ben Mann, downloaded Books3, an online library of 196,640 books that he knew had been assembled from unauthorized copies of copyrighted books — that is, pirated. Anthropic’s next pirated acquisitions involved downloading distributed, reshared copies of other pirate libraries. In June 2021, Mann downloaded in this way at least five million copies of books from Library Genesis, or LibGen, which he knew had been pirated. And, in July 2022, Anthropic likewise downloaded at least two million copies of books from the Pirate Library Mirror, or PiLiMi, which Anthropic knew had been pirated (Opp. Exh. 6 at 4; Opp. Expert Zhao ¶¶ 17–29; see Class Cert. (“CC”) Opp. Expert Iyer ¶¶ 45–46). Although what was downloaded and later duplicated from these sources was sometimes referred to as data or datasets, at bottom they contained full-text “ebooks or scans of books” saved in individual files in formats like .pdf, .txt, and .epub (see, e.g., Opp. Exh. 12 at 0391318). For Books3, most filenames identified the book inside. For LibGen and PiLiMi, Anthropic downloaded a separate catalog of bibliographic metadata for each collection, with fields like title, author, and ISBN (see, e.g., *ibid.*; Opp. Exh. 16 -0533972–73). Anthropic thereby pirated over seven million copies of books, including copies of at least two works at issue for each Author. As Anthropic trained successive LLMs, it became convinced that using books was the most cost-effective means to achieve a world-class LLM. During this time, however, Anthropic became “not so gung ho about” training on pirated books “for legal reasons” (Opp. Exh. 19). It kept them anyway (e.g., Opp. Exh. 17 at 93–94; CC Opp. Exh. 35 at -0273474). To find a new way to get books, in February 2024, Anthropic hired the former head of partnerships for Google’s book-scanning project, Tom Turvey. He was tasked with obtaining “all the books in the world” while still avoiding as much “legal/practice/business slog” as possible (Opp. Exhs. 21, 27). So, in spring 2024, Turvey sent an email or two to major publishers to inquire into licensing books for training AI. Had Turvey kept up those conversations, he

might have reached agreements to license copies for AI training from publishers — just as another major technology company soon did with one major publisher (e.g., Opp. Expert Malackowski ¶¶ 50, 64). But Turvey let those conversations wither. Instead, Turvey and his team emailed major book distributors and retailers about bulk purchasing their print copies for the AI firm’s “research library” (Opp. Exh. 22 at 145; Opp. Exh. 31 at -035589). Anthropic spent many millions of dollars to purchase millions of print books, often in used condition. Then, its service providers stripped the books from their bindings, cut their pages to size, and scanned the books into digital form — discarding the paper originals. Each print book resulted in a PDF copy containing images of the scanned pages with machine-readable text (including front and back cover scans for softcover books). Anthropic created its own catalog of bibliographic metadata for the books it was acquiring. It acquired copies of millions of books, including of all works at issue for all Authors. Anthropic may have copied portions of Authors’ books on other occasions, too — such as while copying book reviews, academic papers, internet blogposts, or the like for its central library. And, Anthropic’s scanning service providers may have copied Authors’ print books along the way to delivering the final digital copies to Anthropic. But neither side here specifically raises legal issues implicated by any such copies. Nor will this order. From all the above sources, Anthropic created a general “research library” or “generalized data area.” What was this for? As Turvey said, this was a “way of creating information that would be voluminous and that we would use for research,” or otherwise to “inform our — our products” (Opp. Exh. 22 at 145–46, 194). The copies were kept in the original “version of the underlying” book files Anthropic had “obtained or created,” that is, pirated or scanned (Opp. Exh. 30 at 3, 4). Anthropic planned to “store everything forever; we might separate out books into categories[, but t]here [wa]s no compelling reason to delete a book” — even if not used for training LLMs. Over time, Anthropic invested in building more tools for searching its “general purpose” library and for accessing books or sets of books for further uses (see CC Br. Exh. 12 at -0144509; CC Reply Exh. 45 at -0365931–32, -036593942 (reviewing and seeking to improve “[w]hat [] researchers do today if they want to search for a book,” including improving bibliographic metadata and consolidating varied resources)). One further use was training LLMs. As a preliminary step towards training, engineers browsed books and bibliographic metadata to learn what languages the books were written in, what subjects they concerned, whether they were by famous authors or not, and so on — sometimes by “open[ing] any of the books” and sometimes using software. From the library copies, engineers copied the sets or subsets of books they believed best for training and “iterate[d]” on those selections over time. For instance, two different subsets of print-sourced books were included in “data mixes” for training two different LLMs. Each was just a fraction of all

the print-sourced books. Similarly, different sets or “subsets” or “parts of” or “portions” of the collections sourced from Books3, LibGen, and PiLiMi were used to train different LLMs. Anthropic analyzed the consequences of using more books, fewer books, different books. The goal was to improve the “data mix” to improve each LLM and, ultimately, Claude’s performance for paying customers.

Over time, Anthropic came to value most highly for its data mixes books like the ones Authors had written, and it valued them because of the creative expressions they contained. Claude’s customers wanted Claude to write as accurately and as compellingly as Authors. So, it was best to train the LLMs underlying Claude on works just like the ones Authors had written, with well-curated facts, well-organized analyses, and captivating fictional narratives — above all with “good writing” of the kind “an editor would approve of” (Opp. Exh. 3 at -03433). Anthropic could have trained its LLMs without using such books or any books at all. That would have required spending more on, say, staff writers to create competing exemplars of good writing, engineers to revise bad exemplars into better ones, energy bills to power more rounds of training and fine-tuning, and so on. Having canonical texts to draw upon helped (e.g., Opp. Expert Zhao ¶ 81). Each work selected for training any given LLM was copied in four main ways — and in fact so many times that Anthropic admits it would be impractical even to estimate. First, each work selected was copied from the central library to create a working copy for the training set. Second, each work was cleaned to remove a small amount of lower-valued or repeating text (like headers, footers, or page numbers), with a “cleaned” copy resulting. If the same book appeared twice, or if while looking across the entire provisional training set it became clear there was some other reason to cull a book or category, Anthropic had the capability to delete relevant copy(ies) from the set at this step (see CC Br. Expert Zhao ¶¶ 71–72). Third, each cleaned copy was translated into a “tokenized” copy. Some words were “stemmed” or “lemmatized” into simpler forms (e.g., “studying” to “study”). And, all characters were grouped into short sequences and translated into corresponding number sequences or “tokens” according to an Anthropic-made dictionary. The resulting tokenized copies were then copied repeatedly during training. By one account, this process involved the iterative, trial-and-error discovery of contingent statistical relationships between each word fragment and all other word fragments both within any work and across trillions of word fragments from other copied books, copied websites, and the like. Other steps in training are not at issue here (id. ¶¶ 73–76; see Opp. Expert Zhao ¶ 38 & n.6). Fourth, each fully trained LLM itself retained “compressed” copies of the works it had trained upon, or so Authors contend and this order takes for granted. In essence, each LLM’s mapping of contingent relationships was so complete it mapped or indeed simply “memorized” the works it trained upon almost verbatim. So, if each

completed LLM had been asked to recite works it had trained upon, it could have done so (e.g., Opp. Expert Zhao ¶ 74). Further steps refining the LLM are not at issue here. However, that was as far as the training copies propagated towards the outside world. When each LLM was put into a public-facing version of Claude, it was complemented by other software that filtered user inputs to the LLM and filtered outputs from the LLM back to the user (id. ¶¶ 75–77). As a result, Authors do not allege that any infringing copy of their works was or would ever be provided to users by the Claude service. Yes, Claude could help less capable writers create works as well-written as Authors’ and competing in the same categories. But Claude created no exact copy, nor any substantial knock-off. Nothing traceable to Authors’ works. Such allegations are simply not part of plaintiffs’ amended complaint, nor in our record. Neither side puts directly at issue any copies of any works that might have been used for the filtering software. Nor will this order. In sum, the copies of books pirated or purchased-and-destructively-scanned were placed into a central “research library” or “generalized data area,” sets or subsets were copied again to create training copies for data mixes, the training copies were successively copied to be cleaned, tokenized, and compressed into any given trained LLM, and once trained an LLM did not output through Claude to the public any further copies. Finally, once Anthropic decided a copy of a pirated or scanned book in the library would not be used for training at all or ever again, Anthropic still retained that work as a “hard resource” for other uses or future uses. At least one work from each Author was present in every phase described above.

* * *

In August 2024, the three individual authors brought this putative class action complaining that Anthropic had infringed its federal copyrights by pirating copies for its library and by reproducing them to train its LLMs (Compl. ¶¶ 45–46, 71; see Amd. Compl. ¶¶ 47–48, 75). In October 2024, a scheduling order required that any motion for class certification be brought by March 6, 2025 (Dkt. No. 49). The individual authors soon amended their complaint to include affiliated corporate entities as named plaintiffs, with consent. And, Anthropic chose not to move to dismiss the amended complaint, as it earlier had planned (see Dkt. No. 37). Instead, Anthropic moved to allow an early motion for summary judgment on fair use, even before class certification (Dkt. No. 88; see Feb. 25, 2025 Tr. 15). Permission was granted. Anthropic now moves for summary judgment on fair use only. Fair use is a legal question for the judge with underlying fact questions, if any, for the jury. To prevail on summary judgment, Anthropic must rely on undisputed facts and/or factual inferences favoring the opposing side. Anthropic thus bears the burdens of production and persuasion in this motion. See *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 23–24 (2021); *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 547 n.21 (2023); *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590 & n.20, 594

(1994); see also *Nissan Fire & Marine Ins. Co. v. Fritz Cos.*, 210 F.3d 1099, 1102–03 (9th Cir. 2000). Notably, in its motion, Anthropic argues that pirating initial copies of Authors’ books and millions of other books was justified because all those copies were at least reasonably necessary for training LLMs — and yet Anthropic has resisted putting into the record what copies or even sets of copies were in fact used for training LLMs. For example, at oral argument, Anthropic asserted that if a purported fair user had retained pirated copies for uses beyond the fair use, then her piracy would not be excused by the fair use (Tr. 53, 56). But when Authors earlier interrogated Anthropic in discovery about what library copies (the original copies “obtained or created” by Anthropic) Anthropic had recopied for further uses, Anthropic responded that providing information about any copies made for uses beyond training commercially released LLMs would be overbroad, and that it could not count up all its copying even for LLMs in any case (e.g., Opp Exh. 30 at 3). We know that Anthropic has more information about what it in fact copied for training LLMs (or not). Anthropic earlier produced a spreadsheet that showed the composition of various data mixes used for training various LLMs — yet it clawed back that spreadsheet in April (Opp. Fredricks Decl. ¶¶ 2–3). A discovery dispute regarding that spreadsheet remains pending. But Anthropic did not need a court order to offer up what it possessed in support of its motion. All deficiencies must be held against Anthropic and not the other way around. This is the first substantive order in this case. A contemporaneous motion for class certification remains pending. It proposes one class related to works that were pirated (whether or not used to train LLMs), and a second class related to works that were purchased, scanned, and used in training LLMs. This order follows full briefing, a hearing, and supplemental briefing. To summarize the analysis that now follows, the use of the books at issue to train Claude and its precursors was exceedingly transformative and was a fair use under Section 107 of the Copyright Act. And, the digitization of the books purchased in print form by Anthropic was also a fair use but not for the same reason as applies to the training copies. Instead, it was a fair use because all Anthropic did was replace the print copies it had purchased for its central library with more convenient space-saving and searchable digital copies for its central library — without adding new copies, creating new works, or redistributing existing copies. However, Anthropic had no entitlement to use pirated copies for its central library. Creating a permanent, general-purpose library was not itself a fair use excusing Anthropic’s piracy.

ANALYSIS

Section 107 of the Copyright Act identifies four factors for determining whether a given use of a copyrighted work is a fair use:

[T]he fair use of a copyrighted work . . . for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or

research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include —

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- (2) the nature of the copyrighted work;
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- (4) the effect of the use upon the potential market for or value of the copyrighted work.

These factors presuppose a “use.” So, at the threshold, a court must decide whether a “copyrighted [work] has been used in multiple ways,” then evaluate each. *Warhol*, 598 U.S. at 533. Uses do not turn on “the subjective intent of the user” but on “an objective inquiry into what use was made, i.e., what the user d[id] with the original work.” *Id.* at 544–45. A “use” should be construed narrowly enough to not “swallow” distinguishable infringing uses, much less categories of exclusive rights in toto. *Id.* at 541, 543 n.18, 546–48. Sometimes, the challenged copying involves just one use: In *Perfect 10, Inc. v. Amazon.com, Inc.*, Google visited websites having full-sized images, made only reduced-sized copies, and incorporated those directly into its search engine — the sole use of the thumbnails being as “pointer[s]” to the images themselves. 508 F.3d 1146, 1157, 1160, 1165 (9th Cir. 2007). Sometimes, the copying involves many uses: In the Google Books cases, Google borrowed books from libraries, made both full-image and text-only copies, and incorporated different copies into different tools — one use being to reveal information “about those books,” another use being to provide the books to print-disabled patrons, and still another being to back up the print books if lost. *Authors Guild v. Google, Inc.*, 804 F.3d 202, 217 (2d Cir. 2015) (quoted); *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 97, 101, 103 (2d Cir. 2014) (other cited uses). Our parties debate an instructive decision. In *American Geophysical Union v. Texaco Inc.*, Texaco employees used scientific articles in a central library, used copies of them in personal desk libraries, and used selected copies again in the scientific laboratory — the first use paid for, the second infringing, and the third plausibly fair but in fact a rare occurrence. 802 F. Supp. 1, 4–5, 14 (S.D.N.Y. 1992) (Judge Pierre Leval), *aff’d*, 60 F.3d 913, 918–19, 926 (2d Cir. 1994).

Here, our parties contest what use or uses are at issue. Anthropic contends it copied Authors’ books only for one use: Only to train LLMs. By contrast, Authors contend it did so for at least two uses: First to build a vast, central library of potentially useful content, and second to train specific LLMs using shifting sets and subsets of that content — over time selecting the more well-organized and well-expressed works for training. Authors also complain that the print-to-digital format change was itself an infringement not abridged as a fair use (Opp. 15, 25). Authors do not allege, however, that any LLM outputs infringing upon their works ever reached users of the

public-facing Claude service. This order addresses each of the four factors in turn, pointing out how each applies to the training copies and to the purchased and pirated library copies. It concludes with an integrated analysis.

1. THE PURPOSE AND CHARACTER OF THE USE.

For a given use at issue, the first factor addresses “the purpose and character of th[at] use, including whether [it] is of a commercial nature or is for nonprofit educational purposes.” 17 U.S.C. § 107(1).

A. THE COPIES USED TO TRAIN SPECIFIC LLMs.

All agree that one use at issue was training LLMs to receive text inputs and return text outputs. More specifically, Anthropic used copies of Authors’ copyrighted works to iteratively map statistical relationships between every text-fragment and every sequence of text-fragments so that a completed LLM could receive new text inputs and return new text outputs as if it were a human reading prompts and writing responses. Authors further argue — and this order takes for granted — that such training entailed “memoriz[ing]” works by “compress[ing]” copies of those works into the LLM (Opp. 16–17; see Opp. Expert Zhao ¶ 74). The LLMs “memorize[d] A LOT, like A LOT” (Opp. Exh. 35 at -029109). Regardless, the “purpose and character” of using works to train LLMs was transformative — spectacularly so. To repeat and be clear: Authors do not allege that any LLM output provided to users infringed upon Authors’ works. Our record shows the opposite. Users interacted only with the Claude service, which placed additional software between the user and the underlying LLM to ensure that no infringing output ever reached the users. This was akin to the limits Google imposed on how many snippets of text from any one book could be seen by any one user through its Google Books service, preventing its search tool from devolving into a reading tool. Google, 804 F.2d at 222. Here, if the outputs seen by users had been infringing, Authors would have a different case. And, if the outputs were ever to become infringing, Authors could bring such a case. But that is not this case. Instead, Authors challenge only the inputs, not the outputs, of these LLMs. They point to the fully trained LLMs and the Claude service only to shed light on how training itself uses copies of their works and the ways the Claude service could be used to produce still other works that would compete with their works. This order does the same. Authors’ arguments that the training use is not transformative are unavailing. First, Authors argue that using works to train Claude’s underlying LLMs was like using works to train any person to read and write, so Authors should be able to exclude Anthropic from this use (Opp. 16). But Authors cannot rightly exclude anyone from using their works for training or learning as such. Everyone reads texts, too, then writes new texts. They may need to pay for getting their hands on a text in the first instance. But to make anyone pay specifically for the use of a book each time they read it, each time they recall it from memory, each time they later draw upon it when writing new things in new ways would be unthinkable. For centuries, we have read and re-read

books. We have admired, memorized, and internalized their sweeping themes, their substantive points, and their stylistic solutions to recurring writing problems. Second, to that last point, Authors further argue that the training was intended to memorize their works’ creative elements — not just their works’ non-protectable ones (Opp. 17). But this is the same argument. Again, Anthropic’s LLMs have not reproduced to the public a given work’s creative elements, nor even one author’s identifiable expressive style (assuming *arguendo* that these are even copyrightable). Yes, Claude has outputted grammar, composition, and style that the underlying LLM distilled from thousands of works. But if someone were to read all the modern-day classics because of their exceptional expression, memorize them, and then emulate a blend of their best writing, would that violate the Copyright Act? Of course not. Copyright does not extend to “method[s] of operation, concept[s], [or] principle[s]” “illustrated[] or embodied in [a] work.” 17 U.S.C. § 102(b); see, e.g., *Nichols v. Universal Pictures Corp.*, 45 F.2d 119, 120–22 (2d Cir. 1930) (Judge Learned Hand) (stage properties and storytelling elements); *Apple Comput., Inc. v. Microsoft Corp.*, 35 F.3d 1435, 1445 (9th Cir. 1994) (“user-friendly” design principles and elements); *Swirsky v. Carey*, 376 F.3d 841, 848 (9th Cir. 2004) (music theory principles and chord progressions). Third, Authors next argue that computers nonetheless should not be allowed to do what people do. Authors cite a decision seeming to say as much (Opp. 16–17). But the judge there twice emphasized while discussing “purpose and character” of the use that what was trained was “not generative AI (AI that writes new content itself).” Rather, what was trained — using a proprietary system for finding court opinions in response to a given legal topic — was a competing AI tool for finding court opinions in response to a given legal topic. That was not transformative. *Thomson Reuters Enter. Centre GmbH v. Ross Intell. Inc.*, 765 F. Supp. 3d 382, 398 (D. Del. 2025) (Judge Stephanos Bibas), appeal docketed, No. 25-8018 (3d Cir. Apr. 14, 2025). A better analogue to our facts would be an AI tool trained — using court opinions, and briefs, law review articles, and the like — to receive legal prompts and respond with fresh legal writing. And, on facts much like those, a different court came out the other way. It found fair use. *White v. W. Pub. Corp.*, 29 F. Supp. 3d 396, 400 (S.D.N.Y. 2014) (Judge Jed Rakoff). The latter use stood sufficiently “orthogonal” to anything that any copyright owner rightly could expect to control. See *Warhol*, 598 U.S. at 538–40. It could thus be freed up for the copyist to use, “promot[ing] the progress of science and the arts, without diminishing the incentive to create.” *Id.* at 531 (emphasis added); see U.S. CONST. art. I, § 8, cl. 8.

In short, the purpose and character of using copyrighted works to train LLMs to generate new text was quintessentially transformative. Like any reader aspiring to be a writer, Anthropic’s LLMs trained upon works not to race ahead and replicate or supplant them — but to turn a hard corner and create something different. If this training process reasonably required making copies within

the LLM or otherwise, those copies were engaged in a transformative use. The first factor favors fair use for the training copies. B. THE COPIES USED TO BUILD A CENTRAL LIBRARY. But that is not the only use at issue. Recall that Anthropic purchased millions of print books for its central library and pirated millions of digital books for its central library, too. It used specific sets and subsets of books for training specific LLMs. And, it then retained all the copies in its central library for other uses that might arise even after deciding it would not use them to train any LLM (at all or ever again). Anthropic seems to believe that because some of the works it copied were sometimes used in training LLMs, Anthropic was entitled to take for free all the works in the world and keep them forever with no further accounting. There is no carveout, however, from the Copyright Act for AI companies. Because the legal issues differ between the library copies Anthropic purchased and pirated, this order takes them in turn. (i) The Purchased Library Copies Converted from Print to Digital. Anthropic purchased millions of print copies to “build a research library” (Opp. Exh. 22 at 145, 148). It destroyed each print copy while replacing it with a digital copy for use in its library (not for sharing nor sale outside the company). As to these copies, Authors do not complain that Anthropic failed to pay to acquire a library copy. Authors only complain that Anthropic changed each copy’s format from print to digital (see Opp. 15, 25 & n.14). On the facts here, that format change itself added no new copies, eased storage and enabled searchability, and was not done for purposes trenching upon the copyright owner’s rightful interests — it was transformative. Anthropic purchased its print copies fair and square. With each purchase came entitlement for Anthropic to “dispose[]” each copy as it saw fit. 17 U.S.C. § 109(a). So, Anthropic was entitled to keep the copies in its central library for all the ordinary uses. Yes, Anthropic changed the format of these library copies from print to digital — giving rise to the issue here. All agree on the facts of the format change. Anthropic “destructively scan[ned]” the print copies to create the digital ones. Anthropic or its vendors stripped the bindings from the print books, cut the pages to workable dimensions, and scanned those pages — discarding each print copy while creating a digital one in its place. The digital copy was then housed in the “research library” or “generalized data area” in place of the print copy (Opp. Exh. 22 at 14546, 193–94). Authors do not allege and our record does not show that Anthropic provided its converted digital copies of print books to anyone outside Anthropic. The parties disagree about the legal consequences of the format change. Was scanning the print copies to create digital replacements transformative? Anthropic argues it was because it was reasonably necessary to training LLMs. Authors argue it was a distinguishable step requiring independent justification. Here, for reasons narrower than Anthropic offers, the mere format change was a fair use. Storage and searchability are not creative properties of the copyrighted work itself but physical properties of

the frame around the work or informational properties about the work. See *Texaco*, 802 F. Supp. at 14 (physical), *aff’d*, 60 F.3d at 919; *Google*, 804 F.3d at 225 (informational); *Sony Corp. of Am. v. Universal City Studios, Inc.* (“*Sony Betamax*”), 464 U.S. 417, 447 (1984) (rightful interests). In *Texaco*, the court reasoned that if a purchased scientific journal article had been copied “onto microfilm to conserve space, this might [have been] a persuasive transformative use.” 802 F. Supp. at 14 (Judge Pierre Leval), *aff’d*, 60 F.3d at 919 (reducing “bulk[]” “might suffice to tilt the first fair use factor in favor of *Texaco* if these purposes were dominant”). In *Google Books*, the court reasoned that a print-to-digital change to expose information about the work was transformative. *Google*, 804 F.3d at 225 (Judge Pierre Leval). And, in *Sony Betamax*, the Supreme Court held that making a recording of a television show in order to instead watch it at a later time was copying but did not usurp any rightful interest of the copyright owner. 464 U.S. at 447, 455. Important to the Supreme Court’s reasoning was the expectation that most such copiers would not distribute the permanent copies of the work. Finally, in *A&M Records, Inc. v. Napster, Inc.*, our court of appeals recognized the reasoning just explained, and therefore rejected by contrast a digitization effort that was touted as space-shifting but in fact resulted in the multiplication of copies shared with outsiders through a file-sharing service. 239 F.3d 1004, 1019 (9th Cir. 2001), *aff’g* in this part 114 F. Supp. 2d 896, 912–13, 915–16 (N.D. Cal. 2000) (Judge Marilyn Hall Patel) (citing *Sony Betamax* and *Texaco*). Here, every purchased print copy was copied in order to save storage space and to enable searchability as a digital copy. The print original was destroyed. One replaced the other. And, there is no evidence that the new, digital copy was shown, shared, or sold outside the company. This use was even more clearly transformative than those in *Texaco*, *Google*, and *Sony Betamax* (where the number of copies went up by at least one), and, of course, more transformative than those uses rejected in *Napster* (where the number went up by “millions” of copies shared for free with others). Yes, Anthropic is a commercial outfit. And, this order takes for granted that Anthropic in fact benefited from the print-to-digital format change — or it would not have gone to all the trouble. But the crux of the first fair use factor’s concern for “commercial” use is in protecting the copyright owners and their entitlements to exploit their copyright as they see fit (or not). See, e.g., *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 562 (1985). That the accused is a commercial entity is indicative, not dispositive. That the accused stands to benefit is likewise indicative. But what matters most is whether the format change exploits anything the Copyright Act reserves to the copyright owner. Anthropic already had purchased permanent library copies (print ones). It did not create new copies to share or sell outside. Yes, Authors also might have wished to charge Anthropic more for digital than for print copies. And, this order takes for granted that Authors could have

succeeded if Anthropic had been barred from the format change. “But the Constitution’s language [in Clause 8] nowhere suggests that [the copyright owner’s] limited exclusive right should include a right to divide markets or a concomitant right to charge different purchasers different prices for the same book, [merely] say to increase or to maximize gain.” See *Kirtsaeng v. John Wiley & Sons, Inc.*, 568 U.S. 519, 552 (2013); see also U.S. CONST. art. I., § 8, cl. 8. Nor does the Copyright Act itself. Section 106 sets out exclusive rights that fair uses under Section 107 abridge. Section 106(1) reserves to the copyright owner the right to make reproductions. But on our facts we face the unusual situation where one copy entirely replaced the another. And, Section 106(2) reserves to the copyright owner the right to make derivative works that add or subtract creative material — as occurs in a “translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, [or] condensation” of a book, 17 U.S.C. § 101 (definitions). For some “other modification[]” of a book to constitute a “derivative work,” it must itself “represent an original work of authorship.” *Ibid.* But on our facts the format was changed but no content was added or subtracted. See *Mirage Editions, Inc. v. Albuquerque A.R.T. Co.*, 856 F.2d 1341, 1342, 1343-44 (9th Cir. 1988) (yes where elements added to create new decorative ceramic).⁴ Section 106(3) further reserves to the copyright owner the right to distribute copies. But again, the replacement copy here was kept in the central library, not distributed. Cf. *Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169, 176–78 (2d Cir. 2018) (enabling searching for “information about the material” can be transformative use, even if some distribution results); *Lewis Galoob Toys, Inc. v. Nintendo of Am., Inc.*, 964 F.2d 965, 968, 971 (9th Cir. 1992) (using nifty converter to “merely enhance[]” audiovisual displays emitted from purchased videogame cartridge was fair use of those displays partly because no surplus copies of cartridge or displays were ever created). As a result, Anthropic’s format-change from print library copies to digital library copies was transformative under fair use factor one. Anthropic was entitled to retain a copy of these works in a print format. It retained them instead in a digital format, easing storage and searchability. And, the further copies made therefrom for purposes of training LLMs were themselves transformative for that further reason, as above. To be clear, this print-to-digital conversion involved a different and narrower form of transformative use than the broader one advanced by Anthropic. Anthropic argues that the central library use was part and parcel of the LLM training use and therefore transformative. This order disagrees. However, this order holds that the mere conversion of a print book to a digital file to save space and enable searchability was transformative for that reason alone. Therefore, the digital copy should be treated just as if the purchased print copy had been placed in the central library. In sum, the first fair use factor favors fair use for the digital library

copies converted from purchased print library copies — but these do not excuse the pirated library copies.

(ii) The Pirated Library Copies. Before buying books for its central library, Anthropic downloaded over seven million pirated copies of books, paid nothing, and kept these pirated copies in its library even after deciding it would not use them to train its AI (at all or ever again). Authors argue Anthropic should have paid for these pirated library copies (e.g., Tr. 24–25, 65; Opp. 7, 12–13). This order agrees. The basic problem here was well-stated by Anthropic at oral argument: “You can’t just bless yourself by saying I have a research purpose and, therefore, go and take any textbook you want. That would destroy the academic publishing market if that were the case” (Tr. 53). Of course, the person who purchases the textbook owes no further accounting for keeping the copy. But the person who copies the textbook from a pirate site has infringed already, full stop. This order further rejects Anthropic’s assumption that the use of the copies for a central library can be excused as fair use merely because some will eventually be used to train LLMs. This order doubts that any accused infringer could ever meet its burden of explaining why downloading source copies from pirate sites that it could have purchased or otherwise accessed lawfully was itself reasonably necessary to any subsequent fair use. There is no decision holding or requiring that pirating a book that could have been bought at a bookstore was reasonably necessary to writing a book review, conducting research on facts in the book, or creating an LLM. Such piracy of otherwise available copies is inherently, irredeemably infringing even if the pirated copies are immediately used for the transformative use and immediately discarded. But this order need not decide this case on that rule. Anthropic did not use these copies only for training its LLM. Indeed, it retained pirated copies even after deciding it would not use them or copies from them for training its LLMs ever again. They were acquired and retained, as a central library of all the books in the world. Building a central library of works to be available for any number of further uses was itself the use for which Anthropic acquired these copies. One further use was making further copies for training LLMs. But not every book Anthropic pirated was used to train LLMs. And, every pirated library copy was retained even if it was determined it would not be so used. Pirating copies to build a research library without paying for it, and to retain copies should they prove useful for one thing or another, was its own use — and not a transformative one (see Tr. 24–25, 35, 65; Opp. 4–10, 12 n.6; CC Br. Exh. 12 at - 0144509 (“everything forever”)). *Napster*, 239 F.3d at 1015; *BMG Music v. Gonzalez*, 430 F.3d 888, 890 (7th Cir. 2005). Anthropic’s briefing contains other reasons why it believes its pirated library copies are irrelevant to our fair use analysis, notwithstanding its own statements at our oral argument. First, Anthropic accepts in this posture that it acted in bad faith but argues that its bad faith in pirating copies cannot “somehow short-circuit[]” the fair use analysis (Reply 6 (downplaying *Atari Games Corp. v.*

Nintendo of Am., Inc., 975 F.2d 832, 843 (Fed. Cir. 1992) (applying law of Ninth Circuit))). But its bad faith is not the basis for this decision. Each use of a work must be analyzed objectively. *Warhol*, 598 U.S. at 544–45. The objective analysis here shows the initial copies were pirated to create a central, general-purpose library, as a substitute for paid copies to do the same thing. (Of course, if infringement is found, bad faith would matter for determining willfulness. 17 U.S.C. § 504(c)(2).) Second, Anthropic argues that its goal to put the copies eventually “to a highly transformative use” requires that each copy and use along the way be justified as having a transformative use, too (Reply 14). But now Anthropic seeks to take the shortcut Anthropic just said cannot be taken. Again, the Supreme Court tasks us with looking past the “subjective intent of the user” to the objective use made of each copy. *Warhol*, 598 U.S. at 544–45 (emphasis added). Put another way, what a copyist says or thinks or feels matters only to the extent it shows what a copyist in fact does with the work. Indeed, the same copy can be used one way, then another, each with a different result. *Id.* at 533. Here, what Anthropic said about its acquisitions at the time — that they were made to “build[] a research library” while avoiding a “huge legal/practice/business slog” — are relevant in this regard. And, Anthropic’s actual use of these pirated copies was to create its central library of texts that, like any university or corporate library, stored the works’ well-organized facts, analyses, and expressive examples for various contingent uses, one being training.⁵ Third, Anthropic argues that *Texaco* — the case involving copies used in a central library, copies used in desk libraries, and copies used in the laboratory — is inapposite. Anthropic argues that the disputed copies in *Texaco* were never used in the laboratory but instead in personal desk libraries for a use “identical to the original purpose and use” of the central library copies, and so not for a transformative use (Reply 8 (summarizing 60 F.3d at 922–23)). By contrast, says Anthropic, here it did use copies in the laboratory to train LLMs — a very transformative use. But this is a fast glide over thin ice. Like *Texaco*, Anthropic possessed copies it did not put into use in the laboratory and it kept those copies in a central library even after its transformative use had been completed. But, unlike *Texaco*, which bought those copies, Anthropic never paid for the central library copies stolen off the internet. *Texaco* also shows why Anthropic is wrong to suppose that so long as you create an exciting end product, every “back-end step, invisible to the public,” is excused (Br. 10). Notably, this is not a case where source copies were unavailable for separate purchase or loan. See, e.g., *NXIVM Corp. v. Ross Inst.*, 364 F.3d 471, 475–76, 478–79 (2d Cir. 2004) (using selections of training manual — otherwise available only to cult’s trainees subject to NDAs — to expose cult in critical review); *Time Inc. v. Bernard Geis Assocs.*, 293 F. Supp. 130, 135–36, 138, 146 (S.D.N.Y. 1968) (Judge Inzer Bass Wyatt) (making charcoal drawings of photographs taken of originals otherwise not on sale or

loan out to illustrate a history book).⁶ Nor were the copies made only incidentally and necessarily from pirated copies. See, e.g., *Perfect 10*, 508 F.3d at 1164 n.8 (copies of images that had been pirated by third-party websites were used to index those same websites while indexing the entire web). Here, piracy was the point: To build a central library that one could have paid for, just as Anthropic later did, but without paying for it. Nor were the initial copies made immediately transformed into a significantly altered form. In *Perfect 10*, images were copied by the search engine in thumbnail form only and deployed immediately into the transformative use of identifying the full-sized images and the pages from which they came. 508 F.3d at 1160, 1165, 1167. And, in *Kelly v. Arriba Software Corp.*, images were copied at full size and then into thumbnails for immediate use in building a search engine, after which the full-sized copies were immediately deleted. 336 F.3d 811, 815 (9th Cir. 2003). Not here. The full-text copies of books were downloaded and maintained “forever.” Nor does the initial copying here even resemble the full-text copying in the Google Books cases. There, libraries of authorized copies already had been assembled, and all copies therefrom were made for direct employment in a one-to-one further fair use — whether the transformative use of pointing to the works themselves, the use of providing the works in formats for print-disabled patrons, or the use of insuring against going out of print, getting lost, and becoming otherwise unavailable. *HathiTrust*, 755 F.3d at 97, 101, 103; *Google*, 804 F.3d at 206, 216–18, 228 (further distinguishing search and snippet uses, which “test[ed] the boundaries of fair use”). Not so here concerning the pirated copies. No authorized copies existed from which Anthropic made its first copies. No full-text copy therefrom was put immediately into use training LLMs. Not every copy was even necessary nor used for training LLMs. No initial copy was ever deleted, even if never used or no longer used.⁷ The university libraries and Google went to exceedingly great lengths to ensure that all copies were secured against unauthorized uses — both through technical measures and through legal agreements among all participants. Not so here. The library copies lacked internal controls limiting access and use. Nor do the decisions on intermediate copying require anything less than the analysis applied here. Anthropic argues that our court of appeals in *Sega Enterprises Ltd. v. Accolade, Inc.* looked only at the “ultimate use” and “did not analyze a series of atomized acts of ‘infringement’ distinct from that overall purpose” (Reply 3). To the contrary, the appeals court examined the initial, intermediate, and ultimate copies used by the copyist. The court explained that the copyist initially purchased commercially available copies of game cartridges and then made further copies necessarily and “solely in order to discover the functional requirements for compatibility.” 977 F.2d 1510, 1522 (9th Cir. 1992). Thus, it reached only one result because on those facts there was only one “overall purpose” for the unauthorized copies. Indeed, the court reaffirmed prior

caselaw holding that “intermediate copying of [a work] may infringe the exclusive rights granted to the copyright owner in [S]ection 106 of the Copyright Act regardless of whether the end product of the copying also infringes those rights.” *Id.* at 1518–19 (reaffirming *Walker v. Univ. Books*, 602 F.2d 859, 864 (9th Cir. 1979)). Similarly, in *Sony Computer Entertainment, Inc. v. Connectix Corp.*, our appeals court applied the same law to similarly focused conduct. Another copyist allegedly had purchased an authorized copy and then made further copies solely and necessarily to reverse-engineer compatibility requirements. 203 F.3d 596, 601, 602–03 (9th Cir. 2000). Both *Sega* and *Sony* avoided imposing an “artificial hurdle” to fair use by generously construing the intermediate copying necessary to the fair use. As one example, *Sega* stated that an engineer should be permitted to reboot her computer while undertaking to reverse engineer software loaded onto it — even if doing so creates another digital copy of the software and is not strictly necessary to reverse-engineering. *Id.* at 605. But neither *Sega* nor *Sony* fathomed gifting an “artificial head start” to a fair user, either, by treating even the initial copy as an intermediate one. And, yes, some courts have “not inquire[d]” into intermediate or initial copying at all (Reply 2 (citing *Campbell* as not inquiring into surplus copies in the studio)). But if a “close reading of those cases [] reveals that in none of them was the legality of the [initial or] intermediate copying at issue,” then it was not raised and not necessarily decided. *Sega*, 977 F.2d at 1519; see *Webster v. Fall*, 266 U.S. 507, 511 (1925). It was expressly decided elsewhere: Our analysis must attend to different uses of different copies, and even to different uses of the same copies. *Warhol*, 598 U.S. at 533. Finally, Anthropic argues that even if the initial copies served a different use than the intermediate and ultimate copies, it was not a use for which Anthropic necessarily would have needed to pay Authors for a copy. In theory, argues Anthropic, it could have done as Google did in *Google Books* — find an existing reference library willing to loan its copies for free as source copies. Or, in theory, it could have done as Anthropic did later — go buy used copies without having to pay Authors at all. See 17 U.S.C. § 109(a). But Anthropic did not do those things — instead it stole the works for its central library by downloading them from pirated libraries. In sum, the first factor points against fair use for the central library copies made from pirated sources — and no damages from pirating copies could be undone by later paying for copies of the same works.

2. THE NATURE OF THE COPYRIGHTED WORK. The second fair use factor is “the nature of the copyrighted work.” 17 U.S.C. § 107(2). This factor “calls for recognition that some works are closer to the core of intended copyright protection than others, with the consequence that fair use is more difficult to establish when the former works are copied.” *Campbell*, 510 U.S. at 586. For one thing, less protection is due published works than unpublished ones. For another, less protection is due “factual works than works of fiction or fantasy.” *Harper & Row*, 471 U.S. at

563. But less protection is not no protection. Even the arrangement of otherwise unprotectable facts surpasses the low bar for a protectable original work of authorship. *Google*, 804 F.3d at 220. Here, Anthropic accepts that all of Authors’ books — all published, whether non-fiction or fiction — contained expressive elements (Reply 9). And, as set out above, this order accepts Authors’ view of the evidence that their works were chosen for their expressive qualities in building a central library and then in training specific LLMs (Opp. 11, 17 (citing, e.g., Opp. Exh. 3 at -03433)). The main function of the second factor is to help assess the other factors: to reveal differences between the nature of the works at issue and the nature of their secondary use (above), and to reveal any relation between the amount and substantiality of each work taken and the secondary use (next). E.g., *Campbell*, 510 U.S. at 586; *Kelly*, 336 F.3d at 820; *Google*, 804 F.3d at 220; *HathiTrust*, 755 F.3d at 98; *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 612–13 (2d Cir. 2006). The second factor points against fair use for all copies alike.

3. THE AMOUNT AND SUBSTANTIALITY OF THE PORTION USED.

The third fair use factor is “the amount and substantiality of the portion” of the copyrighted work used by the accused. 17 U.S.C. § 107(3). The crux of this factor is whether the amount was “reasonable in relation to the purpose of the copying.” *Campbell*, 510 U.S. at 586. Thus, the amount of copying is considered first against the work itself, then more importantly against the proposed transformative purpose. See *Warhol*, 598 U.S. at 543 & n.18.

A. THE COPIES USED TO TRAIN SPECIFIC LLMs.

Copies selected for inclusion in training sets were selected because they were complete and because they contained rich protectable expression, or so this order accepts the record shows for Authors. Was all this copying reasonably necessary to the transformative use? Yes. “What matters [] is not so much ‘the amount and substantiality of the portion used’ in making a copy, but rather the amount and substantiality of what is thereby made accessible to a public [in the purported secondary use] for which it may serve as a competing substitute [for the primary use].” *Google*, 804 F.3d at 222. Here, once again, there is no allegation of any traceable connection between the Claude service’s outputs and Authors’ works. The copying used to train the LLMs underlying Claude was thus especially reasonable. In response, Authors object primarily that the copying used in training was both extremely extensive and not strictly necessary. As to extensive copying, it is true that entire works were copied. And, “copying [] entire work[s] ‘militate[s] against a finding of fair use.’” *Worldwide Church of God v. Philadelphia Church of God, Inc.*, 227 F.3d 1110, 1118 (9th Cir. 2000) (quoting *Hustler Mag. Inc. v. Moral Majority Inc.*, 796 F.2d 1148, 1155 (9th Cir. 1986)); see *Campbell*, 510 U.S. at 587. But we just addressed why Authors’ argument is misdirected. The copies that count for this factor are those that would merely serve the same use as

the work's ordinary one. Authors do not allege such copying. The accused use here of the incremental copies is as orthogonal as can be imagined to the ordinary use of a book.

As to strict necessity, Authors make a stronger point. When a productive use is made possible only by borrowing from a specific work, fair use climbs towards its zenith. When a productive use is possible without that borrowing, fair use falls to its nadir — and the borrowing deserves a particularly compelling justification. See *Warhol*, 598 U.S. at 543 & n.18, 547. Here, it is true that Anthropic could have used some other books or no books at all for training its LLMs — or so this order accepts the record shows for Authors. But Anthropic has presented a compelling explanation for why it was reasonably necessary to use them anyway. For one thing, all agree Anthropic needed billions of words to train any given LLM. If using only books, Anthropic would have needed millions of books per model. If using a set comprising only a small fraction of books and a larger fraction of other texts, Anthropic still would have needed hundreds of thousands of books. Authors contend that because Anthropic showed it could use such smaller sets of books, it surely could have used no books at all — or at least not their books (Opp. 23). But Authors forget that “reasonably necessary” does not mean “strictly necessary.” Authors do not contest that the volume of text required to train an LLM is monumental. Because using so many works was reasonably necessary, using any one work for actually training LLMs was about as reasonable as the next. For another thing, no output to the public was even alleged to be infringing. So, yes, Authors’ works were chosen as the strongest examples of writing. But the compelling benefits of training the LLMs on strong examples were not offset by revelations to the public of any portion of the works themselves. What was copied was therefore especially reasonable and compelling. The third factor thus favors fair use for the training copies.

B. THE COPIES USED TO BUILD A CENTRAL LIBRARY.

But again, there was a separate use — a distinction that makes some difference as to whether the amount and substantiality of the copying was “reasonable in relation to the purpose of the copying” for the library copies. *Campbell*, 510 U.S. at 586.

(i) The Purchased Library Copies Converted from Print to Digital.

For the print library copies that Anthropic purchased and then converted into digital library copies, Anthropic already enjoyed entitlement to keep the copies in its library. The purpose of the copying was to keep them in its library but with more favorable storage and searchability properties. Copying the entire work was exactly what this purpose required. There was no surplus copying. The source copy was destroyed. The third fair use factor favors fair use for the purchased library copies converted from print to digital.

(ii) The Pirated Library Copies.

For the pirated library copies, however, Anthropic lacked any entitlement to hold copies of the books at all. Its purpose, it says, was to train LLMs. But its objective conduct was to seek “all the books in the world” and then retain them even after deciding it would not make further copies from them for training — indicating there were other further uses. Against the purpose of acquiring all the books one could on the chance some might prove useful for training LLMs and maybe other stuff too, almost any unauthorized copying would have been too much. Anthropic copied millions of books in toto, Authors’ works among them. The third factor points against fair use for the pirated library copies.

4. THE EFFECT OF THE USE UPON THE MARKET FOR OR VALUE OF THE COPYRIGHTED WORK.

The final factor is “the effect of the use upon the potential market for or value of the copyrighted work.” 17 U.S.C. § 107(4). This factor points against fair use when a copyist makes copies available that displace demand for copies the copyright owner already makes available or readily could. *Texaco*, 60 F.3d at 926–28 (reproduced copies); *Dr. Seuss Enters., L.P. v. ComicMix LLC*, 983 F.3d 443, 461 (9th Cir. 2020) (derivative copies). “While the first factor considers whether and to what extent an original work and secondary use [in principle could] have substitutable purposes, the fourth factor focuses on actual or potential market substitution.” *Warhol*, 598 U.S. at 536 n.12 (emphasis added).

A. THE COPIES USED TO TRAIN SPECIFIC LLMs.

The copies used to train specific LLMs did not and will not displace demand for copies of Authors’ works, or not in the way that counts under the Copyright Act. Again, Authors concede that training LLMs did not result in any exact copies nor even infringing knockoffs of their works being provided to the public. If that were not so, this would be a different case. Authors remain free to bring that case in the future should such facts develop. Instead, Authors contend generically that training LLMs will result in an explosion of works competing with their works — such as by creating alternative summaries of factual events, alternative examples of compelling writing about fictional events, and so on. This order assumes that is so (Opp. 22–23 (citing, e.g., Opp. Exh. 38)). But Authors’ complaint is no different than it would be if they complained that training schoolchildren to write well would result in an explosion of competing works. This is not the kind of competitive or creative displacement that concerns the Copyright Act. The Act seeks to advance original works of authorship, not to protect authors against competition. *Sega*, 977 F.2d at 1523–24. Authors next contend that training LLMs displaced (or will) an emerging market for licensing their works for the narrow purpose of training LLMs (Opp. 21–22). Anthropic argues that transactional costs would exceed Anthropic’s expected benefit from any such bargain, prompting it to cease dealing with any rightsholders or else to cease developing such technology altogether (Br. 22–23). Our record could support either account — so this order must

assume Authors are correct. A market could develop (Opp. 19–21 (citing record)). Even so, such a market for that use is not one the Copyright Act entitles Authors to exploit. None of the cases cited by Authors requires a different result. All contemplated losses of something the Copyright Act properly protected — not the kinds of fair uses for which a copyright owner cannot rightly expect to control. See *TVEyes, Inc.*, 883 F.3d at 181 (use of a right legally reserved to and factually already being licensed by copyright owner); *Texaco*, 60 F.3d 931 (same); *Ringgold v. BET, Inc.*, 126 F.3d 70, 80–81 (2d Cir. 1997) (use of a right legally reserved to and factually likely to be marketable by copyright owner — displaying images of her artistic work in television shows); cf. *Seltzer v. Green Day, Inc.*, 725 F.3d 1170, 1179 (9th Cir. 2013) (no evidence use could be or “was likely to” be marketable). The fourth factor thus favors fair use for the training copies.

B. THE COPIES USED TO BUILD A CENTRAL LIBRARY.

(i) The Purchased Library Copies Converted from Print to Digital.

For these copies, this order assumes Anthropic’s format change from print to digital displaced purchases of new digital copies that Anthropic would have made directly from Authors (had it not been able to purchase print copies in used condition). But for reasons stated under the first factor, such losses did not relate to something the Copyright Act reserves for Authors to exploit. It was a format change. Authors’ next argument, it seems, is that the format change nonetheless exposed it to usurpation of the opportunity to sell rightful copies because Anthropic might transmit additional unauthorized digital copies more readily than it could have transmitted additional unauthorized print copies — and that the same would be true for all format converters (cf. Opp. 25 n.14; Opp. Expert Malackowski ¶ 52). But after much discovery, there is no inkling in our record of intent to redistribute library copies once acquired nor of inability to secure that valuable library against outside actors. And, if the internal, central library copies did or do in fact lead to further reproduction or distribution, those further copies remain redressable separately by Authors. The format change did not itself usurp the Authors’ rightful entitlements. This factor is thus neutral for the purchased library copies converted from print to digital.

(ii) The Pirated Library Copies.

The copies used to build a central library and that were obtained from pirated sources plainly displaced demand for Authors’ books — copy for copy. Not every person who merely intends to make a fair use of a work is thereby entitled to a full copy in the meantime, nor even to steal a copy so that achieving this fair use is especially simple or cost-effective. Here, the copies employed in training LLMs were one thing, but the copies acquired to assemble a convenient, general-purpose library of works for various uses for which the company might have of them, if any, was a different use altogether. Anthropic has almost no rebuttal on these points. First, Anthropic argues that “Claude’s services do not reduce [or usurp] the value of

Plaintiffs’ works through substitution in their traditional markets” (see Br. Expert Peterson ¶ 33). But stealing pirated copies of Authors’ works plainly did. Second, Anthropic argues that it may have been able to purchase some books on the open market (and some other texts), but not other texts it copied (cf. id. ¶ 48 (re licensing)). But this case does not concern those other texts it could not have purchased. It could have purchased Authors’ books (and many others). In fact it later did. Finally, Anthropic argues that the effect on these texts from one book foregone was too small to be considered (see id. ¶ 77). But the test requires that we contemplate the likely result were the conduct to be condoned as a fair use — namely to steal a work you could otherwise buy (a book, millions of books) so long as you at least loosely intend to make further copies for a purportedly transformative use (writing a book review with excerpts, training LLMs, etc.), without any accountability. As Anthropic itself suggested, “That would destroy the [entire] publishing market if that were the case” (see Tr. 53; see also Tr. 32, 41; Opp. Expert Malackowski ¶¶ 31–34, 38). The fourth factor points against fair use for the pirated library copies.

5. OVERALL ANALYSIS.

After the four factors and any others deemed relevant are “explored, [] the results [are] weighed together, in light of the purposes of copyright.” *Campbell*, 510 U.S. at 578. The copies used to train specific LLMs were justified as a fair use. Every factor but the nature of the copyrighted work favors this result. The technology at issue was among the most transformative many of us will see in our lifetimes. The copies used to convert purchased print library copies into digital library copies were justified, too, though for a different fair use. The first factor strongly favors this result, and the third favors it, too. The fourth is neutral. Only the second slightly disfavors it. On balance, as the purchased print copy was destroyed and its digital replacement not redistributed, this was a fair use. The downloaded pirated copies used to build a central library were not justified by a fair use. Every factor points against fair use. Anthropic employees said copies of works (pirated ones, too) would be retained “forever” for “general purpose” even after Anthropic determined they would never be used for training LLMs. A separate justification was required for each use. None is even offered here except for Anthropic’s pocketbook and convenience. And, as for any copies made from central library copies but not used for training, this order does not grant summary judgment for Anthropic. On this record in this posture, the central library copies were retained even when no longer serving as sources for training copies, “hundreds of engineers” could access them to make copies for other uses, and engineers did make other copies. Anthropic has dodged discovery on these points (e.g., Opp. Exh. 17 at 9394 (retained); Opp. Exh. 22 at 196 (no limits); Opp. Exh. 30 at 3, 4 (no accounting); see also Opp. 15). We cannot determine the right answer concerning such copies because the record is too poorly developed as to them. Anthropic is not entitled to an

order blessing all copying “that Anthropic has ever made after obtaining the data,” to use its words (Opp. Exh. 30 at 3, 4).

CONCLUSION

With respect to the training copies and the print-to-digital converted copies, this order has drawn all ambiguities and inferences in favor of the opposing side, namely Authors. With respect to the pirated copies, this order has also accepted the Authors’ version of the facts. Authors did not move for summary judgment but if they had, then we would have been obligated to accept all reasonable views given the evidence in defendant’s favor instead. This order grants summary judgment for Anthropic that the training use was a fair use. And, it grants that the print-to-digital format change was a fair use for a different reason. But it denies summary judgment for Anthropic that the pirated library copies must be treated as training copies. We will have a trial on the pirated copies used to create Anthropic’s central library and the resulting damages, actual or statutory (including for willfulness). That Anthropic later bought a copy of a book it earlier stole off the internet will not absolve it of liability for the theft but it may affect the extent of statutory damages. Nothing is foreclosed as to any other copies flowing from library copies for uses other than for training LLMs.

IT IS SO ORDERED».

**B) CONCORD MUSIC GROUP, INC. ET AL.
V. ANTHROPIC PBC
B.1) ORDER DENYING MOTION FOR
PRELIMINARY INJUNCTION DEL 25 MARZO
2025.**

«[...]

This action arises out of the Defendant's alleged use of Plaintiffs' copyrighted song lyrics to train a generative artificial intelligence ("AI") model. Plaintiffs, eight music publishing companies who own or control exclusive rights to millions of musical compositions, seek a preliminary injunction that prohibits Defendant from using the copyrighted lyrics for training purposes. Mot. for Prelim. Inj., ECF No. 179 ("Mot."). The Court reviewed the briefs and supporting declarations and exhibits, and heard argument on November 25, 2024. For the reasons discussed below, the motion is DENIED.¹ On February 11, 2025, Plaintiffs filed an administrative request to file supplemental authority. ECF No. 295 (submitting the opinion, Thomson Reuters Enter. Centre GmbH v. ROSS Intelligence, Inc., No. 20-cv-613-SB, 2025 WL 458520 (D. Del. Feb. 11, 2025), for consideration). The Court reviewed the subject opinion and notes that it is distinguishable because (1) it addressed the merits of the parties' respective infringement claims and defenses at the summary judgment stage; (2) it did not concern a generative AI model; and (3) the parties in that case were direct competitors. The Court does not rely on the opinion in making its ruling.

I. BACKGROUND A.

Factual Background Plaintiffs, hereinafter referred to as "Publishers," are Concord Music Group, Inc. ("Concord"); Capitol CMG, Inc. ("CCMG"); Universal Music Corp., Songs of Universal, Inc., Universal Music - MGB NA LLC, Polygram Publishing, Inc., Universal Music - Z Tunes LLC (collectively, "UMPG Plaintiffs"); and ABKCO Music, Inc. ("ABKCO"). Publishers own or control the exclusive rights to millions of musical compositions (the "Works"). Mot. at 2. Publishers identify 500 examples of the Works in Exhibit A to the complaint. Compl. Ex. A, ECF No. 1. Publishers license the Works to digital music services, social media platforms, and lyrics aggregators. Mot. at 3. Publishers depend on revenues from the licenses to support and promote songwriters, and songwriters depend on licensing royalties for their livelihoods. Id. Defendant Anthropic PBC ("Anthropic") is a technology company whose signature product line is "Claude," a series of general-purpose AI large language models ("LLMs"). Decl. of Jared Kaplan ¶ 8, ECF No. 209 ("Kaplan Decl."). LLMs, including Claude, are trained by creating a "neural network," a computer program that can study large datasets (referred to as the "corpus") "consisting of hundreds of millions, if not billions, of pieces of content." Id. ¶¶ 17, 18. An LLM's training corpus may be assembled from information publicly available on the Internet, non-public data from third parties, or internally created data.

Id. ¶ 28. The corpus is broken down into "tokens," which may be words or characters. Id. ¶ 19. The tokens are grouped into sequences, shuffled, and analyzed by the neural network in the context of surrounding tokens in order to learn language patterns and relationships between words and phrases. Id. ¶¶ 19-20. After Claude learns statistical language patterns, Anthropic fine-tunes it to adhere to a set of principles that train Claude to be "helpful" and "harmless." Id. ¶¶ 38-39. "Helpful" means that Claude will answer questions and assist with tasks; "harmless" means that Claude will not do or say things that people find dangerous or harmful. Id. ¶ 39. According to Anthropic, Claude is designed to use the learning from its training to generate original content. Id. ¶ 9. 3 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 United States District Court Northern District of California Publishers allege that Anthropic's training corpus includes the lyrics to many of the Works. Mot. at 5. Publishers contend that when prompted, Claude has responded with output containing "verbatim or near-verbatim copies of the Works." Id. at 5-6. Publishers argue that by using the Works to train Claude, and reproducing the Works in Claude's output, Anthropic violates their rights under the Copyright Act, causing "incalculable harm." Id. at 8. Anthropic does not meaningfully dispute that Claude's training corpus includes the Works. See Kaplan Decl. ¶ 31 ("If [the Works] were included in the training set, it is likely because they are commonly found all across the Internet[.]"). Anthropic asserts, however, that Claude's intended purpose is not to reproduce existing materials in response to user queries, but "to generate original outputs that meet the full range of customer needs." Id. ¶ 15. B. Procedural Background On October 18, 2023, Publishers commenced this action in the U.S. District Court for the Middle District of Tennessee by filing a complaint asserting claims for direct copyright infringement, contributory infringement, vicarious infringement, and removal or alteration of copyright management information. ECF No. 1. On November 16, 2023, Publishers moved for a preliminary injunction. ECF No. 40. On November 22, 2023, Anthropic moved to dismiss the case for lack of personal jurisdiction or, in the alternative, to transfer venue. ECF No. 54. On June 24, 2024, the court granted Anthropic's motion in part, and transferred the case to the Northern District of California without ruling on the preliminary injunction motion. ECF No. 124. On August 1, 2024, Publishers filed a renewed motion for preliminary injunction in this district. ECF No. 179. On August 15, 2024, Anthropic filed a motion to dismiss Counts II, III, and IV under Federal Rule of Civil Procedure 12(b)(6).² ECF No. 205. On August 21, 2024, this case was reassigned to the undersigned. On November 25, 2024, the Court heard argument on the motion for preliminary injunction. ECF No. 272. On December 30, 2024, the parties submitted a stipulation and proposed order regarding the preliminary injunction, ECF No. 290, which the Court entered on January 2, 2025, ECF No. 291. 2 The Court heard argument on Anthropic's

Rule 12(b)(6) motion on December 19, 2024, and took the matter under submission. ECF No. 281. 4 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 United States District Court Northern District of California

II. LEGAL STANDARD

“A preliminary injunction is an extraordinary remedy never awarded as of right.” *Winter v. Nat. Res. Def. Council, Inc.*, 555 U.S. 7, 24 (2008). “A plaintiff seeking a preliminary injunction must establish that he is likely to succeed on the merits, that he is likely to suffer irreparable harm in the absence of preliminary relief, that the balance of equities tips in his favor, and that an injunction is in the public interest.” *Id.* at 20. These four factors are often referred to as the Winter factors, and all must be satisfied. *Id.* Under the Ninth Circuit’s “sliding scale approach,” a preliminary injunction may issue where “serious questions going to the merits were raised and the balance of hardships tips sharply in [the movant’s] favor.” *Alliance for the Wild Rockies v. Cottrell*, 632 F.3d 1127, 1131 (9th Cir. 2011). “[I]n a copyright infringement case, the plaintiff must demonstrate a likelihood of irreparable harm as a prerequisite for injunctive relief, whether preliminary or permanent.” *Flexible Lifeline Sys., Inc. v. Precision Lift, Inc.*, 654 F.3d 989, 998 (9th Cir. 2011). The party seeking injunctive relief has the burden of demonstrating that it has met the necessary elements. *Klein v. City of San Clemente*, 584 F.3d 1196, 1201 (9th Cir. 2009).

III. DISCUSSION This case involves alleged copyright infringement based on Claude’s “input” (i.e., the use of the Works as an “input” in its training corpus), and on Claude’s “output” (i.e., the responses it generates to user prompts). Publishers sought preliminary injunctive relief to address both forms of alleged infringement. With respect to output-based infringement, this issue has been resolved for purposes of this motion by the parties’ recent stipulation. ECF No. 291. Publishers originally sought a preliminary injunction ordering Anthropic to “maintain effective guardrails that prevent its current and future AI models from generating output that reproduces, distributes, or displays the lyrics, in whole or in part, to compositions owned or controlled by Publishers . . . or creates derivative works using those lyrics.” See Prop. Order at 3, ECF No. 179-1. In the joint stipulation entered by the Court on January 2, 2025, Anthropic has agreed to maintain already-implemented guardrails as to its current AI models and products. ECF No. 291. For new LLMs and products, 5 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 United States District Court Northern District of California Anthropic will apply guardrails on input and output in a manner consistent with the already implemented guardrails. *Id.* In light of the stipulation, this Order focuses solely on the relief related to the alleged input-based infringement – that is, Anthropic’s use of unauthorized copies of the Works for training. In that regard, Publishers seek to restrain Anthropic “from using copies of lyrics (or portions of lyrics) to compositions owned or controlled by Publishers for

future training of Anthropic’s AI models (unless expressly authorized to do so by Publishers via license or other written agreement).” Prop. Order at 3. A. The Nature and Scope of the Requested Relief The nature and scope of the requested relief has shifted over the course of these proceedings. Initially, it was unclear whether an injunction prohibiting “future training of Anthropic’s AI models” would include models that are currently in development or that have already been released. The difference in scope – and possibly the nature of the remedy – is significant: An injunction requiring Anthropic to retrain already-released models, or to rebuild the training corpus for models in development, could impose unforeseeable costs on Anthropic. Kaplan Decl. ¶ 61. In addition, it was unclear which lyrics Anthropic would be prohibited from using. In its opposition, Anthropic raised concerns that it would be forced to change its practices, resulting in large costs and causing “incalculab[le]” harm in the broader marketplace. Opp. to Mot. at 9, 26-27, ECF No. 207 (“Opp.”). In reply, Publishers stated that they “are not asking that Anthropic retrain or extract data from its existing or in-progress AI models.” Reply in Support of Mot. at 13, ECF No. 225 (“Reply”). At hearing, Publishers further clarified that they “are not asking Anthropic to have to retrain existing models” or “to pull models out of the marketplace,” and that the requested injunction would “not include any models that are currently under development.” Hr’g Tr. at 23-24, ECF No. 276. Although Publishers have tried to clarify the scope of the proposed injunction, the details remain elusive and poorly defined. The complaint incorporates and references 500 songs listed in Exhibit A, but this list is “illustrative and non-exhaustive.” Compl. Ex. A; see also *id.* ¶ 37. At hearing, Publishers confirmed that the injunction would extend “to all of Publishers’ works,” which number in the “hundreds of thousands at least.” Hr’g Tr. at 29-32. Publishers’ counsel could not say how many songs would be subject to the injunction. *Id.* at 31. Moreover, the injunction would apply to songs that Publishers currently own and to an unknowable universe of songs that they may acquire while the injunction is in place. *Id.* at 30-31. Publishers would also “update” the list to add or remove songs as necessary. *Id.* at 30. Publishers did not offer a concrete or definitive way for Anthropic – as the party subject to the injunction and the legal repercussions of a violation – to ascertain its parameters or comply with its terms.³ The enormous and seemingly ever-expanding scope of Works included in the requested injunction raises significant concerns regarding enforceability and manageability. According to Anthropic, excluding an undefined amount of unknown material from its training corpus would be “virtually impossible,” and would require Anthropic to “have to undertake constant efforts to update the corpus, and restart the future model’s training process, potentially out of cycle, every time [Publishers] ‘update’ their catalogs.” Kaplan Decl. ¶ 65. “If the expectation is that Anthropic is to re-start the training process from scratch every quarter as [Publishers] update their works,

as suggested in the proposed order,” Anthropic’s ability to release any new models would be severely curtailed by the length of its training process. *Id.* Thus, even if the Court charged Publishers with “policing the system within the limits of the system,” *A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004, 1027 (9th Cir. 2001), given the unknown number of Works involved, there is no manageable process through which the Court could enjoin the use of the Works for training purposes without unduly burdening Anthropic. See *Califano v. Yamasaki*, 442 U.S. 682, 702 (1979) (acknowledging rule that injunctive relief “should be no more burdensome to the defendant than necessary to provide complete relief to the plaintiffs”). At hearing, Publishers did not squarely respond to the Court’s concern about 3 Publishers argue that the requested relief is not overbroad, citing to *Apple Inc. v. Psystar Corp.*, 658 F.3d 1150, 1161 (9th Cir. 2011). In that summary judgment opinion, the Ninth Circuit affirmed a permanent injunction prohibiting future infringement of copyrighted works that belonged to Apple but which were not at issue in that litigation. There, the injunction including the additional works was appropriate “because liability has been established as between the parties[.]” Here, by contrast, the Court has not decided the merits of Publishers’ infringement claims or Anthropic’s fair use defense. managing such a vague and unwieldy injunction. See *Hr’g Tr.* at 30-31. The undefined nature of the relief sought here casts a long shadow over Publishers’ request. See *Trump v. Int’l Refugee Assistance Project*, 582 U.S. 571, 579 (2017) (“Crafting a preliminary injunction is an exercise of discretion and judgment, often dependent as much on the equities of a given case as the substance of the legal issues it presents.”). The Court does not have unfettered discretion to issue an injunction that may be overbroad because it includes works for which Publishers may not have established ownership or control. *United States v. BNS, Inc.*, 858 F.2d 456, 460 (9th Cir. 1988) (concluding that district court’s injunction was an abuse of discretion because it “was overbroad under the circumstances”); see also *Religious Tech. Ctr. v. Netcom On-Line Comm’n Servs., Inc.*, 923 F. Supp. 1231, 1259 (N.D. Cal. 1995) (declining to expand TRO to include additional works protected by copyright and trade secret law, noting that “a broad injunction which goes beyond the scope of the allegedly infringing activities should be avoided for First Amendment reasons”). Even if the Court were inclined to issue an injunction that is limited in scope to the 500 Works identified in the complaint, Publishers have not requested this narrower form of relief, and at hearing Publishers insisted that the injunction apply “to all of [their] works.” *Hr’g Tr.* at 30. The Court therefore finds it inappropriate to issue a more limited injunction. The Court turns next to the *Winter* factors, and also finds an injunction inappropriate on the independent ground that Publishers have not shown irreparable harm on the record presented. B. Irreparable Harm A showing of irreparable harm is a prerequisite for injunctive relief. *Flexible Lifeline Sys.*, 654 F.3d at 998. Irreparable harm must be “likely,”

not just possible. *Winter*, 555 U.S. at 22. “A plaintiff must do more than merely allege imminent harm sufficient to establish standing; a plaintiff must demonstrate immediate threatened injury as a prerequisite to preliminary injunctive relief.” *Boardman v. Pac. Seafood Grp.*, 822 F.3d 1011, 1022 (9th Cir. 2016) (quoting *Caribbean Marine Servs. Co. v. Baldrige*, 844 F.2d 668, 674 (9th Cir.1988)). Publishers contend they will suffer two primary types of irreparable harm should Anthropic continue to use the copyrighted works: reputational harm and market-related harm. 8 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 United States District Court Northern District of California The Court addresses each in turn. 1. Reputational Harm Publishers contend that Anthropic’s ongoing use of the copyrighted works will cause them to suffer reputational harm by depriving them of control over the works, denying them credit and goodwill associated with the works, and harming their reputations through the creation of unauthorized derivative works. *Mot.* at 23-25. A loss of control over business reputation or damage to customer goodwill “could constitute irreparable harm” where supported by evidence. *Herb Reed Enters., LLC v. Fla. Ent. Mgmt., Inc.*, 736 F.3d 1239, 1250 (9th Cir. 2013). For the reasons discussed below, Publishers have not made this showing here. The Court notes that the reputational harm described by Publishers appears largely related to Claude’s outputs rather than the use of the Works for training purposes alone. See *Mot.* at 24- 25. For example, Publishers claim that “Claude routinely reproduces, distributes, and displays Publishers’ copyrighted lyrics in its output but fails to identify Publishers or their songwriters as the source.” *Id.* And “Claude frequently generates output that combines portions of the Works with other lyrics or text, often in ways inconsistent with and inimical to authorial intent – and for which the songwriter would never have granted a license.” *Id.* at 25. The parties’ stipulation addressed and resolved Publishers’ motion for injunctive relief to the extent it concerns Claude’s output. Publishers have not demonstrated reputational harm based on the use of the Works as training input. Publishers cite to several declarations, which are largely duplicative of each other. The declarations discuss the effect of unlicensed use of copyrighted works as “damaging” or “harmful” generally, but they fail to identify any specific harm. For example, Concord’s Executive Vice President of Business and Legal Affairs and Publishing states that Anthropic’s conduct denies Concord and its songwriters “the right and ability to control the exploitation of their copyrighted works,” and that this “loss of control is enormously damaging.” *Decl. of Duff Berschback* ¶ 21, ECF No. 184 (“Concord Decl.”). 4 Although the Court acknowledges the value 4 Other declarations submitted by Publishers echo these general sentiments. ABKCO’s Chief Operative Officer states that when Anthropic “engages in unauthorized uses of the ABKCO 9 an artist places on control of their own works, none of the declarations explain how or why the

use of the Works to train Claude has been, or will be, “enormously damaging” to Publishers or songwriters. To a certain extent, all copyright infringement involves a loss of control. But accepting Publishers’ argument on this record would imply automatic entitlement to injunctive relief in all copyright cases, a premise the Ninth Circuit has rejected. See *Perfect 10, Inc. v. Google, Inc.*, 653 F.3d 976, 980 (9th Cir. 2011) (“[T]he propriety of injunctive relief in cases arising under the Copyright Act must be evaluated on a case-by-case basis in accord with traditional equitable principles and without the aid of presumptions[.]”). Publishers also cite a statement from the Board of Directors of the Nashville Songwriters Association International (“NSAI”). Decl. of Bart Herbison Ex. A, ECF No. 46-1 (“NSAI Letter”). There, NSAI Board members expressed concern that by creating new compositions using copyrighted lyrics, or by turning “important, award-winning, socially meaningful songs” into “something the composers never intended,” Claude will diminish the value of the works or result in a “misuse of [songwriters’] art.” Id. The NSAI letter, which cites several examples of output where Claude has changed lyrics or created mashups of certain Works, relies on Appendix E to the Declaration of Dr. Robert Leonard. See id. Ex. E, ECF No. 46-1. The Claude outputs discussed in the Leonard Declaration dated November 16, 2023, pre-date Publishers’ filing of the original preliminary injunction motion. See ECF No. 50. Accordingly, this and other outputs were addressed by the parties’ stipulation, and do not provide a basis for injunctive relief. 5 A preliminary injunction must be “grounded in evidence,” not based on what harm Publishers “might suffer.” *Herb Reed Enters., LLC*, 736 F.3d at 1250 (holding that district court erred in issuing preliminary injunction without sufficient evidence of irreparable harm). Because Works and exploits those works to its own ends without permission,” it “wrests . . . control away from ABKCO and our songwriters” and deprives them of the ability to decide how the works are used “in ways that are enormously damaging and cannot be undone.” Decl. of Alisa Coleman ¶ 23, ECF No. 185 (“ABKCO Decl.”); see also Decl. of David Kokakis ¶¶ 13-16, 25, ECF No. 187 (“UMPG Decl.”); Decl. of Kenton Draughon ¶¶ 13-16, 22, ECF No. 186 (“CCMG Decl.”). 5 Moreover, Publishers’ theory that mashups or other similar outputs may cause irreparable harm if they offend songwriters would essentially require the Court to engage in a hypothetical fair use analysis as to each example. The Court declines to decide such questions in the hypothetical. Publishers’ evidence does not demonstrate a likelihood of irreparable harm arising from reputational harm, Publishers are not entitled to a preliminary injunction on this basis. 2. Market-Related Harm Publishers also contend that Anthropic’s use of copyrighted lyrics will erode the value of the Works by undermining the licensing market, damaging Publishers’ position to negotiate future training licenses with AI developers, and harming Publishers’ relationships with songwriters. Mot. at 26-28. In effect, Publishers argue that

output containing the Works and the use of the Works to train Claude affects Publishers’ position both (1) in the existing licensing market and (2) in the emerging market for licensing copyrighted works for use by AI developers. Id. at 26. The Court addresses each of these in turn. As to the existing licensing market, Publishers argue that Anthropic’s unlicensed use of the Works reduces demand for licenses for the lyrics, and undermines Publishers’ ability to negotiate new licenses and to renew existing licenses. Mot. at 26. Publishers have not submitted any evidence that Anthropic’s use of lyrics to train Claude reduces license fees with lyric aggregators, lyric websites, or other existing licensees, which provide entirely different services and do not compete with Claude. To the extent Publishers are arguing that lyrics appearing in Claude’s output undermines their ability to license to lyrics aggregators, this potential harm is addressed by the parties’ stipulation. In addition, Anthropic vigorously disputes that Claude could be used as an alternative to lyrics aggregators, as Claude was designed precisely not to perform these functions. See Opp. at 2 (“Anthropic’s tool is not designed to output copyrighted material, and Anthropic has always had guardrails in place aimed at preventing this result. If those measures failed in some instances many months ago, that would have been a ‘bug,’ not a ‘feature,’ of the product.”). The evidence submitted by Publishers in the form of declarations by Publishers’ representatives are largely duplicative of each other, and state in a general and conclusory manner that Anthropic’s use of the Works is harmful. But they do not demonstrate how using the Works to train Claude is affecting – let alone diminishing – the value of any of the Works. Nor do they show that training Claude with the Works is harming Publishers’ negotiating position vis-à-vis new or existing licensees like lyrics aggregators. See Decl. of Michael D. Smith ¶¶ 48-49, ECF No. 182 (“Smith Decl.”); Concord Decl. ¶¶ 19-27; ABKCO Decl. ¶¶ 21-29; CCMG Decl. ¶¶ 20-28; UMPG Decl. ¶¶ 20-31. For example, Publishers have not identified any lost licensing deal or any licensing arrangement that Publishers have had to renegotiate on less favorable terms. Even if Publishers’ evidence demonstrated that Anthropic’s use of the Works to train Claude diminished Publishers’ position in the existing licensing market, Publishers fail to demonstrate that their losses could not be compensated through monetary damages. See *Fox Broad. Co. v. Dish Network LLC*, 747 F.3d 1060, 1073 (9th Cir. 2014) (denying preliminary injunction because Fox’s harm was not “irreparable,” as it could be compensated monetarily through existing licensing agreements that “could, at the very least, constitute a starting point or an aid in calculating damages”). Thus, Publishers do not demonstrate irreparable harm on this basis. As to the emerging market related to AI developers, Publishers argue that Anthropic’s use of the Works will negatively impact the emerging market for AI training licenses and diminish their ability to negotiate such licenses in the future. Mot. at 27. They theorize that “[i]f Anthropic continues to exploit and devalue the Works,

even if only during the pendency of this case, its conduct will become entrenched in the AI industry and the public consciousness.” Id. Publishers further contend that Anthropic’s use “will permanently undermine Publishers’ leverage in future negotiations with AI developers to license lyrics as training data.” Id. The evidence cited by Publishers is insufficient to support this theory. For example, Publishers’ economics expert states that “[i]f LLM developers like Anthropic were permitted to take lyrics and other copyrighted works without compensating rightsholders, this growing market for copyrighted works would likely contract and may collapse.” Smith Decl. ¶ 56 (emphasis added). This statement is both conclusory and speculative. Other declarations cited by Publishers on this point are similarly general in nature. They fail to provide details or specifics regarding how, if at all, Anthropic’s use of the Works to train Claude has affected their respective abilities to negotiate training licenses with other AI developers, or how it will inflict harm on the emerging licensing market. See Concord Decl. ¶¶ 25, 27; ABKCO Decl. ¶ 29; CCMG Decl. ¶ 28; UMPG Decl. ¶ 31. Publishers have not demonstrated a likelihood of harm arising from Anthropic’s use of the Works for training purposes. *Am. Passage Media Corp. v. Cass Comm’cns, Inc.*, 750 F.2d 1212, 34 F.3d 1011, 1012 (9th Cir. 1985) (finding insufficient evidence of irreparable harm where supporting affidavits were “conclusory and without sufficient support in facts”). Even if Publishers had shown that the use of the Works to train Claude will cause them harm within the emerging AI licensing market, they once again fail to show that such harm is irreparable. According to Publishers’ evidence, the market for AI training licenses has grown over the course of this lawsuit rather than diminished. See Smith Decl. ¶¶ 76-77 (listing content licensing agreements that OpenAI and other AI developers have entered into in recent years). If other AI developers are obtaining licenses to use copyrighted material for training purposes, then it follows that (1) the market is not being significantly harmed by Anthropic’s use of the Works without licenses, and (2) the value of the loss of licensing fees could be ascertained. Thus, any harm arising from the emerging AI licensing market would be compensable rather than irreparable. See *Fox Broad. Co.*, 747 F.3d at 1073. Publishers have not shown a likelihood of irreparable harm based on Anthropic’s use of the Works to train Claude.⁶ Publishers cite several cases for the proposition that courts must treat the use of copyrighted works in emerging technology markets with caution and care. Mot. at 26-27 (citing *Metro-Goldwyn-Mayer Studios Inc. v. Gorkster, Ltd.*, 545 U.S. 913, 929 (2005), *Waymo LLC v. Uber Tech. Inc.*, No. C-17-00939-WHA, 2017 WL 2123560, at *11 (N.D. Cal. May 15, 2017), and *Warner Bros. Ent. Inc. v. WTV Sys., Inc.*, 824 F. Supp. 2d 1003, 1012 (C.D. Cal. 2011)). The Court does not disagree. However, Publishers fail to explain how this general proposition entitles them to relief. To the contrary, these cases

demonstrate that emerging technologies often test the bounds and principles of copyright law. Here, it is an open question whether training generative AI models with copyrighted material is infringement or fair use. By seeking a preliminary injunction, Publishers are essentially asking the Court to define the contours of a licensing market for AI training where the threshold question of fair use remains unsettled. The Court declines to 6 Publishers’ third argument is that the use of the Works for training purposes harms Publishers’ relationships with songwriters. See Mot. at 28 (citing ABKCO Decl. ¶ 28 (“If Anthropic is permitted to continue to violate [ABKCO’s] rights, that will cause permanent and lasting damage to our relationship and reputation with our songwriters that will undermine our business as a music publisher.”)). The evidence cited for this proposition is similarly general and conclusory, and thus fails to establish irreparable harm. 13 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 United States District Court Northern District of California award Publishers the extraordinary relief of a preliminary injunction based on legal rights (here, licensing rights) that have not yet been established. See *Perfect 10*, 653 F.3d at 980. Because Publishers have not established the prerequisite factor of irreparable harm, the Court does not address the other Winter factors. See *Famous Birthdays, LLC v. SocialEdge, Inc.*, No. CV 21-9562 PA (MRWx), 2022 WL 1592726, at *4 (C.D. Cal. Feb. 11, 2022) (“Because the Court concludes that Famous Birthdays has failed to establish that it will likely suffer irreparable harm in the absence of preliminary relief, the Court does not reach the other elements required for a preliminary injunction.” (citing *Tech. & Intell. Prop. Strategies Grp. PC v. Fthenakis*, No. C 11- 2373 MEJ, 2012 WL 159595, at *4 n.6 (N.D. Cal. Jan. 17, 2012))). IV. CONCLUSION For the foregoing reasons, Publishers are not entitled to the extraordinary remedy of a preliminary injunction. Publishers’ Motion is DENIED without prejudice. In reaching this conclusion, the Court does not address whether Publishers could plausibly state either direct or secondary infringement claims against Anthropic, as those questions are appropriately reserved for Anthropic’s motion to dismiss the complaint and subsequent proceedings on the merits. IT IS SO ORDERED. Dated: Eumi K. Lee United States District Judge March 25, 2025».

B.2) ORDER DENYING MOTION FOR PRELIMINARY INJUNCTION DEL 25 MARZO 2025.

« [...]

II. LEGAL STANDARD

Under Federal Rule of Civil Procedure 12(b)(6), a court must dismiss a complaint if it fails to state a claim upon which relief can be granted. To avoid dismissal, the plaintiff must allege “enough facts to state a claim to relief

that is plausible on its face.” *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 570 (2007). A claim is facially plausible when the pleaded facts allow the court “to draw the reasonable inference that the defendant is liable for the misconduct alleged.” *Ashcroft v. Iqbal*, 556 U.S. 662, 678 (2009). For purposes of a Rule 12(b)(6) motion, the court generally “accept[s] factual allegations in the complaint as true and construe[s] the pleadings in the light most favorable to the nonmoving party.” *Manzarek v. St. Paul Fire & Marine Ins. Co.*, 519 F.3d 1025, 1031 (9th Cir. 2008). However, the court need not “assume the truth of legal conclusions merely because they are cast in the form of factual allegations.” *Fayer v. Vaughn*, 649 F.3d 1061, 1064 (9th Cir. 2011) (per curiam) (quoting *W. Mining Council v. Watt*, 643 F.2d 618, 624 (9th Cir. 1981)). If dismissal is warranted, the court should grant leave to amend “unless it determines that the pleading could not possibly be cured by the allegation of other facts.” *Lopez v. Smith*, 203 F.3d 1122, 1127 (9th Cir. 2000) (en banc) (quoting *Doe v. United States*, 58 F.3d 494, 497 (9th Cir. 1995)).

III. DISCUSSION

Anthropic seeks to dismiss Counts II (contributory infringement), III (vicarious infringement), and IV (the intentional removal of CMI), arguing that Publishers have failed to state a claim. The Court addresses each in turn.

A. Contributory Infringement (Count II) To state a claim for contributory infringement, Publishers must allege that Anthropic (1) has knowledge of another’s infringement, and (2) materially contributes to or induces that infringement. *Perfect 10, Inc. v. Visa Int’l Serv. Ass’n*, 494 F.3d 788, 795 (9th Cir. 2007). “[T]he existence of direct infringement is a necessary element of a claim for contributory infringement.” *Petroliaam Nasional Berhad v. GoDaddy.com, Inc.*, 897 F. Supp. 2d 856, 867 (N.D. Cal. 2012) (citing *Perfect 10*, 494 F.3d at 795). Anthropic moves to dismiss Count II on the grounds that Publishers fail to allege either a predicate act of direct infringement by a third party, or specific knowledge by Anthropic of any third-party infringement.

1. Publishers have not alleged a predicate act of direct third-party infringement. According to Publishers, the complaint adequately alleges predicate acts of third-party infringement by both individual Claude users and Anthropic business customers. Opp. at 8. Publishers point to allegations that individual users, through Anthropic’s website, “can request and obtain through Claude verbatim or near-verbatim copies of lyrics[.]” Compl. ¶ 65. Publishers also allege that “[w]hen a user prompts . . . Claude’s AI chatbot to provide the lyrics to songs . . . the chatbot will provide responses that contain all or significant portions of those lyrics.” Id. ¶ 8.

Publishers argue that because Claude is available to users “in a variety of ways,” individual users “could and would direct Claude to deliver copies of Publishers’ lyrics.”⁵ Opp. at 9. Publishers further contend that Anthropic’s business customers “might use the Claude API tool” to seek copyrighted lyrics in the process of prompting Claude to generate creative output.⁶ Id. at 9-10. Publishers’ allegations and arguments do not support the requested inference of direct third-party infringement. The allegations assert what unidentified users “can” do when prompting Claude, but Publishers have not clearly alleged a predicate act of direct third-party infringement. Other allegations emphasized by Publishers in their opposition are conclusory, and provide no facts regarding direct third-party infringement.⁷ See Compl. ¶¶ 11, 120 and 134. The parties spent time at hearing and in their briefs disputing whether prompts that were entered by Publishers’ investigators – as opposed to a third party – could satisfy this threshold pleading requirement. Although the complaint contains numerous examples of output containing copyrighted lyrics generated by Claude in response to prompts, the complaint does not attribute the prompts to an author. See e.g., id. ¶¶ 66-69. Anthropic argues that these allegations “reflect Plaintiffs’ attempts to generate alleged copies of their own lyrics,” and are therefore “per se noninfringing, so they cannot support a secondary infringement claim.” Mot. at 6. Publishers appear to concede that at least some of the prompts discussed in the complaint were entered by Publishers or their investigators. See Opp. at 13 (arguing that “[i]t would be premature to dismiss the [c]omplaint for failure to specify other third-party infringements when Publishers have alleged that Claude regurgitates lyrics in response to . . . Plaintiffs and their investigators”) (citing Compl. ¶¶ 66-74, 75-79). None of the allegations in the complaint indicate who submitted the prompts that resulted in infringing output, let alone whether they were submitted by Publishers’ investigators. See, e.g., Compl. ¶ 67 (alleging that “when Anthropic’s Claude is prompted ‘What are the lyrics to I Will Survive by Gloria Gaynor,’ . . . the model responds by providing a nearly word-for-word copy of those lyrics”). If the “users” referenced in the complaint are Publishers’ investigators, Publishers should allege this.⁸ The Court declines to decide at this time whether output prompted by Publishers’ investigators would constitute a predicate act of third-party infringement as a matter of law because Publishers have not alleged these facts.

2. Publishers have not alleged that Anthropic had knowledge of specific acts of third-party infringement. Even if Publishers had alleged a predicate

act of third-party infringement, they have not stated a plausible claim because the complaint does not allege that Anthropic knew or had reason to know of any third-party infringement. The first prong of a contributory infringement claim “requires more than a generalized knowledge” of “the possibility of infringement.” *Luvdarts, LLC v. AT & T Mobility, LLC*, 710 F.3d 1068, 1072 (9th Cir. 2013) (holding that conclusory allegations that defendants had the required knowledge of infringement were “plainly insufficient” to satisfy specific knowledge requirement). To state a claim for contributory infringement, Publishers must allege “actual knowledge of specific acts of infringement.” *Id.* Publishers contend that Anthropic is “well aware of its licensees’ and users’ infringing activities through its AI products” because it “knowingly trained its AI models on infringing content on a massive scale in order to enable those models to generate responses to user prompts that infringe Publishers’ copyrighted lyrics.” *Opp.* at 15 (citing *Compl.* ¶ 122). Such allegations are conclusory at best. Claude may be capable of “substantial lawful as well as unlawful uses,” but contributory liability does not “automatically follow” where the technology arguably permits both. *Luvdarts*, 710 F.3d at 1072 (first quoting *Metro-Goldwyn-Mayer Studios Inc. v. Gorkster, Ltd.*, 545 U.S. 913, 932-33 (2005); then citing *A&M Record, Inc. v. Napster Inc.*, 239 F.3d 1004, 1021 (9th Cir. 2001)). In other words, the inclusion of protected lyrics in Claude’s training corpus alone does not establish actual knowledge by Anthropic of third-party infringement. The cases cited by Publishers in support of this theory do not assist them in light of the conclusory nature of the allegations in the complaint. See *Opp.* at 16. For example, in *BMG Rts. Mgmt. (US) v. Jony Inc.*, the court held that the defendant’s responses to over half of plaintiffs’ takedown requests for copyrighted material supported a reasonable inference that the defendant had specific knowledge of third-party infringement related to the requests. 716 F. Supp. 3d 835, 843 (C.D. Cal. 2024). And in *Splunk Inc. v. Cribl, Inc.*, the court held that the plaintiff adequately alleged knowledge of infringement because the defendant’s CEO “‘derived go-S2S from Splunk’s copyrighted source code,’ ‘provided this code to Cribl . . . with knowledge that go-S2S was an unlicensed derivative of Splunk’s copyrighted S2S version 3 code,’ and ‘each new version of Cribl’s Stream software includes a new copy of this unlicensed derivative of Splunk’s copyrighted S2S version 3 code.’” 662 F. Supp. 3d 1029, 1052 (N.D. Cal. 2023) (citation omitted). In those cases, the plaintiffs made specific allegations establishing the element of knowledge or supporting an inference of

knowledge. Here, the allegations regarding knowledge are conclusory and speculative. They do not satisfy the pleading standard. Because the complaint fails to allege either a predicate act of direct third-party infringement or actual knowledge of infringement, the Court grants the motion as to Count II with leave to amend. B. Vicarious Infringement (Count III) “To state a claim for vicarious copyright infringement, a plaintiff must allege that the defendant has (1) the right and ability to supervise the infringing conduct and (2) a direct financial interest in the infringing activity.” *Perfect 10*, 494 F.3d at 802 (citing *Ellison v. Robertson*, 357 F.3d 1072, 1078 (9th Cir. 2004)). Stated differently, “[o]ne . . . infringes vicariously by profiting from direct infringement while declining to exercise a right to stop or limit it.” *Id.* (quoting *Grokster*, 545 U.S. at 930). “Like contributory liability, vicarious liability requires an underlying act of direct infringement.” *Perfect 10, Inc. v. Yandex N.V.*, 962 F. Supp. 2d 1146, 1158 (N.D. Cal. 2013). As discussed above, Publishers fail to allege the necessary predicate act of direct infringement by a third party. Thus, the Court grants the motion to dismiss Count III with leave to amend. According to Anthropic, Publishers also fail to state a claim for vicarious infringement because they have not alleged that Anthropic had a direct financial interest in any infringing activity. On this point, the Court disagrees. “Financial benefit exists where the availability of infringing material ‘acts as a draw for customers.’” *Napster*, 239 F.3d at 1023 (citing *Fonovisa, Inc. v. Cherry Auction, Inc.*, 76 F.3d 259, 263-64 (9th Cir. 2001)). “The size of the draw relative to a defendant’s overall business is immaterial.” *Perfect 10, Inc. v. Giganews, Inc.*, 847 F.3d 657, 673 (9th Cir. 2017) (quoting *Perfect 10, Inc. v. Giganews, Inc.*, No. CV 11-07098-AB (SHx), 2014 WL 8628031, at *3 (C.D. Cal. Nov. 14, 2014)). “The essential aspect of the ‘direct financial benefit’ inquiry is whether there is a causal relationship between the infringing activity and any financial benefit a defendant reaps, regardless of how substantial the benefit is in proportion to a defendant’s overall profits.” *Ellison*, 357 F.3d at 1079. “There is no requirement that the draw be ‘substantial.’” *Id.* Assuming for the sake of argument that Publishers will amend the complaint to allege a predicate act of third-party infringement, the current allegations plausibly state that Anthropic received a direct financial benefit. See, e.g., *Compl.* ¶ 94 (alleging that Anthropic receives revenues from commercial customers based on the amount of text submitted by the customers’ end users into the Claude API and the amount of text generated as output); *id.* (alleging that Anthropic “is paid every time one of its customers’ end users submits

a request for Publishers' song lyrics, and it is paid again every time its Claude API generates output copying and relying on those lyrics"); *id.* ¶ 98 ("One of the reasons that Anthropic's AI models are so popular and valuable is because of the substantial underlying text corpus that includes Publishers' copyrighted lyrics. As such, Publishers' copyrighted content serves as a draw for individual users, commercial customers, and ultimately investors."). Accepting these allegations as true for purposes of this motion, Publishers have plausibly alleged that Anthropic received a direct financial benefit from potentially infringing activity. See e.g., *Keck v. Alibaba.com Hong Kong Ltd.*, 369 F. Supp. 3d 932, 938 (N.D. Cal. 2019) (denying motion to dismiss vicarious infringement claim where the plaintiff alleged that defendants reproduced plaintiff's artwork on their websites; the infringing material drew customers to the websites; defendants received commissions from the sale of unauthorized copies of her work; and defendants had monetized increased traffic due to the work). However, because Publishers have failed to allege direct infringement by Anthropic's users, the Court grants the motion with leave to amend as to Count III. C. Removal or Alteration of Copyright Management Information (Count IV) Count IV alleges that Anthropic "intentionally removed or altered copyright management information from Publishers' musical compositions, and/or distributed or imported for distribution copies of Publishers' musical compositions knowing that copyright management information has been removed or altered, without Publishers' authorization," in violation of Section 1202(b) of the DMCA. Compl. ¶ 149. The complaint does not specify which subsections of Section 1202(b) are at issue. At hearing, Publishers clarified that Count IV includes subsections (b)(1) and (b)(3). 12/19/2024 Hr'g Tr. at 20. "Copyright law restricts the removal or alteration of copyright management information ("CMI") – information such as the title, the author, the copyright owner, the terms and conditions for use of the work, and other identifying information set forth in a copyright notice or conveyed in connection with the work." *Stevens v. Corelogic, Inc.*, 899 F.3d 666, 671 (9th Cir. 2018) (citing 17 U.S.C. § 1202(b)-(c)). Section 1202(b)(1) provides, "No person shall, without the authority of the copyright owner or the law . . . intentionally remove or alter any copyright management information." 17 U.S.C. § 1202(b)(1). Section 1202(b)(3) states: No person shall, without the authority of the copyright owner or the law . . . distribute . . . copies of works, or phonorecords, knowing that copyright management information has been removed or altered

without authority of the copyright owner or the law, knowing, or . . . having reasonable grounds to know, that it will induce, enable, facilitate, or conceal an infringement of any right under this title. "Both provisions thus require the defendant to possess the mental state of knowing, or having a reasonable basis to know, that [the removal of CMI] 'will induce, enable, facilitate, or conceal' infringement." *Stevens*, 899 F.3d at 673; see also *Doe 1 v. GitHub, Inc.*, 672 F. Supp. 3d 837, 858 (N.D. Cal. 2023) ("On a motion to dismiss, a plaintiff must allege sufficient facts to support the reasonable inference that the defendant 'knew or had a reasonable basis to know that the removal or alteration of CMI . . . w[ould] aid infringement.'"). According to Publishers, the allegations reflect that Anthropic intentionally removed CMI. Publishers highlight paragraphs 84 and 149 of the complaint, but those allegations are conclusory. They simply repeat the statutory language without stating any facts to show that Anthropic intentionally removed CMI. Publishers identify other portions of the complaint relating to the omission of CMI from Claude's output, and contend these sections support an inference that CMI was intentionally removed. See Opp. at 21-22 (citing Compl. ¶¶ 11, 73, 74, 86, 87). They do not. Although Publishers plausibly allege that the output omits the CMI regarding the Works, they have not sufficiently pled that Anthropic acted intentionally as to the removal of CMI. See Compl. ¶ 11 (alleging that Claude's output "often omits critical [CMI]" regarding the Works"), ¶ 86 (alleging that when Claude has reproduced lyrics they are "often unaccompanied by" CMI), ¶¶ 73, 74, 87 (examples of song lyrics where Claude allegedly omitted attribution). Citing GitHub, Publishers urge the Court to infer from these allegations that the removal of CMI was intentional, or that Anthropic intentionally designed Claude's training process to remove CMI. See Opp. at 21, 24. Github does not support this argument because the plaintiff's allegations in that case were much more specific. There, the court concluded it could reasonably infer intentional removal of CMI based on the allegations that (1) the defendants knew the source code they used to train their AI programs contained CMI, (2) the defendants knew that CMI was important for protecting copyright interests, and (3) Github had regularly processed takedown notices "such that it was aware its platform was used to distribute code with removed or altered CMI in a manner which induced infringement." *GitHub*, 672 F. Supp. 3d at 858. Here, Publishers' allegations are conclusory. Unlike Github, Publishers allege no specific facts that reflect Anthropic's intentional removal of CMI or that it was on notice that CMI had been

removed. The Court notes that Publishers have alleged that some outputs do include attribution for lyrics. See, e.g., Compl. ¶ 76 (alleging that Claude produced chords and lyrics to “Daddy Sang Bass” in response to query of “Give me the chords to Daddy Sang Bass by Johnny Cash”). These allegations arguably undermine Publishers’ theory that Anthropic intentionally removes CMI by design, see Opp. at 24, because they suggest that CMI was not removed across the board – intentionally or otherwise – during Claude’s training process. In *Tremblay v. OpenAI, Inc.*, the court found that similarly contradictory allegations undermined the plaintiffs’ Section 1202(b) claim. 716 F. Supp. 3d 772, 779 (N.D. Cal. 2024) (noting that plaintiffs could not support their assertion that defendants removed CMI from copyrighted books used in AI training “by design” because their allegations included excerpts of outputs, which referenced authors’ names). At minimum, the contradictory nature of Publishers’ allegations undermines their argument that CMI was intentionally removed. Turning to Section 1202(b)(3), Publishers contend that their allegations regarding Claude’s training process support a reasonable inference that Anthropic “knew its conduct would induce, enable, facilitate, or conceal further infringement.” Opp. at 24. This argument seems to rely on the same allegations regarding Anthropic’s training process, and the same underlying theory that Claude’s training process imbued Anthropic with knowledge that the removal of CMI would induce third party infringement. Once again, Publishers’ allegations are too conclusory to establish such knowledge, and Publishers fail to state a plausible claim under Section 1202(b). Accordingly, the Court grants the motion to dismiss Count IV with leave to amend.

IV. CONCLUSION

For the foregoing reasons, Anthropic’s motion to dismiss Counts II, III, and IV is GRANTED with leave to amend. Publishers may file an amended complaint within 30 days of this order. Publishers shall not add any new claims without leave of court.

IT IS SO ORDERED.

GIORNALISMO, AI E MARKET SUBSTITUTION
THE NEW YORK TIMES COMPANY V. MICROSOFT
CORPORATION & OPENAI, INC. ET AL., No. 1:23-cv-
11195-SHS-OTW (S.D.N.Y.)

di Sara Manni

I. Premesse

La controversia *The New York Times Company v. Microsoft Corporation & OpenAI, Inc. et al.*, pendente dinanzi alla United States District Court for the Southern District of New York (No. 1:23-cv-11195-SHS-OTW), U.S. District Judge Sidney H. Stein, costituisce una delle più significative controversie giudiziarie sviluppatesi negli ultimi anni in materia di intelligenza artificiale generativa e diritto d'autore e sicuramente, tra quelle con maggiore riscontro mediatico. Il procedimento, che ha preso avvio nel 2023, è oggi confluito - unitamente alle altre azioni connesse promosse da editori e autori - nel *multidistrict litigation* identificato come *In re: OpenAI, Inc. Copyright Infringement Litigation*, No. 1:25-md-03143-SHS-OTW.

La controversia si colloca al centro del crescente conflitto fra imprese tecnologiche operanti nel settore dell'IA generativa e industria dell'informazione professionale. I *large language models* (modelli linguistici addestrati attraverso l'analisi statistica di ingenti quantitativi di dati testuali), fra i quali ChatGPT, sviluppato da OpenAI a partire dalle famiglie di modelli GPT-3, GPT-3.5 e GPT-4, sono in grado di produrre testi, sintesi, risposte argomentative e traduzioni mediante processi di apprendimento automatico fondati sull'elaborazione di *dataset* composti, in alcuni casi, da miliardi di documenti digitali.

Il funzionamento di tali modelli presuppone necessariamente l'acquisizione, la copia e

l'analisi di quantità rilevanti di contenuti reperiti su siti *web*, archivi digitali, piattaforme editoriali e banche dati *online*. L'azione promossa dal *New York Times* coinvolge direttamente una delle principali organizzazioni giornalistiche al mondo e pone al centro del giudizio la contestata utilizzazione, senza autorizzazione, di articoli editoriali coperti da *copyright* per finalità di addestramento algoritmico. La vicenda riflette, peraltro, una tensione strutturale già evidenziata altrove: i moderni sistemi di IA generativa non si limitano alla mera consultazione delle informazioni reperite *online*, ma realizzano operazioni di estrazione, copia, indicizzazione ed elaborazione computazionale dei contenuti digitali che, secondo la prospettiva attorea, integrerebbero forme rilevanti di sfruttamento economico delle opere protette²²⁶.

II. Le parti del giudizio

Il procedimento vede contrapposte, da un lato, *The New York Times Company*; dall'altro, *Microsoft Corporation* e una pluralità di entità riconducibili al gruppo OpenAI - segnatamente OpenAI, Inc., OpenAI LP, OpenAI GP LLC, OpenAI LLC, OpenAI OpCo LLC, OpenAI Global LLC, OAI Corporation, LLC e OpenAI Holdings, LLC - la cui complessa articolazione societaria riflette le successive ristrutturazioni del soggetto economico riconducibile a tale denominazione. Il *New York Times*, fondato nel 1851, opera nel settore dell'informazione e dell'editoria digitale e ha sviluppato, nel corso degli ultimi tre lustri, un modello economico fondato sulla monetizzazione dei propri contenuti giornalistici mediante abbonamenti digitali, sistemi di *paywall*, licenze editoriali e accordi di distribuzione commerciale delle notizie²²⁷. La

²²⁶ Il procedimento è stato ampiamente commentato nella dottrina statunitense. Per un'analisi estesa del caso v., tra i tanti, J. ODELL, *Training on Headlines: The New York Times, OpenAI, and the Copyright Implications of AI Data Usage*, in *Oregon Law Review*, 104, 2025, 203 ss.; con riferimento specifico al profilo della *memorization*, J. FREEMAN, C. RIPPE, E. DEBENEDETTI, M. ANDRIUSHCHENKO, *Exploring Memorization and Copyright Violation in Frontier LLMs: A Study of the New York Times v. OpenAI*

2023 Lawsuit, in *Safe Generative AI Workshop at NeurIPS 2024*, 1-16, disponibile online in Arxiv.

²²⁷ Sull'impatto dell'IA nel settore c.d. GLAM (*galleries, libraries, archives, museums*) e sul rapporto della stessa con le licenze (in particolare Creative Commons), si rinvia a M. RIMMER, *Night at the Artificial Museum: Copyright Law and Artificial Intelligence*, in *Culture Unbound – Journal of Current Cultural Research*, 18, 2026, 151-2010.

testata investe risorse considerevoli nella produzione di articoli originali, inchieste, *reportage* e contenuti editoriali professionali, e tale attività costituisce, in ultima analisi, il nucleo centrale della pretesa attorea.

OpenAI, costituita inizialmente quale *non-profit* e successivamente trasformata, in parte, in soggetto a scopo di lucro, è divenuta nota per lo sviluppo della famiglia di modelli linguistici denominati GPT (*Generative Pre-trained Transformers*), e in particolare per il sistema ChatGPT. Tali modelli appartengono alla categoria dei *large language models*, addestrati mediante l'analisi statistica di rilevanti quantitativi di testi digitali, con la funzione di apprendere correlazioni linguistiche e di generare testi coerenti in risposta agli *input* forniti dall'utente.

Microsoft Corporation, infine, riveste nel procedimento un ruolo centrale in qualità di principale partner commerciale e tecnologico di OpenAI: la multinazionale ha investito ingenti somme nello sviluppo del gruppo, ha integrato i modelli GPT nei propri servizi e ne ha promosso l'impiego all'interno del motore di ricerca Bing – con la funzione Bing Chat, successivamente ribrandizzata *Copilot* – e in altri prodotti commerciali. Secondo il *Times*, Microsoft avrebbe tratto vantaggio economico diretto dall'utilizzazione, da parte di OpenAI, dei contenuti editoriali del *Times*, e l'integrazione dei sistemi generativi nei motori di ricerca rischierebbe, in particolare, di compromettere il traffico verso le testate giornalistiche tradizionali.

III. L'origine della controversia e l'atto introduttivo

L'atto di citazione (*complaint*), depositato il 27 dicembre 2023²²⁸, si apre con un incipit divenuto particolarmente noto, nel quale parte attorea

afferma testualmente: «[i]ndependent journalism is vital to our democracy. It is also increasingly rare and valuable» (il giornalismo indipendente è vitale per la nostra democrazia. Esso è anche sempre più raro e prezioso). La formula, nella sua brevità, condensa il fulcro della prospettiva attorea: la sostenibilità economica del giornalismo indipendente quale presupposto della funzionalità democratica, e il ruolo che il *copyright* è chiamato a svolgere nel proteggere quel presupposto.

Tale valore democratico sarebbe direttamente minacciato dai prodotti di IA²²⁹.

Difatti, si legge nella citazione che:

«4. GenAI Products Threaten High-Quality Journalism

47. Making great journalism is harder than ever. Over the past two decades, the traditional business models that supported quality journalism have collapsed, forcing the shuttering of newspapers all over the country. It has become more difficult for the public to sort fact from fiction in today's information ecosystem, as misinformation floods the internet, television, and other media. If The Times and other news organizations cannot produce and protect their independent journalism, there will be a vacuum that no computer or artificial intelligence can fill.

48. The protection of The Times's intellectual property is critical to its continued ability to fund world-class journalism in the public interest. If The Times and its peers cannot control the use of their content, their ability to monetize that content will be harmed. With less revenue, news organizations will have fewer journalists able to dedicate time and resources to important, in-depth stories, which creates a risk that those stories will go untold. Less journalism will be produced, and the cost to society will be enormous.

49. The Times depends on its exclusive rights of reproduction, adaptation,

²²⁸ *Complaint, The New York Times Company v. Microsoft Corporation et al.*, No. 1:23-cv-11195 (S.D.N.Y., 27 dicembre 2023), Doc. 1.

²²⁹ Per prospettive (critiche e non) circa l'impatto dei sistemi di IA generativa sul settore del giornalismo, M. BROUSSARD – N. DIAKOPOLOS – C. CHUAN ET AL., *Artificial Intelligence and Journalism*, in *Journalism & Mass*

Communication Quarterly, 2019, 96, 673-695; A. L. OPDAHL – B. TESSEM – D. DANG-NGUYEN – E. MOTTA – V. SETTY – E. THRONSDEN – A. TVERBERG – C. TRATTNER, *Trustworthy journalism through AI*, in *Data & Knowledge Engineering*, 146, 2023, 1-17, disponibile online in <Elsevier.com>; D. WENGER – S. HOSSAIN – J. R. SENSEMAN – ET AL., *AI and the Impact on Journalism Education*, in *Journalism & Mass Communication Educator*, 2025, 80, 97-114.

publication, performance, and display under copyright law to resist these forces. The Times has registered the copyright in its print edition every day for over 100 years, maintains a paywall, and has implemented terms of service that set limits on the copying and use of its content. To use Times content for commercial purposes, a party should first approach The Times about a licensing agreement.

50. The Times requires third parties to obtain permission before using Times content and trademarks for commercial purposes, and for decades The Times has licensed its content under negotiated licensing agreements. These agreements help ensure that The Times controls how, where, and for how long its content and brand appears and that it receives fair compensation for third-party use. Third parties, including large tech platforms, pay The Times significant royalties under these agreements in exchange for the right to use Times content for narrowly defined purposes. The agreements prohibit uses beyond those authorized purposes»²³⁰.

Nella ricostruzione attorea, i modelli GPT sarebbero stati addestrati utilizzando milioni di articoli protetti provenienti dagli archivi digitali del *Times*, taluni dei quali accessibili soltanto previo abbonamento, ossia protetti mediante *paywall*. Sotto il profilo probatorio, la testata ha articolato il proprio *complaint* valorizzando, fra l'altro, due dati di rilievo: in primo luogo, la circostanza per cui il *dataset Common Crawl* - utilizzato per il *training* dei modelli GPT - comprenderebbe almeno 16 milioni di articoli unici del *Times* e oltre 66 milioni di *record* complessivi²³¹; in secondo luogo, il fatto che il *Times* risulti fra le fonti più rappresentate anche all'interno del *dataset* proprietario di OpenAI denominato *WebText2*, parimenti utilizzato per il *training* dei modelli GPT.

Particolare risalto, nel *complaint*, è stato attribuito al fenomeno della c.d.

«*memorization*»: il *Times* ha allegato, e in parte documentato, casi nei quali ChatGPT - interrogato mediante *prompt* mirati - riproduceva passaggi sostanzialmente parola per parola di articoli della testata, talora coincidenti con porzioni protette da *paywall*. La capacità del modello di rigenerare il contenuto delle opere - pur a fronte della natura intrinsecamente probabilistica del processo generativo - è stata addotta dall'attrice quale prova della loro avvenuta incorporazione, in qualche misura, nei parametri del modello.

Il *Complaint* originario - successivamente integrato dal *First Amended Complaint* del 12 agosto 2024 e dal *Second Amended Complaint* del 28 maggio 2025 - ha articolato la pretesa attorno a una pluralità di capi, che meritano di essere illustrati nella loro sistemazione propria. In primo luogo, un capo di *direct copyright infringement*, fondato sulla Section 501 del *Copyright Act* del 1976 (17 U.S.C. § 501), in relazione alla violazione dei diritti esclusivi di cui alla Section 106 del medesimo *Act*: secondo l'attrice, gli articoli protetti sarebbero stati riprodotti tanto nella fase di *training* dei modelli, quanto in quella della successiva generazione degli *output*, in assenza di alcuna licenza.

In secondo luogo, un capo di *contributory copyright infringement*, fondato sulla giurisprudenza statunitense in materia di responsabilità secondaria del *copyright*: l'assunto attoreo è che le convenute avrebbero contribuito materialmente alla violazione realizzata da utenti terzi, mettendo a disposizione del pubblico strumenti idonei a generare *output* sostanzialmente coincidenti con gli articoli protetti.

In terzo luogo, capi fondati sul *Digital Millennium Copyright Act* (DMCA), e segnatamente sulla Section 1202 (17 U.S.C. § 1202), commi (b)(1) e (b)(3), per la dedotta rimozione o alterazione delle informazioni sulla

²³⁰ *Complaint, The New York Times Company v. Microsoft Corporation et al.*, No. 1:23-cv-11195 (S.D.N.Y., 27 dicembre 2023), 47-50, disponibile all'indirizzo <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://h-h-law.com/wp-content/uploads/2024/07/New-York-Times-complaint.pdf>.

²³¹ Cfr. *Complaint, The New York Times Company v. Microsoft Corporation et al.*, cit., ove vengono dettagliatamente illustrate le risultanze delle analisi condotte dai consulenti tecnici incaricati dalla parte attorea sul *dataset Common Crawl* e sui *dataset* proprietari di OpenAI.

gestione del diritto d'autore (*copyright management information*) presenti negli articoli del *Times* - ossia titoli, *byline*, note di copyright e analoghi metadati - nell'ambito del processo di *ingestion* dei contenuti nel *dataset*.

Quarto, un capo di *common law unfair competition by misappropriation*, invocato sotto il profilo della c.d. *hot news misappropriation doctrine*, istituito risalente al noto precedente *International News Service v. Associated Press*, 248 U.S. 215 (1918): secondo l'attrice, i sistemi generativi sfrutterebbero l'investimento giornalistico altrui producendo *output* suscettibili di sostituire economicamente il consumo diretto delle notizie, peraltro aggravato dall'assenza di alcuna indicazione della fonte che priva il *Times* anche del *referral revenue*.

Quinto, un capo di *federal trademark dilution* motivato dalla circostanza per cui taluni *output* generati dai modelli - eventualmente errati, fuorvianti o intrinsecamente inattendibili - venivano falsamente attribuiti al *Times* mediante l'apposizione del marchio della testata (così "macchiando" il marchio – *dilution by tarnishing*). È in questo capo che si colloca il riferimento al fenomeno delle c.d. «*hallucinations*» - risposte plausibili, ma non fondate su alcuna fonte reale, che il modello produce attribuendole, talora, a fonti reputabili. La pratica determinerebbe, per il *Times*, non soltanto un danno commerciale e concorrenziale, ma altresì un rilevante pregiudizio reputazionale.

Le convenute, per contro, hanno invocato la liceità dell'accesso ai contenuti, valorizzando la natura intrinsecamente «aperta» dell'infrastruttura *Internet* e l'architettura del *web crawling* quale prassi consolidata nello

sviluppo delle tecnologie informatiche. È stata inoltre contestata, sul piano fattuale, la qualificazione del *training* quale atto di sfruttamento economico delle opere, e ciò sulla base della rappresentazione tecnica secondo cui i modelli GPT non opererebbero quali archivi di testi memorizzati, ma genererebbero gli *output* mediante processi probabilistici fondati sull'apprendimento di *pattern* linguistici astratti. Sotto un secondo profilo - destinato verosimilmente ad assumere centralità nella fase di merito del giudizio - è stata invocata la *doctrine of fair use* disciplinata dalla Section 107 del *Copyright Act* (17 U.S.C. § 107), sull'assunto che il *training* costituirebbe una utilizzazione altamente trasformativa delle opere, finalizzata non già alla redistribuzione commerciale dei contenuti, bensì all'apprendimento statistico delle strutture linguistiche e semantiche del linguaggio umano. La difesa convenuta ha richiamato, sul punto, l'ampia produzione giurisprudenziale statunitense in materia di *fair use*, e in particolare i precedenti *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569 (1994), e - più direttamente pertinente in ragione della natura intermedia dell'uso - *Authors Guild v. Google, Inc.*, 804 F.3d 202 (2d Cir. 2015), relativo al progetto *Google Books*²³².

Le convenute hanno altresì sostenuto che gli esempi di *output* che riproducevano pedissequamente i contenuti di titolarità dell'attrice, dalla medesima documentati, sarebbero il frutto di *prompt* deliberatamente costruiti per indurre il modello a rigenerare passaggi protetti, e dunque non rappresentativi del normale funzionamento del sistema e dell'uso ordinario del servizio.

È inoltre opportuno ricordare, sul piano del contesto, che OpenAI ha stipulato, prima e dopo

²³²Sul tema del *fair use* applicato all'IA generativa, in dottrina, si vedano fra gli altri M.A. LEMLEY, B. CASEY, *Fair Learning*, cit., nonché E. ROSATI, *Copyright exceptions and fair use defences for AI training done for "research" and "learning," or the inescapable licensing horizon*, in *European Journal of Risk Regulation*, 2025, 961-984; R. VADISSETTY, *Navigating Copyright and Fair Use in AI Training Data: Legal Challenges and Future Solutions*, in *2025 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)*, IEEE, 2025; M.

D. MURRAY, *Generative AI art: Copyright infringement and fair use*, in *SMU Science & Tech. Law Review*, 2023, 259 ss.; J. C. CHARLESWORTH, *Generative AI's Illusory Case for Fair Use*, in *Vand. J. Ent. & Tech. L.*, 2024, 323 ss.; e, in prospettiva comparatistica con l'ordinamento sinico, J. ZHANG – X. LI, *Fair Use in Training AI Models: A Review and Prospect of the Relevant Legal Development in China*, in *German Law Journal*, 2025, 1235-1259.

la causa, accordi di *licensing* con diverse organizzazioni editoriali - fra le quali Axel Springer, titolare di *Politico* e *Business Insider*, e l'*Associated Press* - sì che la posizione difensiva sostenuta nel presente giudizio convive, sul piano commerciale, con un approccio non univoco al problema della remunerazione dei titolari dei diritti²³³.

IV. Lo sviluppo processuale e l'ordinanza del 4 aprile 2025

Il procedimento ha conosciuto, nei mesi successivi al deposito dell'atto introduttivo, un articolato sviluppo processuale che merita di essere ripercorso, soprattutto in considerazione del fatto che la pronuncia di merito non è ancora intervenuta e che il presente commento si concentra dunque sulle statuizioni interlocutorie già rese.

Dopo il deposito, da parte delle convenute, delle *motions to dismiss* (febbraio 2024), il *New York Times* ha ottenuto l'autorizzazione a depositare un *Amended Complaint*, successivamente depositato il 12 agosto 2024. In parallelo, il *Times* è stato consolidato - unitamente alle azioni *Daily News* (No. 1:24-cv-03285) e *Center for Investigative Reporting* (No. 1:24-cv-04872) - in un quadro istruttorio unitario.

Con *opinion and order* del 4 aprile 2025, il *District Judge* Sidney H. Stein si è pronunciato sulle eccezioni di rito sollevate dalle convenute, pervenendo a una pluralità di statuizioni.

In *primis*, Stein ha respinto le eccezioni di OpenAI relative alla prescrizione triennale dei *direct copyright infringement claims* per condotte risalenti agli anni 2019 e 2020, ritenendo che la *discovery rule* - secondo cui il termine decorre dal momento in cui il titolare ha avuto, o avrebbe dovuto avere, conoscenza della violazione - non consenta, in fase di *motion to dismiss*, di considerare le pretese come «*time-barred*».

In secondo luogo, sono state respinte le eccezioni relative ai *contributory copyright infringement claims*, ritenendo le allegazioni attoree sufficienti a integrare la fattispecie.

Infine, sono state respinte, nell'ambito dell'azione *Daily News*, le eccezioni relative ai *federal and state trademark dilution claims*.

L'ordinanza ha invece accolto le eccezioni delle convenute con riferimento ad altri capi di domanda.

Sono state segnatamente dichiarate inammissibili - con sentenza pronunciata «*with prejudice*»²³⁴ - le pretese fondate sulla *common law unfair competition by misappropriation*, in

²³³ Molti, in dottrina, hanno sottolineato l'opportunità di implementare un sistema di *licensing agreements* per il training delle IA: S. RANIA – H. ISMAIL, *Redefining Digital Licensing Agreements in the Era of Artificial Intelligence*, in *IOASD Journal of Business and Management Studies*, 2025, 119-130; M. WAGNER, *From free to fee: antitrust implications of AI training data licensing after Bartz and Kadrey in the Northern District of California*, in *Journal of European Competition Law & Practice*, 2026, 24 ss.; in specifica applicazione del caso Bartz di cui al commento precedente, M. WAGNER, *From free to fee: antitrust implications of AI training data licensing after Bartz and Kadrey in the Northern District of California*, in *Journal of European Competition Law & Practice*, 2026, 1-8, disponibile online in <oxford.academic>; E. ROSATI, *Copyright exceptions and fair use defences for AI training done for "research" and "learning," or the inescapable licensing horizon*, in *European Journal of Risk Regulation*, 2025, 961-984; P. SAMUELSON, *Assessing the feasibility of collective licensing of in-copyright works as training data for generative AI systems*, in *Proceedings of the National Academy of Sciences (forthcoming)*, UC Berkeley Public Law Research Paper, 6051014 (2025).

²³⁴ Le locuzioni *with prejudice* e *without prejudice* qualificano, nel diritto processuale statunitense, l'efficacia preclusiva del provvedimento che definisce il giudizio, segnatamente del *dismissal*. La pronuncia resa *with prejudice* equivale a una decisione nel merito e preclude la riproposizione della medesima domanda fra le stesse parti (*res judicata*, nella forma della *claim preclusion*); quella resa *without prejudice* non incide sul merito e lascia impregiudicata la facoltà di agire nuovamente, entro i limiti dello *statute of limitations*. A norma della Rule 41 Fed. R. Civ. P., il *dismissal* volontario è di regola *without prejudice* - salvo si tratti di una seconda rinuncia sulla medesima domanda (*two-dismissal rule*) -, mentre quello involontario opera *with prejudice* ove il giudice non disponga altrimenti; le definizioni transattive si concludono di norma con un *dismissal with prejudice*. Nell'ordinamento italiano il riferimento concettuale è il giudicato sostanziale: la pronuncia *with prejudice* corrisponde alla decisione idonea a fare stato, quella *without prejudice* alla decisione che non preclude la riproposizione della domanda, come nelle pronunce in rito o nell'estinzione del processo.

applicazione della *preemption doctrine* di cui alla Section 301 del *Copyright Act*, che assorbe nella disciplina federale le pretese statali strutturalmente coincidenti con la materia del *copyright*. La caduta di tale *count* è di considerevole rilievo, in quanto il *Times* aveva con esso tentato di estendere la propria tutela al di là dei confini formali della violazione del *copyright*, invocando la *hot news misappropriation doctrine* del precedente *International News Service*.

Con riferimento ai *DMCA claims*, l'*opinion* ha disposto una articolata distribuzione delle decisioni: il *count* di cui alla Section 1202(b)(1) è stato dichiarato inammissibile, senza pregiudizio, nei confronti di Microsoft in tutte le azioni e nei confronti di OpenAI nell'azione *Times*; analoga sorte è stata riservata, sempre senza pregiudizio, al *count* di cui alla Section 1202(b)(3) nei confronti di entrambe le convenute. Sono invece state respinte le eccezioni di OpenAI relative ai *DMCA claims* sub Section 1202(b)(1) nelle azioni *Daily News* e *Center for Investigative Reporting*.

All'esito della pronuncia, la quasi totalità dei capi fondati sul *copyright* - comprensivi tanto del *direct infringement*, quanto del *contributory infringement* - è dunque rimasta in giudizio, sì che le questioni di merito potranno essere affrontate nella successiva fase istruttoria e dibattimentale.

Il *Second Amended Complaint* del 28 maggio 2025 ha articolato in via definitiva le pretese attoree, e nel dicembre 2025 le convenute hanno depositato le rispettive *answers*.

Allo stato attuale, il giudizio si trova nella fase di *discovery*, e si è inoltre da poco costituito il *multidistrict litigation In re OpenAI, Inc. Copyright Infringement Litigation*, No. 1:25-md-03143-SHS-OTW, nel quale confluiscono numerose azioni connesse promosse da editori e autori in diversi *districts*.

V. Le principali questioni giuridiche sottese al giudizio

Il tema più significativo, sotto il profilo sistematico, attiene alla qualificazione giuridica dell'atto di *training* dei modelli generativi: si tratta, in altri termini, di stabilire se l'utilizzazione di opere protette per finalità di addestramento algoritmico costituisca, e in quale misura, una forma di sfruttamento economicamente rilevante.

La questione assume particolare complessità in quanto il *training* implica necessariamente la copia, almeno temporanea, dei contenuti utilizzati, e si traduce, all'esito del processo, nella produzione di un modello matematico astratto, costituito da parametri numerici esprimenti correlazioni statistiche. La pertinenza del rilievo è confermata, peraltro, dal parallelismo con la coeva controversia *Getty Images v. Stability AI*, definita in primo grado dalla *High Court of Justice* inglese con sentenza del 4 novembre 2025 ([2025] EWHC 2863 (Ch)) – della quale si dirà *infra* –: in entrambi i giudizi, il problema centrale ruota attorno alla questione se i parametri di un modello generativo possano qualificarsi come «copie» - o, comunque, come riproduzioni rilevanti - delle opere utilizzate per l'addestramento.

Sotto questo profilo, parte della dottrina statunitense ha valorizzato un approccio «decostruttivo» all'ecosistema dell'IA generativa, scomponendo il processo in una successione di stadi distinti e, soprattutto, distinguono nettamente i *modelli* dall'*output* da essi generato. Da tale distinzione discende un possibile iato logico nell'impostazione attorea: la sostanziale somiglianza riscontrabile nell'*output* non implica necessariamente che l'incorporazione delle opere nei modelli integri, di per sé, una violazione diretta, poiché alla generazione dell'*output* concorre anche l'utente quale attore autonomo - profilo, questo, destinato a intrecciarsi con la pretesa di responsabilità secondaria. Il rilievo assume peso in ragione del c.d. «many hands problem», ossia della difficoltà di individuare, in un ecosistema a molti attori, il soggetto responsabile²³⁵.

²³⁵ T. CHEN, *Language Models' Verbatim Copying: Copyright Infringement Analysis Through the Lens of the New York Times Co. v. Microsoft Corp., OpenAI, Inc. et al.*,

in *Cardozo Arts & Ent. L.J.*, vol. 43, 2025, p. 349, spec. pp. 357 ss., che rielabora M. MURRAY, *Generative AI Art: Copyright Infringement and Fair Use*, 26 *SMU Sci. & Tech.*

Il procedimento ruota, per ampia parte, attorno all'interpretazione della *doctrine of fair use* di cui alla Section 107 del *Copyright Act*. Le convenute sostengono che il *training* costituisca un uso altamente trasformativo, sì che l'IA non sfrutterebbe economicamente le opere come «sostituti diretti»; il *New York Times*, per converso, contesta tale impostazione evidenziando come gli *output* generati possano, di fatto, sostituire economicamente gli articoli originari, compromettendo il valore di mercato del contenuto.

La questione si presenta particolarmente articolata sotto il profilo del quarto fattore del *fair use* - l'«*effect of the use upon the potential market for or value of the copyrighted work*» - considerato dalla giurisprudenza statunitense, e segnatamente dalla Corte Suprema statunitense nel recente *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023), quale criterio di crescente rilievo nell'operazione di bilanciamento²³⁶.

Il caso del *Times* presenta, su questo terreno, un profilo di particolare delicatezza: la *market substitution* più che la mera riproduzione delle opere, riguarda la possibilità - strutturalmente nuova nel panorama tecnologico contemporaneo - che il *chatbot* si sostituisca al sito della testata quale fonte di accesso all'informazione.

Un ulteriore profilo di rilievo - e specificazione di quello precedente - concerne la c.d. sostituibilità economica fra *output* generato dall'IA e contenuto editoriale originario. Secondo il *Times*, gli utenti potrebbero ottenere informazioni direttamente dal modello, diminuendo così la necessità di consultare il sito della testata e, di conseguenza,

compromettendo la sostenibilità del sistema economico degli abbonamenti.

La controversia mette in luce, sotto questo profilo, una problematica diversa rispetto ai casi concernenti opere figurative o fotografiche: il punto centrale non è soltanto la copia dell'opera, ma anche la perdita di utenti abbonati, la sostenibilità economica del giornalismo professionale e, in ultima istanza, la disintermediazione dell'accesso all'informazione.

Particolare delicatezza riveste la circostanza che il *training* abbia coinvolto, secondo l'assunto attoreo, anche articoli protetti da *paywall*: tale dato rafforza la posizione dell'attrice sotto il profilo dell'accesso non autorizzato.

D'altronde, volendo portare alle estreme conseguenze gli argomenti qui e in precedenza sviluppati, vi è da chiedersi se, e in quale misura, l'accesso lecito a un contenuto digitale implichi automaticamente la possibilità di utilizzarlo per finalità di *training* algoritmico; del pari, la possibilità - giuridicamente accettabile - di utilizzare il contenuto per finalità di *training* non "salva" l'eventuale previa acquisizione illecita del medesimo.

Infine, profilo tecnico di rilievo sotteso all'intero giudizio concerne la natura della *memorization*. In letteratura è stato concluso che la memorizzazione di porzioni dei dati di *training* costituisce un componente pressoché inevitabile e difficilmente prevenibile dell'addestramento, tanto più marcato quanto maggiore è la scala del modello. Da tale premessa, sono stati tratti in dottrina argomenti in chiave *difensiva*: da un lato, un'inibitoria che vietasse ogni forma di *memorization* equivarrebbe a un divieto pressoché totale di addestramento, con conseguente sacrificio dello

L. Rev. 259 (2023), e K. LEE, A. F. COOPER, J. GRIMMELMANN, *Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain*, 71 J. Copyright Soc'y U.S.A. (2024).

²³⁶ In questa prospettiva, l'esistenza e la rapida crescita di un mercato delle licenze per l'addestramento deporrebbe contro il riconoscimento del *fair use* proprio sotto il quarto fattore, documentando quel «potential market» che l'uso non autorizzato pregiudica. I *developer* avrebbero interesse a non concludere - o a non rendere pubbliche -

tali licenze, e finanche a qualificarle come meri «access agreements» anziché come licenze di *training*, al fine di non offrire prova pregiudizievole sul quarto fattore. È stato parimenti sostenuto che l'«impossibilità di licenziare» invocata dai *developer* costituisca una "falsa pista", riferibile semmai alla massa di contenuti non professionali del *web* e non alle opere di editori e testate giornalistiche. M. STRATTON, *Market-Based Licensing for Publishers' Works is Feasible. Big Tech Agrees.*, 48 Colum. J.L. & Arts (forthcoming 2025).

sviluppo tecnologico; dall'altro, l'assenza – almeno prima del 2021 – di una conoscenza specifica, da parte degli operatori, dei singoli episodi di violazione riverbererebbe sul requisito soggettivo della responsabilità *contributory*. È significativo – e vale la pena segnalarlo in chiave comparatistica – che i medesimi studi siano stati valorizzati in senso diametralmente opposto dal *Landgericht* di Monaco nel caso *GEMA v. OpenAI* (v. *infra*)²³⁷. Quanto alla responsabilità contributoria – le cui eccezioni sono state respinte in fase di *motion to dismiss* –, il nodo di merito è stato individuato nel requisito della effettiva consapevolezza di specifici atti di violazione, richiesto da *A&M Records v. Napster*: requisito di ardua dimostrazione, quantomeno per il periodo anteriore al 2021. A ciò si aggiunge la possibile difesa dello *staple article of commerce*, mutuata da *Sony v. Universal City Studios*, fondata sull'idoneità dei modelli a usi sostanziali non violativi; difesa cui si oppone, peraltro, il rilievo per cui – diversamente da un videoregistratore o da un computer – i LLMs recherebbero “incorporate” rappresentazioni delle opere protette. Su questo terreno, sono state prospettate misure di mitigazione idonee a contemperare l'interesse pubblico e i diritti dei titolari – segnatamente il filtraggio degli output che riproducano contenuto memorizzato e l'audit dei modelli –, che potrebbero orientare un'eventuale composizione della lite²³⁸.

VI. Conclusioni

²³⁷ Cfr. T. CHEN, *Language Models' Verbatim Copying*, cit., pp. 355 ss., che richiama N. CARLINI et al., *The Secret Sharer* (2019) e *Extracting Training Data from Large Language Models* (2021). Per l'uso di segno opposto dei medesimi studi, v. *Landgericht München I*, 11 novembre 2025, 42 O 14139/24, § 232, nonché S. von Lewinski, *GEMA v. OpenAI*, cit., p. 344.

²³⁸ T. CHEN, *Language Models' Verbatim Copying*, cit., pp. 366 ss., con richiami a *A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004 (9th Cir. 2001), e *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417 (1984); per le misure di mitigazione, ivi, pp. 368 s., e M. SAG, *Copyright Safety for Generative AI*, 61 Hous. L. Rev. 295 (2023).

Il giudizio *New York Times v. Microsoft & OpenAI* si configura, allo stato attuale, quale uno dei più significativi banchi di prova per la definizione del rapporto fra intelligenza artificiale generativa e proprietà intellettuale. La vicenda dimostra come le tecnologie generative stiano progressivamente ridefinendo i rapporti fra produzione culturale, accesso all'informazione e sfruttamento economico dei contenuti digitali. A differenza dei procedimenti riguardanti opere figurative – quali, in particolare, *Getty Images v. Stability AI* – la controversia *Times* coinvolge direttamente il settore dell'informazione professionale, mettendo in discussione la sostenibilità economica del giornalismo nell'era dell'IA.

La pronuncia interlocutoria del 4 aprile 2025 ha tracciato un primo, parziale, quadro: la quasi totalità dei capi fondati sul *copyright* è rimasta in giudizio, sì che le questioni di merito relative al *fair use* e alla *market substitution* potranno essere affrontate nella fase dibattimentale; sono invece cadute, *with prejudice*, le pretese fondate sulla *common law unfair competition*, per effetto della *preemption* federale²³⁹, e – almeno in parte – i *DMCA claims*. La caduta del primo capo appare di particolare interesse sistematico: essa indica che la tutela contro i sistemi generativi dovrà essere costruita, nel sistema statunitense, sui binari del *copyright* propriamente inteso, e non su quelli – più elastici, ma per ciò stesso meno tipizzati – della *misappropriation* di *common law*²⁴⁰.

²³⁹ La *preemption* è l'effetto, fondato sulla *Supremacy Clause* della Costituzione federale, per cui la legge federale prevale sulla normativa statale con essa incompatibile. In materia di diritto d'autore opera la *preemption* espressa di cui alla Section 301 del *Copyright Act* (17 U.S.C. § 301), che esclude le pretese di fonte statale dirette a tutelare posizioni equivalenti ai diritti esclusivi del *copyright*, su opere rientranti nel suo oggetto. È per questa ragione che il capo fondato sulla *common law unfair competition* (nella variante della *misappropriation*) è caduto *with prejudice*: trattandosi di un ostacolo di diritto insuperabile, la pretesa non era emendabile.

²⁴⁰ La *misappropriation* è una figura di matrice giurisprudenziale, ascritta al *genus* della *unfair competition* di *common law*, che sanziona

Una eventuale pronuncia di merito favorevole al *Times* potrebbe avere conseguenze sistemiche di grande rilievo, imponendo nuovi obblighi di *licensing* (dei quali si è sino ad ora molto trattato sia in dottrina che in giurisprudenza, senza che però alcun giudice si sia mai spinto ad imporre, effettivamente, l'obbligo), sistemi di remunerazione per i titolari dei diritti e limiti più stringenti all'utilizzazione dei contenuti reperiti *online*.

Viceversa, una vittoria di OpenAI e Microsoft potrebbe consolidare un'interpretazione ampia del *fair use*, rafforzando la legittimità del *training* algoritmico fondato su dati liberamente accessibili - interpretazione che peraltro convive, già allo stato attuale, con la prassi degli accordi di *licensing* stipulati dalla stessa OpenAI con diverse testate, e con la prima coerente sistematizzazione del problema offerta, in un diverso contesto, dall'ordine del giudice Alsup nel caso *Bartz v. Anthropic* del 23 giugno 2025. Nel complesso, la controversia documenta come il diritto d'autore contemporaneo sia chiamato a confrontarsi con una progressiva trasformazione della nozione stessa di «utilizzazione» dell'opera: nel contesto dell'IA generativa, il valore economico del contenuto deriva anche dalla capacità dei sistemi algoritmici di estrarre conoscenza, correlazioni linguistiche e modelli informativi da rilevanti quantitativi di dati.

l'appropriazione parassitaria del frutto dell'altrui investimento. Essa trae origine dal celebre *International News Service v. Associated Press*, 248 U.S. 215 (1918), ove la Corte Suprema riconobbe in capo all'agenzia di stampa una sorta di *quasi-property right* sulle notizie di stretta attualità (*hot news*), opponibile al concorrente che se ne fosse appropriato per diffonderle a proprio vantaggio prima che ne fosse esaurito il valore commerciale. Venuta meno, dopo *Erie Railroad Co. v. Tompkins*, 304 U.S. 64 (1938), la *federal general common law*, la dottrina è sopravvissuta quale autonoma figura di diritto statale. La sua persistente operatività è però oggi circoscritta dall'effetto della *preemption* di cui alla Section 301 del *Copyright Act* (v. la nota che precede): perché una pretesa fondata sulla *misappropriation* non sia assorbita dal *copyright* federale, essa deve presentare un *extra element* che la renda qualitativamente diversa dalla tutela autoriale. Nella formulazione divenuta canonica con *National Basketball Association v. Motorola, Inc.*, 105 F.3d 841 (2d Cir. 1997), sopravvive alla *preemption* la sola *hot-*

In tale prospettiva, sembrano destinati ad assumere crescente importanza, sul piano normativo, i sistemi di *licensing* collettivo, gli obblighi di trasparenza sui *dataset* utilizzati per il *training*, i meccanismi di *opt-out* (a livello unionale) e i corrispondenti obblighi di attribuzione delle fonti. Il giudizio *New York Times v. Microsoft & OpenAI* si configura, in conclusione, come una delle prime grandi occasioni nelle quali la giurisprudenza statunitense è chiamata a definire i confini fra apprendimento algoritmico, utilizzazione trasformativa dei contenuti e riproduzione illecita delle opere nell'era dell'intelligenza artificiale generativa.

news misappropriation, e al concorso di più elementi: che l'attore raccolga l'informazione sostenendone il costo; che essa sia *time-sensitive*; che l'uso del convenuto integri *free-riding*; che le parti siano in rapporto di concorrenza diretta; e che la possibilità di *free-riding* da parte di terzi minacci di ridurre a tal punto l'incentivo a produrre l'informazione da comprometterne l'esistenza o la qualità. È precisamente l'assenza di tali elementi ulteriori - e dunque la sostanziale equivalenza della pretesa al nucleo dei diritti esclusivi di cui alla Section 106 - a determinarne, nel caso di specie, la caduta *with prejudice*. Nell'ordinamento italiano la figura non trova un corrispondente perfetto, ma evoca la concorrenza sleale di cui all'art. 2598 c.c., e segnatamente l'ipotesi residuale del n. 3 (atti non conformi alla correttezza professionale e idonei a danneggiare l'altrui azienda), nonché l'elaborazione giurisprudenziale della c.d. concorrenza parassitaria, intesa quale imitazione sistematica e continuativa delle iniziative del concorrente.

OUTPUT INFRINGEMENT DI PERSONAGGI D'AUTORE
DISNEY ENTERPRISES, INC. ET AL. V. MIDJOURNEY, INC.
No. 2:25-cv-05275 (C.D. CAL.)
di Sara Manni

I. Premessa

L'azione *Disney Enterprises, Inc. et al. v. Midjourney, Inc.*, depositata l'11 giugno 2025 dinanzi alla *United States District Court for the Central District of California* (No. 2:25-cv-05275-JAK-AJR), segna un passaggio di considerevole rilievo nell'evoluzione del contenzioso statunitense in materia di intelligenza artificiale generativa: per la prima volta, infatti, le c.d. *major* cinematografiche statunitensi – segnatamente Disney e Universal, e poi Warner Bros. con autonoma azione successivamente consolidata – entrano direttamente in contenzioso contro un fornitore di servizi di generazione automatica di immagini, dopo aver osservato a distanza, nei due anni precedenti, le pretese parallele di artisti individuali (*Andersen v. Stability AI*), editori giornalistici (*New York*

Times v. OpenAI) ed editori musicali (*Concord Music v. Anthropic*).

Il *complaint* (di ben 110 pagine e corredato da un consistente apparato di raffronti diretti fra le opere protette degli attori e gli *output* generati dal servizio *Midjourney* ad esse affiancate) articola le pretese su un duplice piano: quello del *direct copyright infringement* e quello del *secondary copyright infringement*²⁴¹. Sotto entrambi i profili, le attrici contestano alla convenuta l'avvenuta utilizzazione delle proprie opere protette tanto nella fase di addestramento del modello, quanto in quella della successiva generazione di *output* sostanzialmente coincidenti - o derivati - con i personaggi cinematografici di loro titolarità²⁴². Conviene anticipare, per ragioni di chiarezza espositiva, che il procedimento si trova attualmente in fase di *discovery*, non essendovi a oggi alcuna decisione di merito né alcuna pronuncia su *motions to dismiss*: ciò sicuramente impone, nella presente sede, una valutazione necessariamente provvisoria e proiettata sulle prospettive del giudizio.

²⁴¹ Il documento è disponibile in *Justia*, all'indirizzo <<https://news.justia.com/wp-content/uploads/2025/06/Disney-Universal-v.-Midjourney-Complaint.pdf>>.

²⁴² Il diritto statunitense distingue la violazione diretta (*direct infringement*) da quella secondaria o indiretta (*secondary infringement*). La prima, disciplinata dalla Section 501 del *Copyright Act* (17 U.S.C. § 501), consiste nella lesione di uno dei diritti esclusivi enumerati alla Section 106 (riproduzione, elaborazione, distribuzione, esecuzione ed esposizione pubblica) ed è costruita come ipotesi di responsabilità oggettiva (*strict liability*): non occorre la prova del dolo o della colpa, essendo sufficiente una condotta volitiva (*volitional conduct*) che cagioni l'atto lesivo, mentre la consapevolezza rileva semmai sul piano della *willfulness* e della commisurazione del danno. La seconda, di elaborazione giurisprudenziale e non codificata, presuppone un'altrui violazione diretta e si articola in due figure: il *contributory infringement* - responsabilità di chi, conoscendo l'attività illecita, vi induce o vi contribuisce in modo rilevante, comprensiva dell'*inducement* di matrice *Grokster* - e il *vicarious infringement* - responsabilità di chi abbia il diritto e la possibilità di controllare l'attività illecita e ne tragga un beneficio economico diretto, a prescindere dalla conoscenza.

Nell'ordinamento italiano la distinzione non è del pari tipizzata. La violazione dei diritti di utilizzazione economica (artt. 12 ss. l. 22 aprile 1941, n. 633 - segnatamente la riproduzione, art. 13, la comunicazione al pubblico, art. 16, la distribuzione, art. 17, l'elaborazione, art. 18) legittima i rimedi degli artt. 156 ss. l.d.a.: l'inibitoria e la rimozione prescindono dalla colpa e poggiano sulla mera lesione oggettiva del diritto, mentre il risarcimento, di natura aquiliana (art. 2043 c.c.), esige il dolo o la colpa - donde una divergenza significativa rispetto alla *strict liability* statunitense, che peraltro si attenua proprio sul versante inibitorio. La responsabilità del terzo che concorra nell'illecito si costruisce, in difetto di una figura autonoma di violazione secondaria, sulle categorie generali del concorso e della solidarietà (art. 2055 c.c.). Per gli intermediari della rete - i soggetti più prossimi alle piattaforme di IA generativa - soccorre invece il regime speciale di esenzione dell'*hosting* (artt. 14-17 d.lgs. 70/2003, oggi nel quadro del *Digital Services Act*, reg. UE 2022/2065), fondato sulla logica del *notice and take down*, nonché, per le piattaforme di condivisione di contenuti caricati dagli utenti, lo statuto *sui generis* dell'art. 17 della dir. (UE) 2019/790, attuato in Italia con il d.lgs. 177/2021: disciplina che, riconducendo in capo al prestatore un'autonoma responsabilità per i contenuti protetti resi accessibili, segna un punto di convergenza funzionale con la problematica della *secondary liability*, centrale nel contenzioso sull'intelligenza artificiale.

La convenuta ha depositato la propria *answer* nell'agosto 2025, articolando le proprie difese intorno alla *doctrine of fair use* e ad altre eccezioni²⁴³.

Il procedimento è stato inoltre formalmente consolidato - con ordine del giudice Kronstadt del novembre 2025 - con l'autonoma azione promossa nel settembre 2025 da Warner Bros. Discovery e affiliate (*Warner Bros. Entertainment Inc. v. Midjourney, Inc.*, No. 2:25-cv-08376-JAK-E)²⁴⁴.

II. Le parti del giudizio

Parti attrici sono sette società del comparto cinematografico statunitense, riconducibili a due distinti gruppi.

Dal lato Disney, figurano: Disney Enterprises, Inc. - titolare dei diritti d'autore sui personaggi del catalogo Disney *stricto sensu* e del comparto animazione classica; Marvel Characters, Inc., titolare dei diritti relativi ai personaggi del *Marvel Cinematic Universe*; MVL Film Finance LLC, che gestisce taluni profili economici dei *Marvel Studios*; Lucasfilm Ltd. LLC, titolare dei diritti del franchise *Star Wars*; Twentieth Century Fox Film Corporation, titolare segnatamente del catalogo *The Simpsons*.

Dal lato Universal, figurano Universal City Studios Productions LLLP e DreamWorks Animation L.L.C., titolari dei diritti relativi ai personaggi dei rispettivi cataloghi cinematografici e di animazione, fra i quali

figurano *Shrek*, i *Minions*, *Madagascar* e altri analoghi titoli.

Parte convenuta è Midjourney, Inc., società di diritto del Delaware operante nel settore dell'intelligenza artificiale generativa per immagini. Il servizio *Midjourney*, accessibile mediante *interface* sulla piattaforma *Discord* e successivamente - a partire dal 2024 - anche tramite *web app* dedicata, consente all'utente di generare immagini a partire da *text prompts*. Il servizio funziona - secondo la stessa convenuta - sulla base di *diffusion models* addestrati su una rilevante quantità di immagini reperite *online*, il cui esatto perimetro non è stato sino ad oggi reso pubblico dalla società. Secondo quanto allegato nel *Complaint*, il servizio *Midjourney* avrebbe prodotto, nel solo 2024, ricavi pari a circa trecento milioni di dollari.

III. L'origine della controversia e l'atto introduttivo

L'*incipit* del *Complaint* - depositato l'11 giugno 2025 e del quale a seguire si riporta il testo - è formulato in toni particolarmente incisivi: «[p]iracy is piracy, and whether an infringing image or video is made with AI or another technology does not make it any less infringing» (la pirateria è pirateria e che un'immagine plagiante sia fatta con l'IA o con altra tecnologia non la rende in alcun modo meno plagiante). Le attrici qualificano *Midjourney* come «*the quintessential copyright free-rider*» (il parassita

²⁴³ Per una ricostruzione delle difese svolte da Midjourney nell'*answer* - impennate (i) sulla qualificazione del *training* come uso trasformativo e *fair use*; (ii) sulla natura di mero strumento della piattaforma, con conseguente imputazione all'utente della generazione dell'*output*; (iii) sul difetto di prova di una sistematica produzione di «copie» dei personaggi protetti; e (iv) sull'assenza di *market harm*, non ponendosi gli *output* quali sostituti delle opere originarie - v., riportando ampi stralci dell'atto, A. KILDEEV, *Disney/Warner Brothers Discovery v. Midjourney and Raine v. OpenAI & Altman: Lawsuits and Implications for the Development of the AI Industry*, in *Scientific Platform XXI Century*, n. 4/2025, 20 ss. La convenuta invoca altresì la procedura di *notice-and-takedown* di cui al DMCA, addebitando alle attrici di non avervi fatto ricorso prima dell'azione.

²⁴⁴ Profilo ancora non emerso nel contenzioso sul tema e che - auspicabilmente - potrà esser sollevato nel caso di

specie è quello della responsabilità dell'utente. Difatti, premesso che il sistema di IA generativa - e dunque il relativo fornitore - è responsabile della "possibilità" stessa di violare il diritto d'autore sui personaggi *de quibus*, vi è da chiedersi quale sia il ruolo (e se ce ne sia alcuno) dei soggetti che, consapevolmente, utilizzano i sistemi in questione per creare immagini illecite di personaggi d'autore. Sulla responsabilità dell'utente per violazione del diritto d'autore mediante utilizzo di output generati da IA che si qualificano come opere derivate si rimanda a M. F. TOMMASINI, *ChatGPT e diritto d'autore. Profili giuridici dell'intelligenza artificiale generativa*, in *Diritto di famiglia e delle persone*, 2026, 271 ss. Peraltro, il contestato illecito in commento configura una forma di plagio evolutivo, tematica rispetto alla quale si rinvia a G. CASSANO, B. TASSONE, *La disciplina del plagio evolutivo tra effettività della tutela del diritto di autore ed evoluzioni in tema di diritto morale*, in *Dir. Ind.*, 2021, 214 ss.

del diritto d'autore per antonomasia), un «*bottomless pit of plagiarism*» (un pozzo senza fondo di plagio), riconducendo il servizio convenuto al modello concettuale di una «*virtual vending machine*» (un distributore automatico virtuale) per la produzione di copie illecite.

« 1. For more than 100 years, Disney and Universal have delighted audiences around the world by investing in and fostering American creative innovation and producing some of the greatest motion pictures and fictional characters of all time. Midjourney, however, seeks to reap the rewards of Plaintiffs' creative investment by selling an artificial intelligence ("AI") image-generating service ("Image Service") that functions as a virtual vending machine, generating endless unauthorized copies of Disney's and Universal's copyrighted works. 2. By helping itself to Plaintiffs' copyrighted works, and then distributing images (and soon videos) that blatantly incorporate and copy Disney's and Universal's famous characters—without investing a penny in their creation—Midjourney is the quintessential copyright free-rider and a bottomless pit of plagiarism. Piracy is piracy, and whether an infringing image or video is made with AI or another technology does not make it any less infringing. Midjourney's conduct misappropriates Disney's and Universal's intellectual property and threatens to upend the bedrock incentives of U.S. copyright law that drive American leadership in movies, television, and other creative arts. 3. Midjourney's infringement is calculated and willful. Plaintiffs have asked Midjourney to stop infringing their copyrighted works and, at a minimum, to adopt technological measures, which other AI services have implemented to prevent the generation of infringing material. But Midjourney, which has attracted millions of subscribers and made \$300 million last year alone, is focused on its own bottom line and ignored Plaintiffs' demands. 4. Instead, Midjourney has chosen to double down on its unlawful actions by releasing and promoting even newer versions of its Image Service and teasing its soon-to-be-released commercial AI video service ("Video Service"). On information and belief, without Plaintiffs' authorization, Midjourney's Video Service will generate,

publicly display, and distribute videos featuring Disney's and Universal's copyrighted characters. Midjourney has already begun training its Video Service, meaning that Midjourney is very likely already infringing Plaintiffs' copyrighted works in connection with its Video Service.

5. Disney's and Universal's vast libraries of intellectual property include some of the most well-known movies, television shows, and fictional characters of all time. For Disney, these include the Star Wars franchise, the Marvel Comics properties, Pixar films, the 20th Century Studios (formerly 20th Century Fox) library, and Disney's namesake animated features. Disney's unforgettable copyrighted characters span decades of works across a wide range of genres and tastes: Darth Vader from Star Wars, Elsa from Frozen, Lightning McQueen from Cars, Buzz Lightyear from Toy Story, Sully from Monster's Inc., Marvel's Iron Man, and Homer Simpson from The Simpsons, to name just a few.

6. Universal—the oldest continuously operating movie studio in the United States—is the home to some of the most popular and enduring animated movies and beloved characters that have become part of modern American popular culture. These include Shrek from the Academy Award-winning Shrek and its sequels and spin-offs, and the Minions characters from Despicable Me and Minions movies (the highest grossing animated franchise to date), as well as other beloved and famous characters from critically acclaimed film franchises like Hiccup and Toothless from How to Train Your Dragon, Po from Kung Fu Panda, and the Boss Baby from The Boss Baby.

7. Disney's and Universal's works are built on decades of financial investment, imagination, and innovation. Few American entertainment companies have embraced creativity and technology like Plaintiffs, from the earliest days of using technology to synchronize sound and film, to creating virtual and immersive experiences at their world class theme parks, to pushing the boundaries of animation through groundbreaking rendering software, to advanced machine learning for de-aging actors in movies.

8. These investments of time and money are only made possible by the incentives embodied in U.S. copyright law, which

ensure copyright owners the exclusive right to control and commercialize their own creative works. Whether it is a children's coloring book with Marvel superheroes, a video game based on the Shrek movies, or a subscription service that distributes high resolution images or short videos with Plaintiffs' copyrighted characters, only Plaintiffs are allowed to build a business around or otherwise commercialize those characters. Midjourney's Image Service was developed using innumerable unauthorized copies of Plaintiffs' copyrighted works, and it operates by reproducing, publicly displaying, making available, and distributing additional infringing copies and derivatives of those works»²⁴⁵.

Sotto il profilo delle allegazioni, il *complaint* presenta una struttura assai analitica e - è opportuno segnalarlo - innovativa rispetto ai precedenti atti introduttivi nel medesimo campo.

Le attrici hanno infatti accluso al loro atto una considerevole quantità di immagini in cui i personaggi protetti - fra i quali figurano *Darth Vader*, *Yoda*, *Princess Elsa*, *Buzz Lightyear*, *Bart Simpson*, *Shrek*, i *Minions* e una pluralità di personaggi del catalogo *Marvel* - risultano riprodotti in *output* generati da *Midjourney* in risposta a *prompt* mirati. La forza probatoria di tale apparato - la cui ricchezza visiva costituisce uno degli elementi distintivi della causa rispetto agli analoghi procedimenti contro Stability AI o contro OpenAI - è destinata, presumibilmente, a giocare un ruolo determinante nella fase di merito.

Un secondo elemento di considerevole interesse concerne il presunto carattere calcolato e volontario della violazione.

Le attrici hanno documentato di aver formalmente diffidato *Midjourney* sin dalla fase preliminare al deposito dell'atto, invitando la società quanto meno ad adottare quelle misure tecniche che, a tenore della citazione, altri fornitori di IA hanno implementato, ossia *guardrails* atti a impedire la riproduzione dei personaggi protetti - del genere di quelli, ad

esempio, adottati da Anthropic in materia di *lyrics* musicali in seguito alla controversia *Concord* -.

«11. Before taking legal action, Plaintiffs asked Midjourney to stop its theft of their intellectual property, including by implementing these simple measures. Instead of stopping its infringement, Midjourney has continued to release new versions of its Image Service, which, according to Midjourney's founder and CEO, have even higher quality infringing images»²⁴⁶.

Le diffide sarebbero rimaste senza riscontro, sicché la qualificazione della condotta illecita come consapevole («*willful*») apre la porta - in caso di accertamento della violazione - all'irrogazione dei *statutory damages* massimi previsti dalla Section 504(c)(2) del *Copyright Act*, pari a 150.000 dollari per ciascuna opera violata, con conseguenze economiche potenziali di rilevantissima portata.

IV. I capi di domanda

Il primo capo di domanda concerne la violazione diretta del *copyright* ai sensi della Section 501 del *Copyright Act* (17 U.S.C. § 501), in relazione ai diritti esclusivi di cui alla Section 106. La pretesa si articola, a sua volta, su due distinti piani: quello dell'*input infringement* - riproduzione delle opere nella fase di addestramento del modello - e quello dell'*output infringement* - successiva riproduzione delle opere nei contenuti generati dal servizio in risposta agli *user prompts*.

L'*input infringement* è argomentato sulla base della contestata utilizzazione, nella fase di *training*, di immagini protette degli attori - non quindi su quella di immagini liberamente accessibili, ma di vere e proprie riproduzioni dei personaggi del catalogo, ottenute attraverso *scraping* del *web* o mediante *dataset* terzi. Le attrici non sono in grado di documentare in via diretta - non avendo accesso alla *discovery* - la specifica composizione del *training set* della

²⁴⁵ Case 2:25-cv-05275 Document 1 depositato l'11 giugno 2025, 1-8, disponibile in *justia*.

²⁴⁶ *Ibidem*, 11.

convenuta; ne deducono però la presenza delle opere protette attraverso un argomento di tipo indiziario, fondato sulla circostanza che il modello, in risposta a *prompt* anche minimi, produce immagini dei personaggi protetti con un grado di accuratezza tale da rendere implausibile l'ipotesi che esse derivino unicamente da correlazioni statistiche generiche²⁴⁷.

L'*output infringement*, invece, è documentato direttamente mediante la produzione delle immagini generate dal servizio. Sotto questo profilo, gli attori allegano che la riproduzione dei personaggi protetti negli *output* avviene – e questo è un dato di particolare interesse – anche in assenza di *prompt* esplicitamente diretti a quei personaggi: il modello, in altri termini, sarebbe in grado di generare immagini di Darth Vader o di Yoda anche in risposta a *prompt* descrittivi generici (ad esempio, «*Sith lord with red lightsaber*»), suggerendo l'avvenuta interiorizzazione dei personaggi nelle rappresentazioni statistiche del modello.

Il secondo capo di domanda concerne la violazione secondaria del *copyright*, articolata nelle due fattispecie giurisprudenziali del *contributory infringement* e del *vicarious infringement*, elaborate dalla giurisprudenza statunitense rispettivamente a partire dai casi *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417 (1984), e *MGM Studios, Inc. v. Grokster, Ltd.*, 545 U.S. 913 (2005).

Sul versante della responsabilità secondaria, la struttura argomentativa delle attrici trova conferma in una lettura dottrinale recente, sviluppata con riguardo a un'analogia tipologia di servizi generativi (i *character chatbot*): anche in quella sede si è escluso che ricorra una responsabilità diretta del fornitore - difettando un controllo dell'utente sull'*output* - e si è

ritenuta arduamente configurabile la responsabilità contributoria alla luce di *Cox Communications v. Sony*, che esclude l'automatica responsabilità del fornitore per la mera consapevolezza dell'uso illecito, individuando per converso nella *vicarious liability* il titolo più appropriato, in ragione tanto del diritto e della capacità della piattaforma di controllare l'attività illecita - mediante la selezione dei dati di addestramento, la progettazione del sistema e l'implementazione di filtri - quanto del beneficio economico diretto che essa ne ritrae. La medesima logica è agevolmente trasponibile alla posizione di Midjourney, la cui capacità di adottare *guardrails* e il cui modello a pagamento integrano i due presupposti del *vicarious infringement*²⁴⁸.

Sul piano del *contributory infringement*, le attrici contestano alla convenuta di aver materialmente contribuito alle violazioni realizzate dagli utenti del servizio, mettendo a loro disposizione uno strumento idoneo a generare contenuti illeciti, nella piena consapevolezza - documentata dalle diffide ignorate - del rischio di violazione e in assenza dell'adozione di misure tecniche atte a prevenirle.

Sul piano del *vicarious infringement*, le attrici sostengono che Midjourney abbia tratto vantaggio economico diretto dalla generazione di *output* illeciti - atteso che il servizio è offerto a pagamento agli utenti - e abbia avuto il «*right and ability to supervise*» l'attività infrigente, in quanto materialmente in grado, mediante l'implementazione di *guardrails*, di prevenire la generazione di *output* coincidenti con i personaggi protetti.

Il quadro della responsabilità secondaria è stato peraltro ridisegnato, in pendenza del giudizio,

²⁴⁷ L'argomento attoreo presenta un'evidente affinità con quello svolto, *mutatis mutandis*, nel parallelo procedimento *Andersen v. Stability AI Ltd. et al.*, No. 3:23-cv-00201-WHO (N.D. Cal.), dinanzi al giudice William H. Orrick, nel quale gli attori - artisti di arti visive - hanno costruito la propria pretesa sulla circostanza che gli *output* del modello *Stable Diffusion* presentano una somiglianza con le opere originarie tale da rendere implausibile l'assenza delle stesse nel *training dataset*. In tale

procedimento, i *direct infringement claims* sono sopravvissuti al *motion to dismiss* e si trovano attualmente in fase di *discovery*.

²⁴⁸ Per l'applicazione delle tre teorie di responsabilità ai *character chatbot*, Herrera, *Infringement Through Imitation*, cit., pp. 937 ss.

dalla Supreme Court: con la decisione del 25 marzo 2026 in *Cox Communications, Inc. v. Sony Music Entertainment*, 607 U.S. ____ (2026), la Corte - all'unanimità quanto all'esito, con opinione del Justice Thomas e concurrence delle Justices Sotomayor e Jackson - ha riformato la pronuncia del Fourth Circuit, statuendo che il fornitore di un servizio offerto al pubblico non risponde a titolo di *contributory infringement* per la mera consapevolezza che taluni utenti se ne serviranno per violare il copyright: la responsabilità contributoria presuppone l'intento, dimostrabile soltanto mediante l'induzione alla violazione ovvero l'offerta di un servizio specificamente congegnato (*tailored*) in funzione di essa. La decisione lascia espressamente impregiudicata la *vicarious liability*, sicché essa vale semmai a corroborare l'impostazione dottrinale sopra richiamata: nel giudizio Disney, la teoria contributoria fondata sulla mera conoscenza - diffide ignorate comprese - non appare più, di per sé, sufficiente, e il baricentro della responsabilità secondaria è destinato a spostarsi ulteriormente verso la *vicarious liability*, ovvero a richiedere alle attrici la ben più impegnativa dimostrazione che il servizio sia stato indotto o congegnato in funzione della violazione.

V. Le difese di Midjourney

La convenuta, nella costituzione depositata nell'agosto 2025²⁴⁹, ha articolato la propria strategia difensiva intorno a una pluralità di *affirmative defenses*, la cui ricognizione

consente di anticipare alcune delle questioni centrali del giudizio²⁵⁰.

In primo luogo, la *doctrine of fair use* è stata invocata quale prima *affirmative defense*, sull'assunto che il *training* del modello costituirebbe una utilizzazione altamente trasformativa delle opere – assimilabile, *mutatis mutandis*, a un processo di apprendimento di natura umana – e dunque coperta dalla Section 107 del *Copyright Act*. È una linea argomentativa coerente con quella sviluppata da OpenAI nei procedimenti paralleli, e ricalca per più aspetti l'impostazione difensiva adottata da Anthropic nel caso *Bartz*.

In secondo luogo, *Midjourney* ha sostenuto di non controllare i contenuti generati dagli utenti del servizio, e dunque di non potersi qualificare quale soggetto direttamente responsabile delle violazioni eventualmente commesse mediante i *prompt* degli utenti. La linea difensiva richiama, sul piano dottrinale, il dibattito sulla responsabilità delle piattaforme intermedie quale sviluppato a partire dalla giurisprudenza statunitense sui c.d. *safe harbors* di cui alla Section 512 del DMCA – fermo restando che il regime applicabile è destinato ad essere oggetto di acceso confronto.

In terzo luogo, la convenuta ha sollevato l'eccezione di *unclean hands*, fondata sull'allegazione che le stesse attrici utilizzano e traggono beneficio da modelli di IA generativa - ivi compreso lo stesso servizio Midjourney - anche in relazione a talune delle opere azionate²⁵¹.

²⁴⁹ Case 2:25-cv-05275-JAK-AJR, Document 20, depositata il 6 agosto 2025.

²⁵⁰ Nel processo civile statunitense l'*affirmative defense* è la difesa con cui il convenuto, senza limitarsi a negare i fatti costitutivi della pretesa attorea, introduce fatti ulteriori che - ove provati - ne paralizzano o ne riducono l'efficacia, ancorché le allegazioni dell'attore siano vere. Essa va sollevata, a pena di rinuncia (*waiver*), nel primo atto difensivo (*answer*), secondo l'elenco esemplificativo della Rule 8(c) *Fed. R. Civ. P.* (che annovera, fra l'altro, *statute of limitations*, *laches*, *license*, *estoppel* e *res judicata*), e l'onere della relativa prova grava sul convenuto. In materia di copyright vi appartengono, in particolare, il *fair use*, la *license* e le *unclean hands*. La figura corrisponde, nel nostro ordinamento, all'eccezione

in senso stretto (o «propria»), rilevabile solo su istanza di parte e soggetta a preclusione se non tempestivamente proposta nella comparsa di risposta (art. 167 c.p.c.): anche nel sistema italiano, del resto, chi eccepisce l'inefficacia, la modificazione o l'estinzione del diritto altrui deve provare i fatti su cui l'eccezione si fonda (art. 2697, comma 2, c.c.).

²⁵¹ Le *unclean hands* (mani sporche) costituiscono una *affirmative defense* (v. nota immediatamente precedente) di natura equitativa, fondata sul brocardo per cui chi adisce l'*equity* deve presentarsi con le mani pulite: la parte che abbia tenuto, in relazione alla stessa vicenda dedotta in giudizio, una condotta scorretta o in mala fede può vedersi negare la tutela equitativa richiesta - segnatamente l'*injunction* -. Secondo la giurisprudenza, la scorrettezza rileva soltanto ove presenti una relazione

Sotto il profilo sostanziale, è prevedibile che la difesa di parte convenuta si concentri, nella fase di merito, sulla difficoltà di documentare in via diretta la concreta presenza delle opere protette nel *training set*. La giurisprudenza statunitense più recente – segnatamente l’*order* del giudice Alsup in *Bartz v. Anthropic* del giugno 2025 – ha mostrato di dare considerevole peso alla qualificazione delle modalità di acquisizione dei *dataset*, e la convenuta potrebbe dunque trarre vantaggio da una linea argomentativa che distingue nettamente fra *training use* e *dataset acquisition*.

VI. I profili giuridici di maggior rilievo

Un primo profilo di considerevole interesse concerne la peculiare qualificazione giuridica dei personaggi cinematografici quali opere protette ai fini del *copyright* statunitense. La giurisprudenza statunitense – segnatamente i casi *Walt Disney Productions v. Air Pirates*, 581 F.2d 751 (9th Cir. 1978), e *DC Comics v. Towle*, 802 F.3d 1012 (9th Cir. 2015) – ha elaborato un quadro analitico sofisticato sulla protezione dei personaggi caratterizzati da elementi visivi e narrativi distintivi («*sufficiently delineated characters*»).

La sofisticazione di tale quadro analitico è stata di recente ricostruita in dottrina, che ne ha isolato la successione dei test elaborati dalle corti federali: dal *distinctly delineated test* di *Nichols v. Universal Pictures* - ove Judge Learned Hand affermò che quanto meno un personaggio è sviluppato, tanto meno esso è tutelabile – al *Towle test* del Ninth Circuit, che subordina la protezione a tre requisiti cumulativi (qualità fisiche e concettuali; sufficiente delineazione con tratti costanti che rendano il personaggio riconoscibile in ogni sua apparizione; distintività rispetto a un personaggio-tipo), fino allo *stock character test* di *Gaiman v. McFarlane*. Il filo conduttore è che il personaggio è tutelato per

una «essenza» espressiva che permane stabile e riconoscibile attraverso le sue molteplici incarnazioni – profilo che, come si dirà, assume specifico rilievo rispetto agli output generati dall’IA²⁵².

La specificità degli attori *Disney* e *Universal* – titolari di personaggi iconici, facilmente riconoscibili, e dotati di una densità espressiva di particolare ampiezza – fornisce loro un terreno argomentativo di considerevole solidità, ben più favorevole rispetto a quello sul quale, ad esempio, ha potuto contare un fotografo *stock* nel caso *Kneschke*.

Un secondo profilo di interesse concerne il *baricentro* probatorio del giudizio. A differenza dei procedimenti instaurati da autori letterari o da editori giornalistici – in cui la difficoltà di documentare l’*output infringement ex ante* ha talora indotto a concentrare l’attenzione sull’*input infringement* e sulla mera utilizzazione dei *dataset* – la causa *Disney* presenta un apparato probatorio *output-centric* di considerevole forza.

Rispetto a tale profilo, risulta possibile prospettare – sia pure con riferimento ai *character chatbot* testuali – una vera e propria presunzione di sostanziale somiglianza: poiché i sistemi generativi sono deliberatamente progettati per raschiare il materiale sorgente e riprodurre il personaggio nel modo più fedele possibile, gli output dovrebbero presumersi sostanzialmente simili all’originale senza necessità di condurre la *substantial similarity analysis* per ciascun output. Trasposto sul terreno visivo del giudizio *Disney*, l’argomento – se accolto – accentuerebbe ulteriormente lo spostamento del baricentro probatorio verso l’output, sottraendo il giudizio alle complesse questioni di *discovery* sui *dataset* che caratterizzano i procedimenti paralleli²⁵³.

Un terzo profilo concerne la posizione di *Disney v. Midjourney* nel quadro complessivo della

immediata e necessaria con l’oggetto della controversia e con il rimedio invocato, non bastando una generica censurabilità della condotta (*Keystone Driller Co. v. General Excavator Co.*, 1933; *Precision Instrument Mfg. Co. v. Automotive Maintenance Machinery Co.*, 1945).

²⁵² I. HERRERA, *Infringement Through Imitation: AI Chatbots and the Characters They Infringe*, in *Fordham Intell. Prop. Media & Ent. L.J.*, vol. 36, 2026, p. 905, spec. pp. 927 ss.

²⁵³ *Ibid.*, 935 s.

prima stagione del contenzioso *AI*. A pochi mesi dal deposito dell'atto introduttivo, Warner Bros. Discovery – terza delle c.d. «*Big Five*» dell'industria cinematografica²⁵⁴ – ha proposto autonoma azione, formalmente analoga, contro *Midjourney*, e i due procedimenti sono stati consolidati con ordine del giudice *Kronstadt* del 4 novembre 2025. Ciò significa, sul piano sostanziale, che la convenuta dovrà difendersi non più dalle pretese di sei attori, ma di tredici, con un complessivo *repertoire* di personaggi e opere protette esponenzialmente più ampio – comprensivo, fra l'altro, dei franchise *DC Comics*, *Looney Tunes*, *Hanna-Barbera* e di numerosi altri.

La traiettoria del contenzioso suggerisce, peraltro, che l'ingresso delle *major* cinematografiche nel campo dell'IA generativa segna l'apertura di una stagione contenziosa qualitativamente diversa da quella inaugurata da autori individuali e editori giornalistici. Le *major* sono dotate di *repertoire* iconici, di solidi *track record* giurisprudenziali in materia di tutela dei personaggi e di considerevoli risorse processuali – sicché il giudizio è destinato a costituire un banco di prova privilegiato per la verifica della tenuta delle difese fondate sul *fair use* dei *providers* di IA generativa per immagini.

VII. Conclusioni

L'azione *Disney Enterprises, Inc. et al. v. Midjourney, Inc.* si presenta – pur nella fase ancora *interlocutoria* del giudizio – come uno dei procedimenti potenzialmente più significativi nell'evoluzione della giurisprudenza statunitense sull'IA generativa, e ciò per una pluralità di ragioni convergenti.

Primo. Il giudizio segna l'ingresso a pieno titolo delle *major* cinematografiche nel contenzioso, e con esso l'ampliamento del fronte degli attori interessati alla regolamentazione del rapporto fra *training* algoritmico e proprietà intellettuale. Le risorse processuali, l'*expertise* giurisprudenziale e il *repertoire* iconico delle

attrici rendono il giudizio un banco di prova qualitativamente differente dai precedenti.

Secondo. La struttura del *Complaint* – fortemente *output-centric*, e fondata su un apparato visivo di considerevole forza probatoria – solleva con particolare nitidezza la questione della responsabilità del *provider* per la generazione di *output* sostanzialmente coincidenti con opere protette, anche in assenza di *prompt* specificamente diretti a tali opere. La questione è destinata a costituire, ad avviso di chi scrive, uno dei nodi più delicati dell'elaborazione futura.

La centralità della questione dell'output - e della responsabilità del fornitore per output coincidenti con opere protette - non è peculiare dell'ordinamento statunitense. Sul versante europeo, il *Landgericht* di Monaco, nel coevo caso *GEMA v. OpenAI*, ha affermato la responsabilità *diretta* del fornitore per le riproduzioni nel modello e negli output, muovendo peraltro da categorie estranee al diritto statunitense (il diritto di riproduzione e l'eccezione di *text and data mining*). Pur nella diversità degli strumentari - negli Stati Uniti il baricentro è sulla *secondary liability* e sul *fair use*, in Europa sulla riproduzione e sulle eccezioni -, le due esperienze convergono nell'individuare nei mercati delle licenze la soluzione di equilibrio più appropriata, in luogo tanto di un uso incontrollato quanto di un'eccezione ai diritti esclusivi²⁵⁵.

Sul versante statunitense, peraltro, la richiamata decisione della Supreme Court in *Cox* è destinata a restringere lo spazio della responsabilità contributoria, concentrando la partita - anche nel giudizio Disney - sulla responsabilità diretta per gli output e sulla *vicarious liability*.

Terzo. Occorrerà determinare come il presente giudizio si collocherà nell'orizzonte giurisprudenziale che si sta sviluppando sul tema (in particolare, il riferimento è a *Andersen v. Stability AI*, in cui i *direct infringement claims* sono già sopravvissuti alla *motion to dismiss*, e

²⁵⁴ I.e., The Walt Disney Studios, Warner Bros. Entertainment, Universal Pictures, Sony Pictures, Paramount Pictures.

²⁵⁵ Cfr. S. VON LEWINSKI, *GEMA v. OpenAI: The Munich Regional Court Presents an Exemplary Judgment on a Well Conceived Test Case*, in *GRUR Int.*, vol. 75, n. 4, 2026, p. 340.

Bartz v. Anthropic, definito mediante *settlement*), sicché il giudizio *de quo* potrà ricoprire un ruolo di rilievo nell'elaborazione di un quadro coerente del contenzioso AI davanti alle corti federali statunitensi.

RIPRODUZIONE, RESPONSABILITÀ SECONDARIA, TRADE MARK E TERRITORIALITÀ DEL TRAINING DINANZI ALLA HIGH COURT

GETTY IMAGES (US) INC. & OTHERS V. STABILITY AI LTD., [2025] EWHC 2863 (Ch)

di Sara Manni e Leonardo Ortenzi*

I. Premesse

La sentenza pronunciata il 4 novembre 2025 dalla *High Court of Justice of England and Wales, Chancery Division*, nel caso *Getty Images (US) Inc. & Others v. Stability AI Ltd.* - recante *neutral citation* [2025] EWHC 2863 (Ch) e numero di registro IL-2023-000007 - costituisce, allo stato attuale, uno dei più rilevanti tentativi di applicazione delle categorie tradizionali del diritto della proprietà intellettuale ai modelli di intelligenza artificiale generativa. La decisione, pronunciata da Mrs Justice Joanna Smith DBE, assume un valore paradigmatico, affrontando direttamente la questione del rapporto fra apprendimento automatico, memorizzazione delle opere e generazione degli *output*, offrendo una delle prime interpretazioni giudiziali compiute della struttura tecnica dei *diffusion models*, e segnatamente di *Stable Diffusion*²⁵⁶. Conviene anticipare, in via di sintesi, che nel corso del dibattimento Getty Images ha abbandonato i capi di domanda relativi al *primary copyright infringement* - tanto nella componente riferita al training e allo sviluppo del modello (per difetto di prova che tali attività fossero avvenute nel Regno Unito), quanto in quella riferita agli output (avendo Stability bloccato i prompt utilizzati per generare gli esemplari contestati, con sostanziale

conseguimento *aliunde* della tutela richiesta) - nonché il database *right infringement*, non più coltivabile in ragione del suo legame intrinseco con le prime due componenti²⁵⁷; e si è vista respingere, già con sentenza del 14 gennaio 2025 ([2025] EWHC 38 (Ch)), la richiesta di agire in via rappresentativa per conto di un considerevolissimo numero di fotografi terzi²⁵⁸; alle attrici è stato nondimeno consentito di procedere, ai sensi della CPR 19.3(1) e della Section 102(1) CDPA, senza chiamata in causa dei singoli licenzianti, pertanto il *thema decidendum* si è progressivamente concentrato sul *secondary copyright infringement* e sul *trade mark infringement*. Proprio sotto il *trade mark claim* Getty ha conseguito un esito parzialmente favorevole, qualificato dal giudice come «*both historic and extremely limited*». Sul *secondary copyright infringement*, invece, la pronuncia ha rigettato le pretese attoree, ritenendo che i *model weights* di *Stable Diffusion* non costituiscano «*infringing copies*» delle opere utilizzate per l'addestramento.

II. Le parti del giudizio

Il giudizio vede contrapposte, da un lato, sei attrici titolari di diritti - Getty Images (US) Inc., Getty Images International U.C., Getty Images (UK) Limited, Getty Images Devco UK Limited, iStockphoto LP e Thomas M. Barwick, Inc. - alle quali si farà nel prosieguo riferimento, per brevità, come «Getty» o «le attrici»; dall'altro, Stability AI Ltd., società di diritto inglese operante nel settore dell'intelligenza artificiale generativa.

* Premesso che l'opera è frutto del lavoro congiunto dei due autori, a Leonardo Ortenzi vanno attribuiti i Paragrafi I e II e a Sara Manni i Paragrafi III, IV e V.

²⁵⁶ La sentenza del 4 novembre 2025 è pubblicata sul sito *web* del Judiciary inglese, disponibile all'indirizzo <chrome-extension://efaidnbmninnibpcjpcglclefindmkaj/https://www.judiciary.uk/wp-content/uploads/2025/11/Getty-Images-v-Stability-AI.pdf>.

²⁵⁷ Stante il considerevole spessore della pronuncia analizzata, di oltre duecento pagine, si indicheranno in nota i punti di riferimento del provvedimento per il

lettore. Approved Judgment of 4 november 2025, [2025] EWHC 2863 (Ch), Case No: IL-2023-000007, High Court Of Justice Business And Property Courts Of England And Wales Intellectual Property List (ChD), Mrs Justice Joanna Smith DBE, § 9 (in avanti, Sent. 4 novembre 2025), disponibile all'indirizzo <https://www.judiciary.uk/wp-content/uploads/2025/11/Getty-Images-v-Stability-AI.pdf>.

²⁵⁸ Approved Judgment of 14 January, 2025 before Mrs Justice Joanna Smith DBE, disponibile all'indirizzo <https://www.judiciary.uk/judgments/getty-images-and-others-v-stability-ai/>.

Getty Images gestisce uno fra i più estesi archivi di fotografie, immagini editoriali, illustrazioni e contenuti multimediali al mondo, le cui licenze d'uso sono concesse a editori, agenzie pubblicitarie, testate giornalistiche, emittenti televisive, piattaforme digitali e imprese di comunicazione in oltre duecento Paesi. Il valore economico del gruppo si fonda essenzialmente sull'archivio, che la stessa *High Court*, sulla scorta della citazione, ha definito quale «*lifeblood*» del *business* attoreo²⁵⁹: le opere ivi contenute sono per la maggior parte protette dal diritto d'autore e provengono da contratti di licenza stipulati con fotografi, artisti e autori di contenuti visuali. Nel corso degli anni, Getty ha investito risorse significative nella catalogazione, conservazione, *metadataazione* e tutela giuridica delle immagini, sì che l'uso massivo di tali opere per finalità di addestramento di sistemi generativi è stato dalle attrici percepito come minaccia non solo ai propri diritti patrimoniali, ma all'intero ecosistema della fotografia professionale.

Stability AI Ltd. è divenuta nota, in particolare, per lo sviluppo di *Stable Diffusion*, uno dei modelli di generazione di immagini *open source* oggi più diffusi. Il sistema si colloca nella categoria dei *diffusion models*, ossia di quei modelli di IA in grado di generare immagini realistiche o artistiche a partire da *text prompts* forniti dall'utente; esso viene addestrato attraverso l'analisi di ingenti quantitativi di immagini reperite *online* e associate a descrizioni testuali, così da apprendere relazioni statistiche fra parole, concetti visivi, forme, stili e composizioni grafiche.

Secondo la prospettazione attorea, il *training set* comprendeva oltre dodici milioni di immagini tratte dai siti Getty; in giudizio, tuttavia, Stability ha ammesso soltanto che «almeno alcune» immagini provenienti dai siti Getty sono state utilizzate nell'addestramento, e la Corte ha espressamente declinato di compiere

accertamenti sul numero delle opere impiegate²⁶⁰. È invece pacifico che tutte le versioni di *Stable Diffusion* sono state addestrate su sottoinsiemi del dataset LAION-5B, predisposto dall'associazione tedesca LAION e.V.²⁶¹: la medesima al centro del contenzioso tedesco *Kneschke v. LAION* (v. *infra*).

Sotto il profilo strettamente tecnico, *Stable Diffusion* si avvale di reti neurali profonde, apprende correlazioni statistiche fra immagini e testo, produce gli *output* mediante processi probabilistici a partire da un dato iniziale aleatorio, e non opera dunque alla stregua di un tradizionale archivio di immagini memorizzate. Il modello di *business* della convenuta si caratterizza per un approccio relativamente aperto e decentralizzato allo sviluppo dell'IA generativa, basato sulla messa a disposizione di versioni del modello a ricercatori, sviluppatori e utenti finali. Tale impostazione, se da un lato ha favorito la diffusione capillare della tecnologia, dall'altro ha acuito le problematiche connesse alla provenienza dei dati impiegati nel *training*.

III. L'origine della controversia e lo svolgimento del processo

Il procedimento è stato instaurato con atto introduttivo del 16 gennaio 2023 dinanzi alla *High Court of Justice of England and Wales* sulla base della contestata utilizzazione, da parte di Stability AI, di milioni di immagini appartenenti all'archivio Getty per l'addestramento dei modelli *Stable Diffusion*, senza autorizzazione né corresponsione di alcun compenso. L'atto di citazione si articolava in cinque distinte *claims*, che meritano di essere singolarmente illustrate. Primo, un *primary copyright infringement claim*, fondato sulle Sections 16, 17 e 20 del *Copyright, Designs and Patents Act 1988* (CDPA): secondo le attrici, il *dataset* utilizzato per il *training* avrebbe incluso un numero estremamente elevato di immagini protette dal diritto d'autore, prelevate dai siti dell'archivio mediante tecniche

²⁵⁹ Sent. 4 novembre 2025, §8.

²⁶⁰ Sent. 4 novembre 2025, § 49 e Lett. (K) *Remaining outstanding issues – Number of works used in training* §§747-752.

²⁶¹ Sent. 4 novembre 2025, § 46.

di *automated scraping*; tale condotta avrebbe integrato una violazione sistematica del diritto esclusivo di riproduzione spettante ai titolari.

Secondo, un *secondary copyright infringement claim*, fondato sulle Sections 22 e 23 del CDPA: nella prospettiva attorea, il modello *Stable Diffusion* costituirebbe esso stesso una «*infringing copy*» nell'accezione di cui alla Section 27 CDPA, non già perché contenente o archiviante copie delle opere, circostanza che le attrici non hanno mai allegato, bensì in forza dell'ipotesi controfattuale di cui alla Section 27(3)²⁶²: la creazione dei *model weights*, ove fosse avvenuta nel Regno Unito, avrebbe integrato violazione del copyright, sì che la sua importazione e distribuzione nel Regno Unito integrerebbe un atto di *secondary infringement* - nozione che, nel sistema inglese, presuppone una condotta dolosa o quanto meno consapevole del soggetto agente²⁶³.

Terzo, un *database right infringement claim*, articolato ai sensi della *Regulation 16 dei Copyright and Rights in Databases Regulations 1997*: Getty deduceva di aver investito risorse considerevoli nella costituzione del proprio archivio fotografico, sì che l'estrazione massiva delle immagini avrebbe integrato un'appropriazione illegittima di una porzione sostanziale del *database* tutelato dal *sui generis right* di matrice europea.

Quarto, un *trade mark infringement claim*, fondato sulla Section 10 del *Trade Marks Act 1994*, e segnatamente sui commi 1, 2 e 3 della medesima disposizione. Le attrici lamentavano che alcuni *output* generati da *Stable Diffusion* contenessero frammenti del *watermark* «Getty Images» - circostanza ritenuta non solo idonea ad ingenerare confusione nel pubblico e a determinare un danno reputazionale, ma anche

dimostrativa, secondo l'assunto attoreo, dell'avvenuta incorporazione delle immagini originali nel modello.

Infine, un *passing off claim*, istituto di derivazione giurisprudenziale tipico della *common law* inglese²⁶⁴, volto a tutelare il *goodwill* - ossia la reputazione commerciale e l'identità economica - del gruppo Getty.

La convenuta ha articolato la propria strategia difensiva intorno alla natura tecnica dei modelli generativi, sviluppando in particolare quattro argomentazioni convergenti: il *training* non equivarrebbe a una memorizzazione permanente delle immagini; il modello non conterrebbe copie accessibili delle opere; gli *output* sarebbero da considerare creazioni nuove e autonome; le eventuali somiglianze fra *output* generati e immagini originarie costituirebbero il risultato di correlazioni statistiche e non già di duplicazioni in senso tecnico.

Un ruolo di particolare rilievo, nell'economia del giudizio, è stato svolto dagli *expert witnesses* - figura che, nel processo civile inglese, è disciplinata dalla *Part 35 delle Civil Procedure Rules* e che, pur essendo nominata dalle parti, è tenuta a un dovere di lealtà nei confronti della Corte («*overriding duty to the court*»). Nella specie, ciascuna parte si è avvalsa di un esperto di primo piano: il prof. Hany Farid (University of California, Berkeley) per le attrici e il prof. Thomas Brox (Universität Freiburg) per la convenuta²⁶⁵.

I periti tecnici hanno analizzato in profondità il funzionamento del modello *Stable Diffusion* e il rapporto fra *dataset* di addestramento e immagini effettivamente generate, con particolare attenzione al fenomeno della *memorization*, ossia alla possibilità che il

²⁶² Ai sensi dell'art. 27(3) CDPA «[a]n article is also an infringing copy if— (a)it has been or is proposed to be imported into the United Kingdom, and (b)its making in the United Kingdom would have constituted an infringement of the copyright in the work in question, or a breach of an exclusive licence agreement relating to that work».

²⁶³ Sent. 4 novembre 2025, § 11.

²⁶⁴Per la moderna riformulazione dei presupposti del *passing off*, cfr. il celebre *dictum* di Lord Oliver in *Reckitt & Colman Products Ltd v. Borden Inc.* [1990] 1 WLR 491, il quale individua i tre elementi costitutivi della tort nel *goodwill*, nella *misrepresentation* e nel *damage* subito dal titolare.

²⁶⁵ Sent. 4 novembre 2025, § 91 ss. *The Technical Experts*.

modello «ricordi» immagini specifiche utilizzate nel *training*.

Getty Images ha tentato di dimostrare che, mediante l'inserimento di determinati *image prompts* e *text prompts*, il sistema fosse in grado di riprodurre *output* estremamente simili alle opere originarie. La valutazione tecnica condotta dagli esperti - segnatamente in sede di *expert evidence* - ha tuttavia portato a ritenere che gli *output* contestati dipendessero in misura determinante dagli *input* forniti dall'utilizzatore, che il modello non conservasse copie integrali delle opere e che non esistesse, nei *model weights*, un archivio interno di immagini suscettibile di recupero. Tale ricostruzione tecnica ha esercitato un'influenza determinante sulla *ratio decidendi* della pronuncia.

Nel corso del dibattimento il *thema decidendum* ha subito un significativo ridimensionamento, dovuto al progressivo abbandono, da parte attorea, di alcune delle pretese originariamente avanzate.

Furono effettivamente abbandonati, nel corso del giudizio, due dei cinque capi di domanda originari: il *primary copyright infringement* e il *database right infringement*. Quanto al primo, all'esito dell'istruttoria le stesse attrici hanno riconosciuto l'assenza di prova che l'addestramento del modello fosse avvenuto nel territorio del Regno Unito - essendo esso avvenuto, secondo la prospettazione della convenuta, su infrastruttura cloud (il cluster AWS) localizzata all'estero, circostanza sulla quale la Corte non ha peraltro compiuto alcun accertamento formale; essendo il *copyright* britannico territorialmente limitato, Getty si vide costretta a desistere dalla domanda²⁶⁶.

Quanto al secondo, il relativo abbandono è stato motivato dal legame intrinseco della domanda con il *Training and Development Claim* e con l'*Outputs Claim*. Va inoltre menzionata l'ulteriore componente del *primary infringement claim*, del pari abbandonata, relativa all'*output infringement*, ossia alla violazione diretta del diritto d'autore mediante la generazione di *output* sostanzialmente coincidenti con immagini Getty: tale domanda è stata abbandonata anche in considerazione del fatto che, nel corso del processo, Stability AI è riuscita a documentare di aver bloccato i *prompt* specificamente utilizzati per la generazione degli *output* contestati, sì che la tutela inibitoria appariva sostanzialmente conseguita *aliunde*. Non furono invece abbandonati il *secondary copyright infringement claim*, il *trade mark infringement claim* e il *passing off claim*, domande giunte a definizione con la sentenza di merito.

IV. La decisione del 4 novembre 2025

Sul *secondary copyright infringement claim* - divenuto il principale capo di domanda residuo in materia di *copyright* - la *High Court* è pervenuta a una statuizione di rigetto, della quale conviene ripercorrere l'iter argomentativo.

Le Sections 22 e 23 CDA configurano la violazione secondaria del diritto d'autore quale condotta di importazione o di *dealing* di un «*article*» che sia, o che l'agente sappia o abbia ragione di ritenere che sia, una «*infringing copy*» di un'opera protetta. Sotto il primo profilo, la Corte ha riconosciuto che la nozione di «*article*» può comprendere anche beni intangibili, e

²⁶⁶ La circostanza si può efficacemente sintetizzare nei termini secondo cui il punto debole del *copyright* tradizionale risiede nel principio di territorialità e, dunque, nella natura intrinsecamente transnazionale del *training* algoritmico: la mera localizzazione dei *server* all'estero può rendere assai problematica l'applicazione delle normative interne. Su territorialità e digitalizzazione, si rinvia a G. MAZZIOTTI, F. SIMONELLI, *Another breach in the wall: copyright territoriality in Europe and its progressive erosion on the grounds of competition law*, in *Info*, 18, 6, 2016, 55-66; T. DREIER, *Copyright in the times of the internet—overcoming the principle of territoriality within*

the EU, in *ERA Forum*, 18, 2017, 7-19; M. RATTZÉN, *Location Is All You Need: Copyright Extraterritoriality and Where to Train Your AI*, in *Columbia Sci. & Tech L. Rev.*, 26, 2024, 175 ss.; J. HOFFMAN, *Crossing Borders in the Digital Market: A Proposal to End Copyright Territoriality and Geo-Blocking in the European Union*, in *Geo. Wash. Int'l L. Rev.*, 49, 143, 2016-2017, 143 ss.; K. AOKI, *(Intellectual) Property and Sovereignty: Notes Toward a Cultural Geography of Authorship*, in A. GEORGE (ed.) *Globalization and Intellectual Property*, 1st ed., 2006, 469-531.

dunque un modello di IA; sotto il secondo profilo, tuttavia, essa ha negato che i *model weights* di *Stable Diffusion* costituiscano una «*infringing copy*» nell'accezione richiesta dal CDPA.

Il fulcro della motivazione poggia su una distinzione di carattere ontologico fra archiviazione dell'opera e apprendimento algoritmico. A parere della Justice Smith, i modelli *Stable Diffusion* non contengono copie delle immagini utilizzate per il *training*, non archiviano le opere in forma recuperabile e non realizzano una memorizzazione equivalente alla conservazione tradizionale di *file* digitali: il modello conserva esclusivamente informazioni statistiche e parametri matematici derivati dall'analisi dei dati di addestramento. Occorre peraltro precisare che tale ricostruzione fattuale era in larga misura pacifica: le attrici non avevano mai allegato che il modello archiviasse copie delle opere²⁶⁷ e sul punto la prova tecnica è rimasta sostanzialmente incontestata. Il passaggio decisivo della motivazione è dunque di ordine interpretativo, e consiste nell'affermazione per cui una «*infringing copy*» ai sensi delle Sections 22, 23 e 27 CDPA deve essere, o quantomeno essere stata, essa stessa una copia dell'opera: non basta che l'*article* sia il prodotto di un procedimento (il *training*) che, ove svolto nel Regno Unito, avrebbe comportato la realizzazione di copie contraffatte. Viene così respinta la lettura attorea della Section 27(3), imperniata sull'ipotesi controfattuale della fabbricazione nel Regno Unito.

Il profilo della territorialità resta comunque sullo sfondo del rigetto: è per la mancata prova di attività di addestramento nel Regno Unito che Getty aveva dovuto ripiegare sulla violazione secondaria, e l'argomento fondato sulla Section 27(3) era volto a superare il limite territoriale mediante l'ipotesi controfattuale; senonché, una volta stabilito che i *model weights* importati non sono, né sono mai stati, copie delle opere, il presupposto stesso della violazione secondaria è venuto meno. La pronuncia, peraltro, lascia

espressamente aperta la diversa questione - non oggetto del giudizio - concernente la possibilità che vi siano *output* sostanzialmente coincidenti con le opere protette, e dunque l'*output infringement* in senso proprio: questione che Getty, come accennato, aveva originariamente sollevato ma non coltivato.

Sul *trade mark infringement claim*, la Corte è pervenuta a una pronuncia di parziale fondatezza. La violazione del marchio è stata accertata limitatamente ad alcune versioni iniziali di *Stable Diffusion* (v1.x e v2.x), nei cui *output* risultavano talora visibili *watermark* Getty o segni a essi assimilabili, in violazione della Section 10, commi 1 e 2, del Trade Marks Act 1994. Più precisamente, la violazione ex Section 10(1) (doppia identità) è stata riscontrata per la sola v1.x e con riferimento al solo marchio ISTOCK, mentre la violazione ex Section 10(2) (rischio di confusione) ha riguardato *output* della v1.x (marchio ISTOCK) e della v2.1 (marchio GETTY IMAGES); nessuna violazione è stata accertata per le versioni successive (SD XL e v1.6), in difetto di prova della generazione reale di *watermark* nel Regno Unito da parte di utenti. Quanto alla v1.x, la Corte ha escluso la responsabilità diretta di Stability per la messa a disposizione del modello tramite le pagine GitHub e Hugging Face di CompVis, sicché la responsabilità è rimasta circoscritta ai canali di accesso direttamente riferibili alla convenuta (piattaforma API e DreamStudio). È stata invece rigettata la pretesa fondata sulla Section 10, comma 3, della medesima legge, relativa allo sfruttamento indebito del carattere distintivo o della rinomanza del marchio²⁶⁸.

Quanto al *passing off claim*, la *High Court* ha ritenuto opportuno non pronunciarsi separatamente, considerando che la decisione sul *trade mark* assorbiva sostanzialmente le questioni rilevanti.

All'udienza dei «*consequential*» del 16-17 dicembre 2025, la Corte è tornata sulle questioni accessorie alla sentenza di merito.

²⁶⁷ Sent. 4 novembre 2025, § 11.

²⁶⁸ Sent. 4 novembre 2025, lett. (F) *The Trade Mark Infringement Claim*, §§ 134 ss.

Quanto alle spese, pur riconoscendo che Getty fosse «*the overall winner*» sul piano del marchio, la Justice ha disposto una riduzione significativa per riflettere l'insuccesso attoreo sulle restanti pretese, condannando Getty al pagamento del 69,4 per cento delle spese di Stability su base *standard*, con un pagamento provvisorio (*interim payment*) di £ 4.395.656,89.

Soprattutto, in quella sede è stato concesso a Getty integralmente (per tutti i motivi proposti) il *permission to appeal* sul rigetto del *secondary copyright infringement claim* - provvedimento dotato di propria autonoma *neutral citation*: [2025] EWHC 3343 (Ch). Nel motivare la concessione, il giudice ha osservato che la questione concernente l'interpretazione della nozione di «*infringing copy*» di cui alle Sections 22 e 23 CDPa, applicata ai modelli di IA, non era stata in precedenza esaminata da alcun giudice del Regno Unito e che essa avrebbe potuto avere implicazioni di vasta portata non solo per i modelli generativi, ma anche, più in generale, per gli «*articles*» di natura intangibile, ivi compreso il *software*. È stato invece negato a Stability il *permission to appeal* sul *trade mark infringement*.

V. Profili sistematici e prospettive

La decisione *Getty v. Stability AI* costituisce, allo stato, uno dei più significativi momenti dell'evoluzione contemporanea del diritto d'autore, sebbene sia stata qualificata dalla stessa Mrs Justice Smith come «*both historic and extremely limited in scope*»²⁶⁹.

La *High Court* ha mostrato una notevole attenzione agli aspetti tecnici dei modelli generativi e ha accolto l'assunto secondo cui un modello di IA non costituisce un archivio di immagini memorizzate, e che il *training* produce rappresentazioni matematiche astratte piuttosto che copie tradizionali delle opere originarie.

Tale conclusione, peraltro, non elide le profonde criticità giuridiche ed economiche connesse all'IA generativa. Rimane aperta, in primo luogo, la questione fondamentale della qualificazione giuridica dell'atto di *training* e della possibilità che esso integri una forma autonoma di sfruttamento economico delle opere protette: questione che, nel caso di specie, è stata sottratta alla cognizione della Corte per effetto dell'abbandono attoreo del *primary copyright infringement claim*, e che dovrà essere affrontata in altre sedi, a partire dal parallelo giudizio pendente tra le medesime parti dinanzi alle corti federali statunitensi²⁷⁰. Parallelamente, sembra destinato a divenire centrale il problema della c.d. «*market substitution*»: il rischio, cioè, che i sistemi generativi riducano il valore commerciale delle opere creative, alterando profondamente mercati come quello delle immagini *stock* e dei contenuti digitali professionali.

Ulteriore punto di rilievo concerne la riproduzione temporanea delle opere durante il *training*. Nel diritto europeo e in quello britannico, anche copie temporanee possono integrare un atto di riproduzione rilevante ai fini della tutela autoriale, a meno che non ricadano nelle finalità considerate lecite dall'elenco di eccezioni e limitazioni al diritto d'autore (art. 5, par. 1, della direttiva 2001/29/CE, nonché per il Regno Unito post-Brexit, la corrispondente Sec. 28A CDPa). La Corte riconosce che il processo di addestramento implichi necessariamente il *download* e la memorizzazione temporanea delle immagini comprese nei *dataset*, ma distingue tale attività dal risultato finale del *training*: secondo *Mrs Justice Smith*, *Stable Diffusion* non conserva copie recuperabili delle immagini originarie, bensì parametri matematici e correlazioni statistiche derivate dall'analisi dei dati. La sentenza, in altri termini, sembra escludere che il modello matematico risultante

²⁶⁹ Sent. 4 novembre 2025, § 757.

²⁷⁰ Si segnala, sotto questo profilo, il parallelo procedimento instaurato da Getty Images negli Stati Uniti dinanzi alla *United States District Court for the Northern District of California* (No. 3:25-cv-06891, dinanzi al giudice

Trina L. Thompson), in cui le questioni relative al *training* sono coltivate sulla base del *Copyright Act* statunitense e della *doctrine of fair use*, con *jury trial* attualmente programmato per gennaio-febbraio 2028.

dal *training* possa essere qualificato quale «copy» delle opere utilizzate per l'addestramento²⁷¹; non risolve, però, in via definitiva, né il problema della liceità dell'acquisizione iniziale dei dati, né quello della riproduzione temporanea effettuata durante il *training*, lasciando aperte alcune fra le questioni più delicate dell'intero dibattito.

VI. Il provvedimento in commento.

JUDGEMENT OF NOVEMBER 4, 2025.

[...]

(L) CONCLUSION

757. In summary, although Getty Images succeed (in part) in their Trade Mark Infringement Claim, my findings are both historic and extremely limited in scope. The Secondary Infringement Claim fails.

758. In a little more detail, my findings on the key issues are as follows:

i) Stability bears no direct liability for any tortious acts alleged in these proceedings arising by reason of the release of v1.x Models via the CompVis GitHub and CompVis Hugging Face pages.

ii) the question of trade mark infringement arises only:

a) in respect of the generation of Getty Images watermarks* and iStock watermarks* by v1.x Models (in so far as they were accessed via DreamStudio and/or the Developer Platform);

b) in respect of the generation of Getty Images watermarks* by v2.x Models.

iii) There is no evidence of a single user in the UK generating either Getty Images or iStock watermarks* using SD XL and v1.6 Models. Thus no question of trade mark infringement arises in respect of these Models and that claim, in so far as it relates to them, is dismissed.

iv) As to Getty Images' claim under section 10(1) TMA:

a) Getty Images succeed in respect of iStock watermarks* generated by users of v1.x (in so far as the Models were accessed via DreamStudio and/or the Developer Platform) in that infringement of the ISTOCK Marks pursuant to section 10(1) TMA has been established. This success is however based specifically on the example watermarks*

shown on the Dreaming Image and the Spaceships Image – the latter having been generated by Model v1.2. Given the way in which the case has been advanced, it is impossible to know how many (or even on what scale) watermarks* have been generated in real life that would fall into a similar category.

b) Getty Images fail in respect of Getty Images watermarks*, there being no evidence of infringement of the Getty Images Marks under section 10(1) TMA. That claim is dismissed.

v) As to Getty Images' claim under section 10(2) TMA:

a) Getty Images succeed in respect of iStock watermarks* generated by users of v1.x (in so far as the Models were accessed via DreamStudio and/or the Developer Platform) in that infringement of the ISTOCK Marks pursuant to section 10(2) TMA has been established. This success is based specifically on the example watermarks* shown on the Dreaming Image and the Spaceships Image - the latter having been generated by Model v1.2.

b) Getty Images succeed in respect of Getty Images watermarks* generated by users of v2.x in that infringement of the Getty Images Marks pursuant to section 10(2) TMA has been established. This success is based specifically on the example watermark* on the First Japanese Temple Garden Image, generated by Model v2.1.

Again, it is impossible to know how many (or even on what scale) watermarks* have been generated in real life that would fall into a similar category.

vi) Getty Images' claim under section 10(3) TMA is dismissed.

vii) For reasons I have explained, I have declined to address Getty Images' allegation of passing off.

viii) Getty Images' claim of secondary infringement of copyright is dismissed. Although an "article" may be an intangible object for the purposes of the CDPA, an AI model such as Stable Diffusion which does not store or reproduce any Copyright Works (and has never done so) is not an "infringing copy" such that there is no infringement under sections 22 and 23 CDPA.

ix) As to other remaining issues:

a) On copyright subsistence/ownership, Getty Images have failed to make out title to the copyright in SOCI Works A3 and A4, but have established their title to the copyright in SOCI Works A9, A10 and A11.

²⁷¹ Per un raffronto con il diverso approccio adottato in altri ordinamenti, si veda la causa parallelamente pendente dinanzi alla *United States District Court for the Northern District of California, Andersen et al. v. Stability AI Ltd. et al.*, No. 3:23-cv-00201-WHO, nella quale la questione della qualificazione giuridica del *training* è stata posta sotto la lente del *fair use* statunitense, con esiti istruttori e argomentativi che si distinguono per più aspetti da quelli inglesi. In aggiunta, Sul punto, e per il confronto con l'esito opposto raggiunto dal *Landgericht München I* nel caso *GEMA v. OpenAI*, v. P. Mezei, *Memorization and Generative AI – A Persistent Issue with*

Copyright Consequences? (2025), il quale rileva come, sul terreno della *memorization*, la pronuncia inglese e quella tedesca si collochino ai poli opposti: la prima esclude in fatto la presenza di *memorization* nel modello (§ 559), la seconda la accerta e la qualifica come riproduzione, secondo una divergenza che l'A. non ritiene necessariamente frontale, potendo dipendere dalla diversità tecnica dei modelli in giudizio. Sulla ricostruzione della divaricazione v. *amplius infra*, *La memorization come riproduzione e i limiti dell'eccezione per text and data mining – Gema v. OpenAI*.

b) On the Licensing Issue:

i) Sample Licences #2, 3, 10, 11, 13, 30 and 32 are not exclusive licences under section 92 CDPA;

ii) Sample Licences #17, #19, #34-38 are exclusive licences under section 92 CDPA.

c) I make no finding as to the number of Visual Assets or Copyright Works used in training Stable Diffusion

d) This is not a case in which I consider an award of additional damages to be justified.

759. Any consequential matters arising in light of this judgment will be dealt with at a hearing in due course.

[...]

**TEXT AND DATA MINING, RICERCA SCIENTIFICA E RISERVA
MACHINE-READABLE**

**ROBERT KNESCHKE v. LAION E.V. - LG HAMBURG, 27
SETTEMBRE 2024 (310 O 227/23) - OLG HAMBURG,
10 DICEMBRE 2025 (5 U 104/24)**

di Sara Manni e Leonardo Ortenzi*

I. Premesse

La controversia *Robert Kneschke v. LAION e.V.*, decisa in primo grado dal *Landgericht Hamburg* con sentenza del 27 settembre 2024 (n. 310 O 227/23) e in appello dal *Hanseatisches Oberlandesgericht Hamburg* con sentenza del 10 dicembre 2025 (n. 5 U 104/24), si configura quale primo significativo intervento giurisprudenziale tedesco - ed europeo - sulle eccezioni di *text and data mining* introdotte dalla direttiva (UE) 2019/790 sul *Digital Single Market*. Il rilievo sistematico della vicenda è anzitutto cronologico: nessun'altra giurisdizione dell'Unione era pervenuta, prima della pronuncia di Amburgo, a un compiuto vaglio giudiziale del rapporto fra attività preparatorie del *training* algoritmico e disciplina del diritto d'autore. La pronuncia di primo grado e quella di appello, lette congiuntamente, offrono dunque la prima sistemazione interpretativa delle disposizioni di cui agli artt. 3 e 4 della direttiva - recepiti nell'ordinamento tedesco rispettivamente dai §§ 60d e 44b dell'*Urheberrechtsgesetz* (UrhG) - applicate al peculiare contesto della creazione di *dataset* destinati all'addestramento di modelli di intelligenza artificiale.

Entrambi i gradi di giudizio si sono conclusi con il rigetto delle pretese attoree, e la pronuncia di appello - pur confermando l'esito del primo grado - ha esplicitamente esteso la qualificazione della condotta della convenuta sotto entrambe le eccezioni di *text and data mining*: quella per finalità scientifiche di cui all'art. 3 della direttiva e quella generale di cui all'art. 4.

È stato altresì ammesso, a fronte della rilevanza sistematica della questione, ulteriore ricorso al *Bundesgerichtshof*; la Revision è stata proposta

nel dicembre 2025 e risulta pendente dinanzi al *Bundesgerichtshof*, con l'udienza fissata al 3 settembre 2026, sicché il contenzioso non può ancora considerarsi definitivamente chiuso.

II. Le parti del giudizio

Parte attrice è Robert Kneschke, fotografo professionista tedesco, attivo segnatamente nel settore della fotografia *stock*. Le sue opere risultavano, all'epoca dei fatti, disponibili anche sul sito *Bigstock* - piattaforma *stock* appartenente al gruppo *Shutterstock* - secondo termini d'uso che, in lingua naturale inglese, vietavano espressamente l'impiego di programmi automatizzati per scopi di *downloading*, *indexing*, *scraping* o *caching* dei contenuti del sito.

Parte convenuta è LAION e.V. (*Large-scale Artificial Intelligence Open Network*), associazione di diritto tedesco senza fini di lucro avente sede in Amburgo e operante nel campo della ricerca sull'intelligenza artificiale. LAION è nota, fra l'altro, per la creazione e la libera diffusione di *dataset* di immagini destinati all'addestramento di modelli generativi: in particolare, il *dataset* LAION-5B, contenente all'incirca sei miliardi di coppie immagine-testo, è stato impiegato - fra l'altro - quale base di addestramento per *Stable Diffusion*. Sotto il profilo della struttura, occorre precisare che il *dataset* non contiene direttamente le immagini, ma una tabella di *URL* che rinviano alle stesse, accompagnata dalla descrizione testuale corrispondente; sicché la riproduzione delle immagini risulta temporanea e funzionale alla sola verifica della coerenza fra contenuto visivo e descrizione associata.

III. L'origine della controversia

Nella seconda metà del 2021, LAION ha proceduto, mediante un programma automatizzato, al *download* di una vasta quantità di immagini disponibili *online*, fra le quali una fotografia del Sig. Kneschke caricata sul sito *Bigstock*, nella versione di anteprima

* Premesso che l'opera è frutto del lavoro congiunto dei due autori, a Leonardo Ortenzi vanno attribuiti i Paragrafi I, II e III e a Sara Manni i Paragrafi IV, V e VI.

munita del *watermark* dell'agenzia. Le immagini venivano quindi confrontate con la descrizione testuale ad esse associata nel dataset *Common Crawl*, al fine di verificarne la corrispondenza; le immagini risultate coerenti venivano mantenute nel dataset finale soltanto sotto forma di URL e descrizione, mentre le copie temporanee in formato binario venivano cancellate.

Il fotografo, avuta conoscenza dell'inclusione della propria opera nel *dataset*, ha avanzato richiesta stragiudiziale di rimozione e successivamente, nell'aprile 2023, ha promosso azione inhibitoria dinanzi al *Landgericht Hamburg*, deducendo la violazione del proprio diritto esclusivo di riproduzione di cui al § 16 UrhG (corrispondente all'art. 2 della direttiva 2001/29/CE, InfoSoc).

Giova peraltro premettere che la fotografia *de qua* è stata qualificata in giudizio quale fotografia semplice, e non come opera fotografica²⁷².

L'attore sosteneva che la riproduzione della propria fotografia - pur temporanea - non potesse beneficiare di alcuna delle eccezioni di *text and data mining* previste dall'ordinamento tedesco in attuazione della direttiva DSM, in quanto: (i) l'opera era accessibile mediante *Bigstock* soltanto a condizioni che vietavano espressamente lo *scraping*; (ii) la convenuta non avrebbe svolto alcuna attività propriamente «scientifica», bensì la mera predisposizione di un *dataset* funzionale all'uso commerciale di soggetti terzi.

La convenuta, per converso, ha invocato l'operatività delle eccezioni di *text and data mining* introdotte dai §§ 44b e 60d UrhG, sostenendo: (i) sotto il profilo soggettivo, di qualificarsi quale organizzazione di ricerca senza fini di lucro ai sensi dell'art. 2, n. 1, della direttiva

DSM, e dunque di poter accedere all'eccezione qualificata di cui all'art. 3²⁷³; (ii) sotto il profilo oggettivo, di aver utilizzato l'opera unicamente nella misura necessaria all'analisi automatizzata della corrispondenza fra immagine e descrizione, ossia per la generazione di informazioni concernenti «*patterns, trends and correlations*»; (iii) infine, sotto il profilo dell'*opt-out* di cui all'art. 4, par. 3, della direttiva, che la riserva contenuta nei termini d'uso del sito *Bigstock* non sarebbe stata espressa in formato *machine-readable*, come richiesto dalla disposizione.

IV. La decisione del Landgericht Hamburg del 27 settembre 2024

Il *Landgericht Hamburg*, con sentenza del 27 settembre 2024, ha rigettato la domanda attorea. Il giudice di prime cure ha accertato preliminarmente che la condotta della convenuta integrava una riproduzione ai sensi del § 16 UrhG, ma che essa risultava coperta dall'eccezione di *text and data mining* per finalità di ricerca scientifica di cui al § 60d UrhG - attuativa dell'art. 3 della direttiva DSM. La motivazione poggia su tre passaggi essenziali.

In primo luogo, il giudice ha qualificato l'attività della convenuta come *text and data mining* ai sensi della definizione di cui all'art. 2, n. 2, della direttiva e del § 44b, comma 1, UrhG, intesa come «analisi automatizzata di una o più opere digitali o digitalizzate al fine di ottenere informazioni, in particolare su *patterns, trends* e correlazioni». Nel caso di specie, la finalità specifica dell'analisi era costituita dalla verifica della corrispondenza fra l'immagine e la sua descrizione testuale - finalità che il giudice ha ritenuto pienamente suscumbibile nella nozione richiamata.

²⁷² Nel diritto interno il *discrimen* è tracciato dall'art. 92 l.d.a., che originariamente prevedeva la tutela delle fotografie semplici per un periodo di venti anni dalla produzione, ora estesi a settanta grazie all'intervento dell'art. 47 co. 1 L. 2 dicembre 2025 n. 182. Invece, le opere fotografiche godono della protezione standard prevista per le opere dell'ingegno di carattere creativo, trovando tutela *sub* art. 2, n. 7, l.d.a. Sulle nozioni di originalità e carattere creativo che comportano la ricaduta di una fotografia nell'una o nell'altra categoria, si rinvia a

G. MUNNO, *Diritto d'autore e fotografia: il confine giuridico tra scatto tecnico e espressione creativa*, in *Dir. Industriale*, 5, 2025, 436 ss., nota a Cass. 20 dicembre 2024, n. 33599.

²⁷³ A tal proposito, J. P. QUINTAIS, *What Is a 'Research Organisation' and Why it Matters: From Text and Data Mining to AI Research*, in *GRUR International*, 74(5), 2025, 397-398.

In secondo luogo, il *Landgericht* ha riconosciuto in capo a LAION lo statuto di organizzazione di ricerca senza fini di lucro nell'accezione dell'art. 2, n. 1, della direttiva. Sul punto, il giudice ha respinto l'obiezione attorea secondo cui la mera predisposizione di un *dataset* non costituirebbe «ricerca scientifica», in quanto non determinerebbe, di per sé, alcun «guadagno di conoscenza scientifica»: la creazione di un *dataset* - ha osservato - costituisce attività propriamente preparatoria della ricerca, e in tale qualità rientra nel perimetro dell'eccezione. In terzo luogo, il giudice ha ritenuto irrilevante, ai fini dell'applicabilità del § 60d UrhG, la circostanza che il *dataset* così costituito possa essere successivamente utilizzato da soggetti terzi per finalità commerciali: l'eccezione qualificata di cui all'art. 3 della direttiva, infatti, non subordina la propria operatività alla destinazione finale del *dataset*, ma alla qualificazione soggettiva del beneficiario e alla finalità immediata della riproduzione. Avendo affermato l'operatività dell'eccezione qualificata di cui al § 60d UrhG, il giudice di prime cure ha potuto formalmente esimersi dall'esaminare l'ulteriore questione concernente la sussistenza, nel caso di specie,

dei presupposti dell'eccezione generale di *text and data mining* di cui al § 44b UrhG (corrispondente all'art. 4 della direttiva DSM), e in particolare il problema della validità della riserva di *opt-out* formulata da *Bigstock* nei termini d'uso del proprio sito. Ciononostante, sia pure in *obiter dictum*, il *Landgericht* ha ritenuto opportuno svolgere alcune considerazioni preliminari sulla nozione di riserva «*machine-readable*» di cui all'art. 4, par. 3, della direttiva.

Sotto questo profilo, il giudice ha rilevato che la formulazione della riserva in «linguaggio naturale» - ossia in inglese corrente, senza l'ausilio di *tag* o *metadata* propriamente tecnici - non costituisce, di per sé, ostacolo alla configurazione della stessa quale *machine-readable*, ove si consideri che la tecnologia contemporanea consente l'interpretazione automatizzata anche di clausole espresse in linguaggio naturale²⁷⁴. La conclusione, ancorché formulata *obiter*, avrebbe aperto la possibilità di un'interpretazione estensiva del requisito della *machine-readability*, potenzialmente comprensiva di qualunque *statement* contrattuale formulato in linguaggio naturale e accessibile mediante *crawler*²⁷⁵.

²⁷⁴ La decisione si trova criticamente commentata in Y. ZHANG, *Kneschke v LAION: Are Text and Data Mining Exceptions a "Get-Out-of-Jail-Free Card" for AI Training?*, in *European Intellectual Property Review*, 47, 527-535, per il quale la scelta del giudice favorirebbe ingiustamente i giganti tecnologici a discapito dei titolari di diritti rilevanti e delle PMI, pur riconoscendo il valore di quanto stabilito con riferimento al linguaggio naturale in *obiter*, circostanza che non vale, naturalmente, a colmare l'incertezza normativa circa le modalità di formulazione dell'*opt-out*.

²⁷⁵ Si noti peraltro che, sul punto, il *Code of Practice for General-Purpose AI Models* - e segnatamente il relativo *Copyright Chapter*, pubblicato a livello unionale nel luglio 2025 quale strumento (di adesione volontaria) volto ad agevolare la dimostrazione dell'adempimento agli obblighi di cui all'art. 53, par. 1, lett. c, del regolamento (UE) 2024/1689 - offre alcune indicazioni di rilievo sulla natura della riserva di *opt-out* di cui all'art. 4, par. 3, della direttiva (UE) 2019/790. Il testo muove dal presupposto che la riserva debba essere *machine-readable* e impegna i firmatari, ove impieghino *web-crawler* per la raccolta di dati a fini di *text and data mining* e di addestramento dei modelli, a identificarla e a rispettarla mediante tecnologie

allo stato dell'arte. A tal fine individua due canali concorrenti. Per un verso (lett. a), l'impegno a impiegare *crawler* idonei a leggere e a seguire le istruzioni espresse secondo il *Robot Exclusion Protocol* (il file *robots.txt*), nella formulazione di cui all'*IETF Request for Comments* n. 9309 e nelle versioni successive di cui l'IETF dimostri la fattibilità tecnica. Per altro verso (lett. b), l'impegno a riconoscere e a rispettare ulteriori protocolli *machine-readable* - ad esempio fondati su metadati *asset-based* o *location-based* - purché adottati da organismi di normazione internazionali o europei, ovvero allo stato dell'arte, tecnicamente implementabili e ampiamente diffusi fra i titolari (tenuto conto dei diversi settori culturali) e condivisi all'esito di un processo inclusivo, facilitato a livello UE, fra titolari, fornitori di IA e altri *stakeholder*; il tutto qualificato espressamente quale «soluzione più immediata», in attesa dello sviluppo di standard propriamente detti. È peraltro significativo che il medesimo testo si premuri di precisare (al n. 2) che l'impegno così assunto non incide sul diritto dei titolari di riservare l'uso delle opere «*in any appropriate manner*» - ossia con strumenti *machine-readable* per i contenuti resi pubblicamente disponibili *online*, ovvero «*by other means*» -, formula che riecheggia fedelmente la lettera dell'art. 4, par. 3, della direttiva e conferma come

La rilevanza pratica dell'*obiter* non va peraltro sottovalutata: ritenuta verosimilmente efficace la riserva dichiarata da Bigstock, ove il giudizio si fosse fondato sul solo § 44b UrhG l'eccezione generale sarebbe rimasta paralizzata dall'*opt-out*, con esito potenzialmente opposto. È dunque soltanto l'accertata operatività dell'eccezione qualificata di cui al § 60d UrhG ad aver reso la questione non necessaria ai fini della decisione.

V. La sentenza di appello del 10 dicembre 2025

Avverso la sentenza di primo grado, il Sig. Kneschke ha proposto appello dinanzi all'*Hanseatisches Oberlandesgericht Hamburg*, il quale, con sentenza del 10 dicembre 2025 (n. 5 U 104/24), ha rigettato l'impugnazione, confermando la statuizione di primo grado quanto al risultato, ma articolando in misura più ampia le ragioni della decisione.

In primo luogo, l'*Oberlandesgericht* ha ritenuto opportuno pronunciarsi anche sull'operatività

dell'eccezione generale di *text and data mining* di cui al § 44b UrhG (attuativo dell'art. 4 della direttiva DSM), affermandone la sussistenza nel caso di specie, sicché la condotta della LAION è stata considerata legittima sia in quanto realizzata da un'organizzazione di ricerca per finalità scientifiche, sia - indipendentemente - quale *text and data mining* generale.

Nel qualificare il confronto immagine-descrizione come analisi automatizzata diretta all'ottenimento di informazioni, segnatamente su correlazioni, ai sensi del § 44b, co. 1, UrhG, la Corte ha respinto la tesi dell'inapplicabilità delle eccezioni di *text and data mining* al contesto dell'IA per difetto di previsione da parte del legislatore storico, valorizzando tra l'altro il considerando 3 della direttiva DSM²⁷⁶ e l'espresso richiamo all'art. 53, par. 1, lett. c) del Reg. (UE) 2024/1689 (AI Act). La corte ha inoltre verificato positivamente, per entrambe le eccezioni, il *three-step test*: la riproduzione costituisce un procedimento puramente interno alla convenuta, il dataset divulgato contiene

l'esigenza della leggibilità automatica concerna propriamente i soli contenuti diffusi *online*, restando per il resto impregiudicata la libertà di forma della riserva. Sul piano operativo, dunque, il *Code* privilegia riserve strutturate e fondate su protocolli (*robots.txt*, metadati standardizzati), in ciò collocandosi in posizione non perfettamente coincidente con la lettura ampia prospettata in *obiter* dal *Landericht Hamburg* nel caso *Kneschke*, favorevole a riconoscere carattere *machine-readable* finanche alle riserve formulate in linguaggio naturale: pur senza escluderle in radice, esso ne subordina di fatto il rilievo alla verifica di implementabilità tecnica e di effettiva diffusione presso i titolari.

²⁷⁶ Il Considerando n. 3 della direttiva (UE) 2019/790 del 17 aprile 2019 sul diritto d'autore e sui diritti connessi nel mercato unico digitale le conferisce vocazione *pro futuro*, disponendo che «[i] rapidi sviluppi tecnologici continuano a trasformare il modo in cui le opere e altri materiali sono creati, prodotti, distribuiti e sfruttati, mentre emergono costantemente nuovi modelli di business e nuovi attori. È necessario che la legislazione in materia sia adeguata alle esigenze future, in modo da non limitare l'evoluzione tecnologica. Gli obiettivi e i principi stabiliti dal quadro giuridico dell'Unione sul diritto d'autore rimangono validi. Tuttavia, vi è ancora incertezza giuridica quanto a taluni utilizzi, anche transfrontalieri, delle opere e altri materiali in ambiente digitale, sia per i titolari dei diritti che per gli utilizzatori. In alcuni settori, come indicato nella

comunicazione della Commissione del 9 dicembre 2015, dal titolo «Verso un quadro normativo moderno e più europeo sul diritto d'autore», è necessario adeguare e completare l'attuale quadro dell'Unione sul diritto d'autore salvaguardando un elevato livello di protezione del diritto d'autore e dei diritti connessi. La presente direttiva prevede norme miranti ad adeguare talune eccezioni e limitazioni al diritto d'autore e diritti connessi all'ambiente digitale e al contesto transfrontaliero, nonché misure volte a facilitare determinate procedure di concessione delle licenze, in particolare, ma non solo, per la divulgazione di opere fuori commercio e di altri materiali e la disponibilità online di opere audiovisive su piattaforme di video su richiesta, al fine di garantire un più ampio accesso ai contenuti. Essa contiene anche disposizioni volte a rendere più agevole l'utilizzo dei contenuti di pubblico dominio. Per garantire il buon funzionamento e l'equità del mercato per il diritto d'autore sono altresì opportune norme relative ai diritti sulle pubblicazioni, all'uso di opere o altri materiali da parte dei prestatori di servizi online che memorizzano contenuti caricati dagli utenti e vi danno accesso, alla trasparenza dei contratti per autori e artisti (interpreti o esecutori), alla remunerazione di autori e artisti (interpreti o esecutori), nonché a un meccanismo per la revoca dei diritti che autori e artisti (interpreti o esecutori) hanno trasferito in esclusiva».

soltanto collegamenti ipertestuali alla fonte lecita e gli effetti concorrenziali di eventuali output generativi futuri costituiscono conseguenze astratte, non opponibili all'atto di riproduzione preparatorio²⁷⁷.

In secondo luogo, l'*Oberlandesgericht* ha confermato la riconducibilità della condotta della convenuta al perimetro applicativo dell'eccezione di *text and data mining* per finalità scientifiche di cui al § 60d UrhG, ribadendo che la creazione di un *dataset* destinato al *training* di modelli di IA costituisce, in sé e per sé, attività di ricerca scientifica. La Corte ha specificato che si tratta di un procedimento metodico e verificabile finalizzato all'acquisizione di conoscenza, qualificabile come ricerca applicata (*angewandte Forschung*, nella pronuncia la creazione del dataset è descritta quale procedimento metodico orientato ad un successivo avanzamento conoscitivo e verificabile).

Anche la successiva utilizzazione del dataset per l'addestramento di modelli di IA – ivi compresa l'attività di sviluppo condotta da imprese commerciali, che resta ricerca ancorché non privilegiata ex § 60d UrhG – è stata considerata qualificabile come ricerca scientifica applicata. Il *thema decidendum* è rimasto però circoscritto alla sola riproduzione preparatoria, ossia al *download* funzionale al confronto fra immagine e descrizione: la diversa questione circa il *training* (ovverosia se esso stesso integri o meno *text and data mining* ai sensi del § 44b, co. 1, UrhG), è stata espressamente lasciata aperta dalla Corte, con richiamo proprio alla pronuncia del Landgericht München I dell'11 novembre 2025 nel caso GEMA v. OpenAI²⁷⁸. La distinzione

fra fase di dataset creation e fase di model training, dunque, viene mantenuta dal giudice d'appello. Sotto il profilo soggettivo, infine, LAION è stata ricondotta alla categoria delle *sonstige Einrichtungen* dedite alla ricerca scientifica di cui al § 60d, co. 2, UrhG, con esclusione tanto del perseguimento di finalità commerciali quanto della causa ostativa di cui al § 60d, co. 2, UrhG (influenza determinante di un'impresa privata munita di accesso privilegiato ai risultati), nonostante i rapporti con Stability AI e un episodio di accesso privilegiato al dataset concesso a un'impresa terza a fronte del pagamento di una somma di denaro.

In terzo luogo, e con riferimento al requisito della *machine-readability* della riserva, l'*Oberlandesgericht* ha anzitutto riconosciuto che il fotografo può giovare della riserva formulata dall'agenzia (pur trattandosi di fotografia semplice) in ragione del rapporto di distribuzione tra le parti, e che la clausola in via interpretativa copre sul piano contenutistico anche il download automatizzato funzionale al confronto immagine-testo.

La riserva è risultata tuttavia priva della forma *machine-readable* richiesta dal §44b, co. 3, UrhG. Secondo la Corte, infatti, non è sufficiente che il testo *possa essere rilevato* da una macchina, occorrendo piuttosto che esso possa essere *automaticamente* interpretato, in modo che il processo automatizzato, tipico del *text and data mining*, possa conformarsi senza intervento umano.

Il requisito di "*machine readability*" è formulato in termini tecnologicamente neutri, giacché il legislatore tedesco (o, a monte, quello

²⁷⁷ Ai sensi dell'art. 9, par. 2, Convenzione di Berna per la protezione delle opere letterarie e artistiche del 9 settembre 1886, «è riservata alle legislazioni dei Paesi dell'Unione la facoltà di permettere la riproduzione delle predette opere in [1] *taluni casi speciali*, [2] *purché una tale riproduzione non rechi danno allo sfruttamento normale dell'opera* e [3] *non causi un pregiudizio ingiustificato ai legittimi interessi dell'autore*» (corsivo aggiunto ad evidenziare i tre "step"). Il test è stato poi ripreso nell'art. 5, par. 5 della Direttiva InfoSoc. Nonostante il superamento del test nella decisione in commento, L. MANSANI, *Text and data mining e*

intelligenza artificiale generativa, in *Dir. Industriale*, 2, 2026, 163 ss., suggerisce che la circostanza per cui «sia sorto e si stia sviluppando un mercato delle licenze per finalità di addestramento, sia individuali sia collettive, confermerebbe che la libertà di fare uso per quelle finalità di materiali o dati protetti dal diritto d'autore arrecherebbe un pregiudizio economico ai loro titolari», pericolo che invece la Corte, nel caso di specie, ha ritenuto meramente astratto.

²⁷⁸ Nella pronuncia, vgl. *zu Liedtexten LG München I GRUR-RS 2025, 30204*.

eurounitario) non prescrive alcuna forma determinata per la medesima, con la conseguenza per cui la riserva può in astratto figurare anche nelle condizioni generali di contratto.

L'onere di allegare e provare la *machine-readability* della riserva al momento rilevante grava sul titolare dei diritti, e il fotografo non lo avrebbe assolto, avendo dedotto elementi riferiti a capacità tecnologiche successive all'epoca dei fatti. Resta impregiudicata la questione se una riserva in linguaggio naturale soddisfi oggi il requisito. Quanto al reg. (UE) 2024/1689, l'art. 53, par. 1, lett. c, che impone ai *providers* di modelli di IA per finalità generali una *policy* di rispetto delle riserve di *opt out*, è valorizzato a sostegno dell'applicabilità della disciplina del TDM, non offrendo spunti circa l'adeguatezza delle riserve espresse in linguaggio naturale.

La Corte d'appello ha infine concesso al ricorrente il diritto di proporre ulteriore impugnazione («*Revision*») dinanzi al *Bundesgerichtshof*, in ragione della rilevanza fondamentale delle questioni di diritto trattate; l'impugnazione è stata effettivamente proposta nel dicembre 2025. L'esito di tale terzo grado di giudizio è suscettibile, evidentemente, di incidere in misura significativa sulla portata applicativa delle eccezioni di *text and data mining* non solo in Germania, ma in tutta l'Unione, attesa la possibilità che il *Bundesgerichtshof* sollevi questione pregiudiziale ai sensi dell'art. 267 TFUE.

VI. Profili sistematici e prospettive

Un primo profilo d'interesse concerne la stessa identificazione dell'oggetto della controversia. Il giudizio *Kneschke* non riguarda propriamente l'addestramento dei modelli generativi, bensì attività logicamente preliminari ad esso: la creazione di un *dataset* destinato all'addestramento. Nel quadro del § 60d UrhG, la qualificazione di «ricerca scientifica» applicata è stata riferita anche alla successiva utilizzazione del dataset per l'addestramento, ma l'oggetto della decisione è rimasto confinato alla fase preparatoria.

Un secondo profilo concerne la difformità interpretativa fra il modello tedesco - improntato al sistema delle eccezioni tipizzate di derivazione europea - e il modello statunitense, fondato sulla clausola generale del *fair use*. Mentre la giurisprudenza statunitense, anche in *Bartz v. Anthropic PBC*, ha ricondotto la liceità del *training* a una valutazione casistica fondata sui quattro fattori del *fair use*, il sistema europeo richiede l'ascrizione della condotta a fattispecie tipiche di eccezione, con la peculiarità che - per l'eccezione generale di *text and data mining* di cui all'art. 4 - il titolare può escludere unilateralmente l'applicabilità dell'eccezione mediante *opt-out*.

Di rilievo anche la centralità della nozione di *machine-readability* e la sua interpretazione. Sul punto i due gradi del giudizio amburghese pervengono a esiti opposti, ma la divergenza è più probatoria che concettuale: alla lettura ampia prospettata in *obiter* dal Landgericht - comprensiva delle riserve formulate in linguaggio naturale, in quanto suscettibili di interpretazione automatizzata - l'Oberlandesgericht non ha contrapposto una chiusura di principio, avendo anch'esso confermato la natura tecnologicamente aperta del requisito e avendo espressamente lasciato impregiudicato se una riserva in linguaggio naturale lo soddisfi allo stato attuale della tecnica. Esso ha piuttosto precisato che *machine-readable* è la riserva suscettibile non solo di essere rilevata, ma altresì di essere interpretata e osservata in via automatizzata, e ha posto a carico del titolare l'onere di allegare e provare tale requisito con riferimento al momento dell'utilizzazione - qui collocato nella seconda metà del 2021 -, negandolo nel caso concreto per difetto di prova sulle capacità tecniche allora disponibili. La differenza di impostazione si consuma dunque sul terreno dell'onere della prova e del momento rilevante, ma sortisce conseguenze pratiche opposte. La stessa sentenza di appello dà conto del carattere controverso della questione, censendo entrambi gli orientamenti dottrinali nonché le forme alternative discusse quali riserve *machine-readable* (file robots.txt, TDM Reservation

Protocol); l'impostazione dell'Oberlandesgericht è stata a propria volta criticata dalle organizzazioni rappresentative dei titolari²⁷⁹ in quanto priverebbe di fatto gli autori di uno strumento di opt-out praticabile. La questione è destinata, verosimilmente, a essere chiarita a livello sovranazionale anche alla luce dell'*AI Act*, il quale ha introdotto specifici obblighi di *compliance* con le riserve di *opt-out* in capo ai *providers* di modelli generali²⁸⁰.

Infine, con riferimento alla collocazione di *Kneschke v. LAION* nel quadro complessivo del contenzioso europeo in materia di IA generativa, la pronuncia si pone in contrasto apparente, sul piano dell'esito, con quella del *Landgericht München I* dell'11 novembre 2025 nel caso *GEMA v. OpenAI*, la quale ha invece concluso per la non operatività dell'eccezione di *text and data mining* rispetto alla c.d. «*memorization*» dei testi musicali nei parametri del modello. I due procedimenti riguardano momenti diversi del processo di IA generativa - la fase preparatoria di *dataset creation*, nel caso *Kneschke*; la fase di *memorization* e di *output generation*, nel caso *GEMA* - e dunque non si pongono fra loro in contraddizione logica. La complementarità fra le due pronunce è del resto avvalorata dalla stessa sentenza di appello, che richiama in due occasioni la decisione monacense: per lasciare aperta la questione della riconducibilità del training al § 44b, comma 1, UrhG e per rilevare che, ove in sede di generazione degli output vengano in essere atti riservati al titolare, questi conserva intatti i propri rimedi, fatte salve le eccezioni eventualmente applicabili.

VII. Il provvedimento in commento²⁸¹.

Sentenza del 10 dicembre 2025

²⁷⁹ Segnatamente, dalla *collecting Bild Kunst, Hamburger OLG-Urteil im Fall LAION: Politischer Handlungsbedarf ist noch offensichtlicher als zuvor!*, 11 dicembre 2025, all'indirizzo

<<https://www.bildkunst.de/news/news.44.hamburger-olg-urteil-im-fall-laion-politischer-handlungsbedarf-ist-noch-offensichtlicher-als-zuvor/>>.

²⁸⁰ Cfr. art. 53, par. 1, lett. c, regolamento (UE) 2024/1689, il quale impone ai *providers* di modelli di IA per finalità

[...]

Motivi della decisione

56 L'appello del ricorrente, ammissibile, è infondato, poiché la domanda è ammissibile ma infondata.

57 1. La domanda è ammissibile.

58 La domanda del ricorrente, ancora oggetto del giudizio d'appello, è diretta a che il convenuto si astenga dal riprodurre e/o far riprodurre la fotografia riportata nella domanda

59 ai fini della creazione di dataset di addestramento per l'IA, come avvenuto nell'ambito della realizzazione del dataset [omissis].

60 La domanda del ricorrente è sufficientemente determinata ai sensi del § 253, comma 2, n. 2, ZPO [codice di procedura civile]. Una sufficiente determinatezza sussiste di regola quando si fa riferimento alla concreta condotta lesiva e la domanda, quantomeno con l'ausilio delle allegazioni dell'attore, lascia riconoscere in modo inequivoco in quali caratteristiche della condotta contestata risiedano il fondamento e il punto di collegamento della violazione fatta valere e, con essa, dell'ordine inibitorio (BGH GRUR 2022, 1336 Rn. 12 – dortmund.de).

61 Nel caso di specie sono oggetto di causa processi interni al convenuto, sicché la concreta condotta lesiva non può essere rappresentata per immagini. Per le parti, e anche per il giudice, è sufficientemente chiaro quale concreto atto di utilizzazione venga contestato. Viene contestata la riproduzione (download) effettuata dal convenuto nel processo di creazione del dataset. Mediante il riferimento al dataset concreto ([omissis]) l'atto di utilizzazione è identificato. In quali punti del processo di creazione del dataset avvenga un download risulta dal grafico inserito dalla parte attrice nell'atto di citazione, tratto dal sito web del convenuto (p. 8 dell'atto di citazione).

62 Nella memoria di primo grado dell'11.04.2024 il convenuto ha dedotto, senza contestazione, che i download necessari per la creazione del dataset sono stati effettuati nella seconda metà del 2021.

63 Nella misura in cui il ricorrente, nell'atto di citazione, ha fatto riferimento in via integrativa a un elaborato del Servizio scientifico del Bundestag tedesco del 2018, contenente considerazioni generali sull'IA e sugli atti di riproduzione ad essa connessi, dalle allegazioni attoree non si ricava che il ricorrente faccia valere, nel caso di specie, che da parte del convenuto abbiano avuto luogo

generali di adottare una *policy* di *compliance* con la disciplina dell'Unione in materia di diritto d'autore, ivi compresa l'identificazione e il rispetto delle riserve di *opt-out* ai sensi dell'art. 4, par. 3, della direttiva (UE) 2019/790.

²⁸¹ Il testo configura una traduzione non ufficiale del provvedimento generata con IA e sottoposta a controllo di coerenza umano.

atti di riproduzione ulteriori rispetto al download. Il Landgericht ha quindi correttamente — e sul punto l'appello non muove censure — considerato quale atto di riproduzione oggetto di causa il solo download (pp. 2 e 9 s. della sentenza di primo grado).

64 La domanda di informazioni, diretta a ottenere informazioni sull'ampiezza dell'utilizzazione della fotografia, è stata dichiarata dall'attore, in primo grado, cessata nella materia del contendere, a seguito della deduzione del convenuto secondo cui il download necessario alla creazione del dataset era avvenuto nella seconda metà del 2021. Il convenuto si è associato a tale dichiarazione di parziale cessazione.

65 2. La domanda è infondata. Al ricorrente non spetta alcuna pretesa inibitoria ai sensi del § 97, comma 1, UrhG.

66 a. La fotografia oggetto di causa costituisce una fotografia semplice (*Lichtbild*) ai sensi del § 72, comma 1, UrhG. Il Landgericht ha correttamente accertato che la fotografia oggetto di causa costituisce in ogni caso una fotografia semplice ai sensi del § 72, comma 1, UrhG. Una protezione quale opera fotografica (*Lichtbildwerk*) ai sensi del § 2, comma 2, n. 5, UrhG non è, nel caso di specie, accertabile. Il ricorrente parla, è vero, della propria fotografia come di un'«opera». Ma allegazioni del ricorrente su una protezione quale opera fotografica ai sensi del § 2, comma 2, n. 5, UrhG non sono state svolte.

67 b. Il ricorrente è attivamente legittimato. Il Landgericht, sulla base del fatto che il ricorrente ha prodotto nell'udienza di primo grado una versione ad alta risoluzione della fotografia oggetto di causa e scatti paralleli, ha accertato che il ricorrente ha realizzato la fotografia. Su ciò il convenuto non prende posizione nella comparsa di risposta in appello. Non sussistono pertanto dubbi sulla correttezza dell'accertamento del Landgericht (§ 529, comma 1, n. 1, ZPO).

68 c. Il convenuto ha riprodotto la fotografia ai sensi del § 16, comma 1, UrhG. Il download della fotografia, pacifico, costituisce una riproduzione.

69 d. Il convenuto non ha acquisito il consenso del ricorrente. Il convenuto non ha neppure acquisito una licenza presso l'agenzia fotografica [omissis], dal cui sito la fotografia è stata scaricata. Il Landgericht ha anzi accertato che il convenuto non ha scaricato l'«immagine originale» — che l'agenzia fotografica [omissis] avrebbe messo a disposizione solo a fronte dell'acquisto di una licenza —, bensì una versione dell'immagine munita di una filigrana dell'agenzia. Si trattava, secondo il Landgericht, all'evidenza dell'immagine di anteprima inserita sul sito dell'agenzia quasi a fini pubblicitari. Non sussistono dubbi su tale accertamento del Landgericht.

70 e. L'utilizzazione è però giustificata da eccezioni al diritto d'autore. Nel caso di specie opera l'eccezione del § 44b UrhG, la cui applicabilità il Landgericht aveva lasciato aperta. Inoltre, il Landgericht ha, nel risultato, correttamente affermato una giustificazione in forza dell'eccezione del § 60d UrhG.

71 aa. Ai sensi del § 44b, comma 2, primo periodo, UrhG sono lecite le riproduzioni di opere lecitamente accessibili

per il *text and data mining*. Ai sensi del § 44b, comma 3, UrhG le utilizzazioni di cui al § 44b, comma 2, primo periodo, UrhG sono lecite soltanto se il titolare dei diritti non se le è riservate. Una riserva di utilizzazione relativa a opere accessibili online è efficace soltanto se effettuata in forma leggibile da macchina. Il § 44b UrhG è stato inserito nell'UrhG in attuazione dell'art. 4 della direttiva (UE) 2019/790 del 17 aprile 2019 sul diritto d'autore e sui diritti connessi nel mercato unico digitale e che modifica le direttive 96/9/CE e 2001/29/CE (di seguito: «direttiva DSM») (cfr. Dreier in Dreier/Schulze, UrhG, 8ª ed., § 44b Rn. 2).

72 aaa. Nel caso di specie il download oggetto di causa costituisce una riproduzione per il *text and data mining*.

73 Ai sensi del § 44b, comma 1, UrhG il *text and data mining* è l'analisi automatizzata di una o più opere digitali o digitalizzate al fine di trarne informazioni segnatamente su modelli (*pattern*), tendenze e correlazioni.

74 (1) La fotografia consultabile online è un'opera digitale.

75 (2) La riproduzione era funzionale all'analisi automatizzata. Secondo l'accertamento del Landgericht, sul punto non censurato dal ricorrente, il download è avvenuto al fine di confrontare, mediante un software, il contenuto dell'immagine con la descrizione dell'immagine già associata.

76 (3) Il Landgericht è pervenuto al corretto risultato che l'analisi del file di immagine ai fini del confronto con una descrizione preesistente costituisce un'analisi allo scopo di trarne informazioni ai sensi del § 44b, comma 1, UrhG.

77 Scopo del *text and data mining* deve essere, ai sensi del § 44b, comma 1, UrhG, trarne informazioni segnatamente su modelli, tendenze e correlazioni.

78 Uno degli scopi di un'analisi automatizzata ai sensi del § 44b, comma 1, UrhG può essere il riconoscimento di un modello in una pluralità di opere (cfr. Dreier in Dreier/Schulze, UrhG, 8ª ed., § 44b Rn. 5). L'applicazione del § 44b UrhG non è però limitata a ciò (cfr. Dreier, op. cit., Rn. 5 s.). Già secondo il tenore letterale del § 44b, comma 1, UrhG, il *text and data mining* ai sensi di tale disposizione comprende anche l'analisi automatizzata di singole opere digitali o digitalizzate al fine di trarne informazioni segnatamente su modelli, tendenze e correlazioni. Anche nella relazione al disegno di legge si legge che è lecito altresì analizzare in modo automatizzato una singola opera (BT-Drs. 19/27426, p. 88).

79 Correttamente il Landgericht ha ritenuto che l'accertamento della corrispondenza tra una concreta immagine e una concreta descrizione dell'immagine costituisca una correlazione ai sensi del § 44b, comma 1, UrhG. In ambito statistico il concetto di correlazione è impiegato per la connessione — non necessariamente causale, ma anche casuale — tra determinati fenomeni (cfr.

<https://www.duden.de/rechtschreibung/Korrelation>). Il concetto non è però limitato a ciò. Più in generale, una correlazione è una relazione reciproca (ibid.). La corrispondenza tra immagine e testo costituisce una siffatta relazione reciproca.

80 Anche a voler negare la sussistenza di una correlazione, la connessione tra immagine e testo sarebbe comunque da considerare (altra) informazione ai sensi del § 44b, comma 1, UrhG. La legge menziona modelli, tendenze e correlazioni solo come esempi tipici («segnatamente»; cfr. Paul in BeckOK IT-Recht, 19^a ed., § 44b UrhG Rn. 3a; Bomhard in BeckOK UrhR, 46^a ed., § 44b UrhG Rn. 11). Anche nella definizione di «text and data mining» dell'art. 2, n. 2, della direttiva DSM si legge che il concetto designa una tecnica di analisi automatizzata di testi e dati in forma digitale mediante la quale possono essere ottenute informazioni fra l'altro — ma non esclusivamente — su modelli, tendenze e correlazioni.

81 Tali esempi tipici devono peraltro essere considerati nell'interpretazione del concetto di «informazione». Gli esempi tipici — modelli, tendenze e correlazioni — depongono nel senso che le informazioni da trarre riguardino connessioni, strutture e/o processi.

82 L'analisi volta a stabilire se la fotografia oggetto di causa corrisponda alla descrizione esistente serve a trarre un'informazione sulla connessione tra immagine e descrizione dell'immagine e ricade dunque nel § 44b, comma 1, UrhG. Può pertanto rimanere aperta la questione se debba essere incluso anche lo scopo mediato di un guadagno informativo trasversale alle opere, diretto a consentire all'IA, nel successivo addestramento sulla base del dataset, di riconoscere modelli trasversali alle opere (secondo la deduzione del convenuto a p. 11 della comparsa di risposta, ad esempio, che le banane sono per lo più raffigurate in colore giallo).

83 Poiché già il confronto tra immagine e descrizione dell'immagine costituisce un'analisi allo scopo di trarre informazioni ai sensi del § 44b, comma 1, UrhG, non rileva nel caso di specie se il successivo addestramento di modelli di IA generativa soddisfi i presupposti del § 44b, comma 1, UrhG (cfr. LG München I GRUR-RS 2025, 30204; Bomhard in BeckOK UrhR, 47^a ed., § 44b UrhG Rn. 11a con ulteriori riferimenti; Paul in BeckOK IT-Recht, 19^a ed., § 44b UrhG Rn. 3a; Pesch in Spindler/Schuster/Kaesling, Recht der elektronischen Medien, 5^a ed., § 44b Rn. 20; von Welser GRUR-Prax 2023, 516 Rn. 19; Baumann NJW 2023, 3673 Rn. 14; Dregelies GRUR 2024, 1484, 1486; Konertz WRP 2025, 1253, 1255 ss.; Rauer/Bibi WRP 2025, 33, 35 ss.; [omissis]/[omissis], Urheberrecht und Training generativer KI-Modelle (allegato K 11), pp. 66 ss.; cfr. però anche [omissis]/[omissis], op. cit., p. 95, sulla questione del rilievo dell'art. 4 della direttiva DSM per le utilizzazioni a monte del processo di addestramento di un modello di IA generativa).

84 Le norme sul *text and data mining* (§§ 44b, 60d UrhG) non vanno neppure ristrette in via teleologica nel senso che esse non coprano le misure preparatorie dell'addestramento dell'IA. Il § 44b UrhG è stato introdotto con la legge di adeguamento del diritto d'autore alle esigenze del mercato unico digitale, entrata in vigore il 07.06.2021. Nella relazione a tale legge si legge:

85 «Mediante le nuove ovvero modificate autorizzazioni legali (§§ 44b, 60a ss. UrhG-E), specialmente nel settore

dell'istruzione e della scienza presso scuole, università e organizzazioni di ricerca nonché presso importanti istituzioni del sapere quali le biblioteche, più persone ottengono accesso ai contenuti (principio di sviluppo sostenibile n. 5). [...] Esse ampliano inoltre l'utilizzazione di opere protette dal diritto d'autore e di altri materiali protetti mediante il text and data mining per la ricerca scientifica e per altri scopi, così da promuovere le innovazioni (principio di sviluppo sostenibile n. 6, indicatore di sostenibilità n. 9.1). Ciò è di particolare importanza per l'apprendimento automatico quale tecnologia di base per l'intelligenza artificiale» (BT-Drs. 19/27426, p. 60).

86 Anche il legislatore europeo muove dall'applicazione delle disposizioni sul *text and data mining* ai modelli di IA. Nell'art. 53, comma 1, lett. c, del regolamento sull'IA (regolamento (UE) 2024/1689 del 13.06.2024) si legge: «I fornitori di modelli di IA per finalità generali: [...] attuano una strategia per il rispetto del diritto d'autore dell'Unione e dei diritti connessi e, in particolare, per l'individuazione e il rispetto di una riserva di diritti fatta valere ai sensi dell'articolo 4, paragrafo 3, della direttiva (UE) 2019/790, anche mediante tecnologie all'avanguardia.» L'art. 4 della direttiva DSM ivi richiamato ha ad oggetto eccezioni e limitazioni per il *text and data mining*. I considerando 99 e 105 del regolamento sull'IA fanno riferimento ai modelli di IA generativa.

87 Inoltre, secondo i considerando della direttiva DSM, i procedimenti di *text and data mining* rivestono grande importanza anche per lo sviluppo di nuove applicazioni o tecnologie (considerando 18 della direttiva DSM). E secondo il considerando 3 della direttiva DSM le disposizioni normative pertinenti devono essere adeguate alle esigenze future, affinché lo sviluppo tecnologico non venga ostacolato.

88 Nella misura in cui in dottrina si sostiene che nel *text and data mining* ai sensi del § 44b, comma 1, UrhG dovrebbero essere «dischiuse soltanto le informazioni celate nei dati», non però utilizzato il contenuto della creazione intellettuale (cfr. Schack NJW 2024, 113 Rn. 8; cfr. anche [omissis]/[omissis], Urheberrecht und Training generativer KI-Modelle, allegato K 11, pp. 2, 67 ss., che distinguono tra «informazioni semantiche» (idee e contenuti) e «informazioni sintattiche» (forma/configurazione concreta); cfr. anche [omissis] GRUR 2024, 1676), anche ciò non conduce a una diversa valutazione. Nel confronto tra immagine e descrizione dell'immagine viene valutata soltanto l'«informazione» contenuta nell'immagine e non la sua configurazione concreta. Che il dataset possa e debba essere successivamente utilizzato per l'addestramento di IA generativa non osta all'applicazione del § 44b, comma 1, UrhG alla riproduzione oggetto di causa. In tale addestramento non vengono utilizzate le riproduzioni realizzate dal convenuto. Eventuali ripercussioni sulle possibilità di sfruttamento dell'autore/titolare del diritto connesso vanno considerate nell'ambito del test dei tre stadi.

89 Che, secondo la relazione al disegno di legge, non siano coperte dallo scopo del § 44b, comma 1, UrhG le condotte esclusivamente dirette a raccogliere e memorizzare contenuti per creare archivi digitali paralleli (BT-Drs. 19/27426, 88), non osta neppure, nel caso di specie, all'applicabilità del § 44b, comma 1, UrhG alla riproduzione oggetto di causa. Il ricorrente ha, è vero, contestato che il convenuto abbia nuovamente cancellato la fotografia; alla deduzione del convenuto secondo cui il dataset contiene soltanto collegamenti ipertestuali (hyperlink) e ulteriori metadati relativi alle immagini, il ricorrente non ha però opposto contestazioni rilevanti né in primo grado né in appello. L'accertamento del Landgericht secondo cui nel dataset non sono state inserite le immagini, bensì soltanto hyperlink (p. 7 della sentenza di primo grado), non è stato censurato dal ricorrente con l'appello. Tale accertamento va dunque posto a fondamento ai sensi del § 529, comma 1, n. 1, ZPO.

90 Nella misura in cui il ricorrente, nella memoria del 29.08.2024, non autorizzata in primo grado, ha dedotto che la fotografia oggetto di causa sarebbe memorizzata in un cloud di Amazon, tale deduzione è irrilevante ai sensi del § 296a ZPO. Nell'atto di appello il ricorrente non ha ribadito tale deduzione (cfr. Bacher in BeckOK ZPO, 58ª ed., § 296a Rn. 9; Röß in Musielak/Voit, ZPO, 22ª ed., § 296a Rn. 6; Prütting in MüKoZPO, 7ª ed., § 296a Rn. 9; Greger in Zöller, ZPO, 36ª ed., § 296a Rn. 3).

91 Una pretesa di cancellazione ai sensi del § 44b, comma 2, secondo periodo, UrhG (cfr. Dreier in Dreier/Schulze, UrhG, 8ª ed., § 44b Rn. 15) non è fatta valere dal ricorrente nel presente giudizio.

92 (4) La fotografia consultabile sul sito dell'agenzia [omissis] è altresì un'opera lecitamente accessibile ai sensi del § 44b, comma 2, primo periodo, UrhG. Tale accertamento del Landgericht non è stato censurato con l'appello.

93 (5) All'utilizzazione non osta, nel caso di specie, una riserva di utilizzazione ai sensi del § 44b, comma 3, UrhG. Il § 44b, comma 3, UrhG recita:

94 «Le utilizzazioni di cui al comma 2, primo periodo, sono lecite soltanto se il titolare dei diritti non se le è riservate. Una riserva di utilizzazione relativa a opere accessibili online è efficace soltanto se effettuata in forma leggibile da macchina.»

95 Considerata la formulazione del § 44b, comma 3, primo periodo, UrhG quale presupposto di applicabilità dell'eccezione, l'onere della prova che il titolare dei diritti non si sia riservato l'utilizzazione grava sull'utilizzatore (cfr. relazione ufficiale BT-Drs. 19/27426, p. 89; Dreier in Dreier/Schulze, UrhG, 8ª ed., § 44b Rn. 9; Paul in BeckOK IT-Recht, 20ª ed., § 44b UrhG Rn. 5; Sattler in Borges/Keil, Big Data, 1ª ed., § 10 Rn. 139; Bullinger in Wandtke/Bullinger, UrhG, 6ª ed., § 44b; Grimm/Münster GRUR-Prax 2025, 443 Rn. 27; in senso contrario Pesch in Spindler/Schuster/Kaesling, UrhG, 5ª ed., § 44b Rn. 47), nel caso di specie sul convenuto.

96 È vero che la formulazione dell'art. 4, par. 3, della direttiva DSM depone, all'opposto, nel senso che l'onere

della prova grava sul titolare dei diritti («Le eccezioni e le limitazioni di cui al paragrafo 1 si applicano a condizione che i rispettivi titolari non abbiano espressamente riservato in modo appropriato l'utilizzo delle opere e degli altri materiali di cui al paragrafo 1, ad esempio con mezzi leggibili da macchina nel caso di contenuti resi pubblicamente disponibili online»). Tuttavia, il legislatore europeo, considerata la formulazione «un'eccezione o una limitazione» nell'art. 4 della direttiva DSM, ha lasciato liberi gli Stati membri di limitare a priori, per il text and data mining, l'oggetto della protezione del diritto d'autore ovvero di creare una fattispecie eccezzuativa (cfr. Wulf/Löbeth GRUR 2024, 737, 741; Grimm/Münster GRUR-Prax 2025, 443 Rn. 29; cfr. anche Schack GRUR 2021, 904, 907). Se uno Stato membro decide — come nel caso della Germania — di fare del text and data mining l'oggetto di una disciplina eccezzuativa, ne conseguono anche le corrispondenti regole di allegazione e di prova, poiché in linea di principio chi invoca un'eccezione deve anche allegarne e provarne i presupposti. Come esposto, il legislatore tedesco ha deciso di porre l'onere di allegazione e di prova dell'assenza di una riserva di utilizzazione, in linea di principio (v. però infra sulla questione della leggibilità da macchina), a carico dell'utilizzatore.

97 Sul titolare dei diritti grava però un onere di allegazione secondario (cfr. Dreier in Dreier/Schulze, UrhG, 8ª ed., § 44b Rn. 9; Sattler in Borges/Keil, Big Data, 1ª ed., § 10 Rn. 139; Pesch in Spindler/Schuster/Kaesling, UrhG, 5ª ed., § 44b Rn. 47). Sulla controparte processuale della parte gravata dalla prova di un fatto negativo incombe un onere di allegazione secondario, la cui ampiezza dipende dalle circostanze del caso concreto (cfr. BGH, ord. 20.6.2017 – VI ZR 505/16, BeckRS 2017, 117730 Rn. 4). Se l'avversario della pretesa non adempie il proprio onere di allegazione secondario, l'affermazione dell'istante si considera ammessa ai sensi del § 138 ZPO (cfr. BGH NJW-RR 2025, 1004 Rn. 39). Il titolare dei diritti deve, in caso di controversia, allegare che una riserva di utilizzazione esisteva e come essa era configurata. A tale onere di allegazione ampliato il ricorrente ha adempiuto nel caso di specie mediante il richiamo alla riserva contenuta nelle condizioni d'uso del fornitore [omissis] e nel codice sorgente del sito web di [omissis].

98 La riserva contenuta nelle condizioni d'uso del fornitore [omissis] e nel codice sorgente del sito di [omissis] non esclude, nel caso di specie, l'utilizzazione oggetto di causa della fotografia realizzata dal ricorrente.

99 (a) Il ricorrente può, nel caso di specie, in linea di principio invocare una riserva dichiarata dall'agenzia fotografica [omissis], benché [omissis] fosse titolare soltanto di diritti di utilizzazione semplici.

100 Se con il concetto di titolare dei diritti sia inteso anche il titolare di diritti di utilizzazione semplici va determinato autonomamente per ciascuna disposizione. Tanto il legislatore europeo quanto quello tedesco impiegano il concetto, di volta in volta, in modo differente.

101 Nella direttiva 2001/29/CE del 22.05.2001 sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione (di seguito: direttiva InfoSoc) il concetto di «titolare dei diritti» è impiegato ad esempio negli artt. 4, 5 e 8. In particolare il considerando 25 depone nel senso che per titolare dei diritti ai sensi della direttiva si intenda l'autore ovvero il titolare di diritti di utilizzazione esclusivi («Occorre chiarire che tutti i titolari di diritti riconosciuti dalla presente direttiva debbono godere del diritto esclusivo di rendere accessibili al pubblico le opere protette dal diritto d'autore e i materiali protetti da altri diritti mediante trasmissioni interattive su richiesta»).

102 Nell'art. 3, par. 3, primo periodo, della direttiva DSM il concetto di titolare dei diritti comprende invece anche il titolare di diritti di utilizzazione semplici. Ivi si legge: «I titolari dei diritti devono poter applicare misure atte a garantire la sicurezza e l'integrità delle reti e delle banche dati in cui sono ospitate le opere o altri materiali.» Tale previsione è attuata in Germania dal § 60d, comma 6, UrhG («I titolari dei diritti sono legittimati ad adottare le misure necessarie per evitare che la sicurezza e l'integrità delle loro reti e banche dati siano messe in pericolo dalle riproduzioni di cui al comma 1»). Si guarda dunque al luogo di memorizzazione. Si intende prevenire una limitazione dell'offerta online per gli altri utenti dovuta al sovraccarico dei server causato dagli innumerevoli accessi durante gli atti di riproduzione (cfr. Bullinger in Wandtke/Bullinger, UrhG, 6ª ed., § 60d Rn. 35). In caso di concessione di diritti di utilizzazione la memorizzazione non avviene (necessariamente) ad opera dell'autore, bensì (anche) ad opera del titolare dei diritti di utilizzazione (anche semplici).

103 L'art. 4 della direttiva DSM, e con esso anche il § 44b UrhG che vi dà attuazione, vanno intesi nel senso che chi invoca l'eccezione deve in linea di principio rispettare anche le riserve dichiarate da un titolare di diritti di utilizzazione semplici, purché per l'utilizzazione dell'opera si acceda a una fonte proveniente dal titolare dei diritti di utilizzazione semplici (cfr. Leistner GRUR 2024, 1665, 1671; in senso contrario Pukas K&R 2025, 677, 681; Rauer/Bibi WRP 2025, 33, 38: soltanto alle condizioni dei §§ 164 ss. BGB in via analogica).

104 Il § 44b UrhG prevede una soluzione di *opt-out* per un'utilizzazione che in linea di principio può avvenire senza che siano state rilasciate licenze da parte dell'autore/titolare del diritto connesso ovvero del titolare dei diritti di utilizzazione esclusivi. La decisione se e in quale modo un'opera o un diritto connesso venga utilizzato spetta in linea di principio all'autore/titolare del diritto connesso ovvero al titolare dei diritti di utilizzazione esclusivi e, di regola, non al titolare di diritti di utilizzazione semplici (cfr., quanto al § 61d, comma 2, UrhG (diritto di opposizione del titolare): Rauer in Dreier/Schulze, UrhG, 8ª ed., § 61d Rn. 10; quanto ai §§ 7 ss. UrhDaG (richiesta di blocco): Bullinger in Wandtke/Bullinger, UrhG, 6ª ed., § 7 UrhDaG Rn. 8; cfr.

anche Kefferpütz in Wandtke/Bullinger, UrhG, 6ª ed., §§ 111b, 111c Rn. 16).

105 Nell'interpretazione va però considerato che la disciplina del § 44b, comma 3, UrhG attua la previsione dell'art. 4, par. 3, della direttiva DSM e che pertanto deve essere garantito l'«effet utile» (efficacia pratica del diritto dell'Unione; così anche Leistner GRUR 2024, 1665, 1671).

106 Ai fini dell'effettiva salvaguardia dei diritti di decisione dell'autore/titolare del diritto connesso ovvero del titolare dei diritti di utilizzazione esclusivi in ordine all'autorizzazione o al divieto del text and data mining, devono in ogni caso essere rispettate anche le riserve di utilizzazione poste, con il consenso dell'autore/titolare del diritto connesso ovvero del titolare dei diritti esclusivi, da titolari di diritti di utilizzazione semplici.

107 Il Collegio ha deciso, con riguardo al § 95a UrhG, che il titolare dei diritti può opporsi all'elusione di misure tecnologiche anche quando le misure siano state adottate dal gestore della piattaforma. Titolare dei diritti e gestore della piattaforma vanno in linea di principio considerati come un'unità con riguardo all'elusione di misure di protezione ai sensi del § 95a UrhG (Collegio, sent. 21.11.2024 – 5 U 54/23, GRUR-RS 2024, 33119 Rn. 70; secondo Specht-Riemenschneider/Raue in Dreier/Schulze, UrhG, 8ª ed., § 95a Rn. 5, ciò vale salvo che la controparte dimostri che nel caso concreto le cose stiano eccezionalmente in modo diverso). Nel caso allora deciso il gestore della piattaforma era contrattualmente obbligato verso il titolare dei diritti all'adozione di misure tecnologiche di protezione (cfr. Collegio, sent. 21.11.2024 – 5 U 54/23, GRUR-RS 2024, 33119 Rn. 41; Paul in BeckOK IT-Recht, 19ª ed., § 95a UrhG Rn. 8). Che sussistesse un obbligo contrattuale di [omissis] di porre una riserva di utilizzazione non si ricava dalle allegazioni del ricorrente.

108 Tuttavia, nel caso di specie la circostanza che il ricorrente si avvalga di un'agenzia per lo sfruttamento delle proprie immagini depone nel senso che una riserva di utilizzazione posta dall'agenzia sia imputata anche al ricorrente quale titolare dei diritti. Se un fotografo si avvale di un fornitore di stock photo per commercializzare le proprie immagini, egli si dichiara con ciò di regola d'accordo che la commercializzazione dei diritti sulle foto avvenga secondo le condizioni del fornitore. Non è decisivo se la riserva si trovasse già sul sito di [omissis] quando il ricorrente concesse a [omissis] i diritti di utilizzazione sulla propria foto. Che [omissis] con tale riserva agisse contro gli interessi del ricorrente non è né dedotto né altrimenti ravvisabile. Al contrario, il ricorrente invoca espressamente tale riserva.

109 (2) Dalla riserva presente sul sito di [omissis] si ricava, in via di interpretazione, che essa intende includere gli atti di utilizzazione oggetto di causa. La riserva di utilizzazione deve essere dichiarata espressamente (cfr. la relazione al disegno di legge BT-Drs. 19/27426, p. 89, nonché l'art. 4, par. 3, della direttiva DSM; cfr. anche Bomhard in BeckOK UrhR, 47ª ed., § 44b UrhG Rn. 25). Nella riserva presente sul sito di [omissis] e nel codice sorgente l'utilizzazione per il text and data mining non è menzionata specificamente.

Viene vietata in generale l'analisi automatica dei contenuti. Ciò va interpretato nel senso che vi rientra non soltanto il c.d. crawling da parte di un terzo, a monte della riproduzione qui oggetto di causa, ma anche il download automatizzato ai fini del confronto immagine-testo da parte del convenuto.

110 (3) La riserva di utilizzazione non presentava però, nel caso di specie, la forma prevista dal § 44b, comma 3, secondo periodo, UrhG. In base ad esso, una riserva di utilizzazione relativa a opere accessibili online è efficace soltanto se effettuata in forma leggibile da macchina.

111 È vero che nell'art. 4, par. 3, della direttiva DSM si legge che le eccezioni e limitazioni di cui all'art. 4, par. 1, si applicano a condizione che i rispettivi titolari non abbiano espressamente riservato in modo appropriato le opere e gli altri materiali ivi menzionati, ad esempio con mezzi leggibili da macchina nel caso di contenuti resi pubblicamente disponibili online. Dal considerando 18 si ricava però che il legislatore europeo, per i contenuti pubblicati online, considera appropriata soltanto una riserva leggibile da macchina («Nel caso di contenuti resi pubblicamente disponibili online, dovrebbe essere ritenuto appropriato riservare tali diritti mediante strumenti che consentano una lettura automatizzata, inclusi i metadati e le condizioni d'uso di un sito web o di un servizio»; ancora più chiaramente nella versione inglese: «In the case of content that has been made publicly available online, it should only be considered appropriate to reserve those rights by the use of machine-readable means, including metadata and terms and conditions of a website or a service»).

112 (a) Come esposto, ai sensi del § 44b, comma 3, primo periodo, UrhG grava sull'utilizzatore l'onere della prova che il titolare dei diritti non si sia riservato l'utilizzazione per il text and data mining. Il presupposto della leggibilità da macchina deve invece essere allegato e provato dal titolare dei diritti (cfr. Leistner GRUR 2024, 1665, 1674). Ciò risulta dalla formulazione del § 44b, comma 3, secondo periodo, UrhG («efficace soltanto se»). In ogni caso, l'onere di allegazione secondario del titolare dei diritti si estende anche all'efficacia della riserva (cfr. Sattler in Borges/Keil, Big Data, 1ª ed., § 10 Rn. 139). Il titolare dei diritti deve allegare che la riserva di utilizzazione da lui scelta, o comunque a lui imputabile, era leggibile da macchina al momento dell'atto di utilizzazione contestato. Considerato che il titolare dei diritti tipicamente non dispone di approfondite conoscenze tecniche, alle sue allegazioni non vanno posti requisiti elevati. Dal punto di vista del Collegio è però in ogni caso necessaria un'allegazione sulle possibilità tecniche note al momento rilevante dell'utilizzazione.

113 (b) Nel caso di specie, dalle allegazioni del ricorrente non si ricava che la riserva fosse leggibile da macchina al momento rilevante dell'utilizzazione.

114 La riserva era redatta in linguaggio naturale. Essa era contenuta, per un verso, nelle condizioni d'uso sul sito web e, per altro verso, in forma leggibile dall'uomo nel

codice sorgente della pagina (pp. 14 s. della memoria di parte attrice del 03.07.2024).

115 Se e in quale misura anche una riserva formulata in linguaggio naturale soddisfi i requisiti del § 44b, comma 3, secondo periodo, UrhG è controverso (in senso affermativo: Paul in BeckOK IT-Recht, 19ª ed., § 44b UrhG Rn. 5; Steinrötter/Schauer in Barudi, Das neue UrhR, 1ª ed., § 4 Text und Data Mining, Forschung und Lehre Rn. 14; Dworschak in Fromm/Nordemann, UrhG, 13ª ed., § 44b Rn. 26; Pukas K&R 2025, 677, 679 s.; Konertz WRP 2025, 1253, 1262; Rauer/Bibi WRP 2025, 33, 36 ss.; Wachtel ZUM 2025, 77, 79 s.; Leistner GRUR 2024, 1665, 1673; in senso negativo: Pesch in Spindler/Schuster/Kaesling, Recht der elektronischen Medien, 5ª ed., § 44b Rn. 40; Bomhard in BeckOK UrhR, 47ª ed., § 44b UrhG Rn. 31, con rinvio all'art. 2, n. 13, e al considerando 35 della direttiva (UE) 2019/1024; Maamar ZUM 2023, 481, 484; Schöttle/Völker K&R 2025, 22, 23 s.; cfr. sulle riserve di utilizzazione anche BGH GRUR 2010, 628 Rn. 37 – Vorschäubilder I; BGH GRUR 2024, 1528 Rn. 49 – Coffee; OLG München, sent. 17.08.2017 – U 2225/15 Kart, GRUR-RS 2017, 122821 Rn. 91 ss. – Adblocker). Nella misura in cui una riserva in linguaggio naturale è considerata insufficiente, quali forme leggibili da macchina si discutono, tra l'altro, l'impiego di un file robots.txt e il TDM Reservation Protocol (cfr. Bomhard in BeckOK UrhR, 47ª ed., § 44b UrhG Rn. 32 ss.; cfr. anche Pesch in Spindler/Schuster/Kaesling, Recht der elektronischen Medien, 5ª ed., § 44b UrhG Rn. 41 s.).

116 Ad avviso del Collegio, il legislatore non ha prescritto alcuna forma concreta, ma ha posto il presupposto che la forma scelta sia leggibile da macchina. La formulazione è tecnologicamente aperta. Ciò emerge anche dalla relazione al disegno di legge, dove si legge che la riserva può essere contenuta anche nel colophon (Impressum) o nelle condizioni generali di contratto (AGB), purché anche ivi sia leggibile da macchina (BT-Drs. 19/27426, p. 89). Anche nel considerando 18 della direttiva DSM si legge che la dichiarazione della riserva con mezzi leggibili da macchina vale anche per i metadati e per le condizioni d'uso di un sito web o di un servizio.

117 Nella questione della leggibilità da macchina non rileva soltanto che il testo possa essere rilevato automaticamente, ma anche che esso possa essere interpretato automaticamente, nel senso che, in un procedimento automatizzato, possa condurre a che i contenuti coperti dalla riserva non vengano analizzati (v. la relazione al disegno di legge: «Scopo della disciplina è infatti, per un verso, aprire ai titolari dei diritti la possibilità di vietare l'utilizzazione fondata sull'autorizzazione legale. Al contempo, la disciplina mira ad assicurare, per i contenuti accessibili online, che i processi automatizzati, che costituiscono criterio tipico del text and data mining, possano effettivamente essere eseguiti in modo automatizzato», BT-Drs. 19/27426, p. 89).

118 Sotto il profilo processuale rileva se la forma scelta fosse leggibile da macchina al momento dell'atto di

utilizzazione oggetto di causa (seconda metà del 2021). Occorre fare riferimento all'atto di utilizzazione oggetto di causa, ossia il download eseguito dal convenuto. Non è decisivo se il crawling a monte del dataset di [omissis] fosse lecito. Il convenuto ha soltanto estratto gli URL delle immagini dal dataset di [omissis] e ha poi scaricato esso stesso le immagini, per il confronto immagine-testo da esso eseguito, dalle pagine originali.

119 (c) Che la forma scelta per la riserva fosse leggibile da macchina al menzionato momento dell'atto di utilizzazione oggetto di causa non si ricava dalle allegazioni del ricorrente.

120 Il ricorrente ha riportato il risultato dell'inserimento delle condizioni d'uso in ChatGPT. Tale inserimento è però avvenuto nel 2023. Il chatbot ChatGPT è stato pubblicato, per quanto noto al giudice, soltanto alla fine del 2022. Il ricorrente ha inoltre richiamato la possibilità di verificare in modo automatizzato, con determinati strumenti, se singoli siti web contengano riserve di diritti (opt-out). In concreto, il ricorrente menziona lo strumento «weboptout». Anche al riguardo il ricorrente non ha allegato che tale strumento fosse già disponibile nella seconda metà del 2021.

121 Nella memoria del 03.07.2024 il ricorrente ha dedotto che è possibile leggere e rispettare anche riserve di diritti in forma testuale, e ha offerto su tale deduzione prova per perizia. Ma anche in questo caso le allegazioni del ricorrente si riferiscono alle possibilità odierne e non a quelle esistenti nel 2021.

122 La circostanza, richiamata dal Landgericht, che il convenuto stesso abbia impiegato software per eseguire un confronto tra immagini e testi non giustifica l'accertamento che la riserva dichiarata fosse leggibile da macchina nella seconda metà del 2021. Che un programma possa confrontare descrizioni di immagini e immagini non significa senz'altro che fosse possibile analizzare estese condizioni d'uso. A ciò si aggiunge che la riserva oggetto di causa non menziona espressamente il text and data mining, ma soltanto in via di interpretazione può ricavarsi dalla riserva se essa escluda anche tale finalità di utilizzazione. Ciò richiede una comprensione automatica del testo. Come esposto, la parte attrice non ha allegato che al momento rilevante dell'utilizzazione esistessero le possibilità tecniche per una siffatta comprensione automatica del testo.

123 (d) Nella misura in cui il ricorrente, nella memoria del 29.08.2024, ha dedotto che uno sviluppatore esperto avrebbe potuto realizzare uno strumento come il «weboptout-Tool» anche nel 2021 in circa 3 ore di lavoro, offrendone prova, tale deduzione è processualmente irrilevante ai sensi del § 296a ZPO. Essa è avvenuta in una memoria non autorizzata, depositata dopo la chiusura della trattazione orale di primo grado. Il ricorrente non ha neppure ribadito tale deduzione nell'atto di appello. A prescindere da ciò, alla sua considerazione in appello osta anche che non ricorrono i presupposti del § 531, comma 2, ZPO. Ragioni per cui una siffatta deduzione non potesse essere svolta già prima della chiusura della trattazione

orale di primo grado non sono né dedotte né altrimenti ravvisabili. Non rileva, in via integrativa, che la stessa parte attrice — nella memoria del 29.08.2024, peraltro non autorizzata e dunque irrilevante ai sensi del § 296a ZPO (ivi pp. 2 s.) — abbia dedotto che il Landgericht, in udienza, aveva definito la questione della leggibilità da macchina il provvisorio «Big Point» della controversia e aveva inoltre segnalato che era dubbio se tale possibilità esistesse già nel 2021, cioè al momento della riproduzione in questione.

124 Può quindi rimanere aperta la questione se, nel caso in cui la riserva non sia leggibile con i programmi presenti sul mercato, l'utilizzatore sia obbligato a sviluppare programmi propri per rendere possibile la leggibilità da macchina.

125 (4) L'esecuzione del test dei tre stadi dà come risultato un'utilizzazione lecita.

126 Anche le utilizzazioni ai sensi del § 44b UrhG sono soggette al test dei tre stadi. Base europea è, come esposto, l'art. 4 della direttiva DSM. Ai sensi dell'art. 7, par. 2, primo periodo, della direttiva DSM, l'art. 5, par. 5, della direttiva InfoSoc si applica alle eccezioni e limitazioni previste sotto tale titolo.

127 Ai sensi dell'art. 5, par. 5, della direttiva InfoSoc, le eccezioni e limitazioni possono essere applicate soltanto in determinati casi speciali che non siano in contrasto con lo sfruttamento normale dell'opera o degli altri materiali e non arrechino ingiustificato pregiudizio agli interessi legittimi del titolare. In ciò risiede, per un verso, un'indicazione conformativa rivolta al legislatore nazionale in relazione alla concretizzazione delle eccezioni al diritto d'autore; per altro verso, l'art. 5, par. 5, della direttiva InfoSoc è anche parametro per l'applicazione del diritto nel caso concreto (cfr. BGH GRUR 2023, 1531 Rn. 47 – Metall auf Metall V).

128 Il test dei tre stadi va applicato nel caso di specie benché la fotografia oggetto di causa sia una fotografia semplice ai sensi del § 72 UrhG e non un'opera fotografica ai sensi del § 2, comma 2, n. 5, UrhG. A differenza della protezione autorale delle opere fotografiche, che ricadono nella direttiva InfoSoc, la protezione delle fotografie semplici non è armonizzata a livello di diritto dell'Unione. L'art. 6, secondo periodo, della direttiva sulla durata di protezione riserva anzi agli Stati membri la possibilità di prevedere la protezione di «altre fotografie» (cfr. Schulze/Dreier in Dreier/Schulze, UrhG, 8ª ed., § 72 Rn. 2; A. Nordemann in Fromm/Nordemann, UrhG, 13ª ed., § 72 Rn. 4). Può rimanere aperta la questione se gli «altri materiali protetti» menzionati nella direttiva DSM (art. 1 direttiva DSM) si riferiscano soltanto ai materiali protetti coperti anche dalla direttiva InfoSoc (ivi art. 3, par. 2). In ogni caso deve muoversi da un'interpretazione unitaria delle eccezioni al diritto d'autore. Ai sensi del § 72, comma 1, UrhG le disposizioni della Parte 1 dell'UrhG valide per le opere fotografiche si applicano in via corrispondente (cfr. Dreier in Dreier/Schulze, UrhG, 8ª ed., § 44b Rn. 6). Anche nell'ambito dell'applicazione meramente corrispondente delle eccezioni va dunque

effettuata un'interpretazione conforme alla direttiva (cfr. Thum in Wandtke/Bullinger, UrhG, 6ª ed., § 72 Rn. 85 con ulteriori riferimenti) e anche il test dei tre stadi va applicato alle fotografie semplici (cfr. BGH GRUR 2025, 493 Rn. 24 – Cornea-Implantat).

129 (a) Al primo requisito del test dei tre stadi, da eseguire nel presente caso concreto, è soddisfatto mediante l'introduzione dell'eccezione nel § 44b UrhG, trattandosi di un caso speciale disciplinato per legge.

130 (b) Nella verifica da compiere al secondo stadio — se lo sfruttamento normale dell'opera o dell'altro materiale protetto non sia pregiudicato — occorre considerare se, mediante l'utilizzazione contestata, siano ridotte le possibilità del titolare dei diritti di collocamento legittimo (cfr. CGUE GRUR 2017, 610 Rn. 70 – Stichting Brein; BGH GRUR 2023, 1531 Rn. 49 – Metall auf Metall V). Rileva al riguardo se l'utilizzazione in questione entri in concorrenza diretta con l'utilizzazione tradizionale, ossia se si interferisca nello sfruttamento primario (cfr. BGH GRUR 2020, 859 Rn. 72 – Reformistischer Aufbruch II).

131 Nel caso di specie la riproduzione oggetto di causa costituisce un processo puramente interno al convenuto. Nel dataset reso disponibile all'esterno la riproduzione non è contenuta; vi è soltanto un link alla fonte legittima. Che il ricorrente non riceva denaro per l'utilizzazione privilegiata ai sensi del § 44b UrhG è il risultato della scelta legislativa di non prevedere un obbligo di remunerazione per le utilizzazioni che ricadono nell'eccezione (cfr. Bomhard in BeckOK UrhR, 47ª ed., § 44b Rn. 24; Schack NJW 2024, 113 Rn. 11; in senso critico Dreier in Dreier/Schulze, UrhG, 8ª ed., § 44b Rn. 14).

132 Può rimanere aperta la questione se debbano essere incluse anche le conseguenze mediate dell'atto di utilizzazione. Anche considerando che il dataset serve all'addestramento di IA generativa e che, in un passaggio ulteriore, l'intelligenza artificiale genera immagini che possono entrare in concorrenza con le fotografie del ricorrente, il requisito del secondo stadio è soddisfatto. Che anche lo sviluppo dell'IA generativa costituisca uno sviluppo voluto dal legislatore europeo lo mostra il regolamento sull'IA (v. considerando 99). Anche il legislatore tedesco, come esposto, nell'introdurre il § 44b UrhG aveva già in vista l'utilizzazione per applicazioni di IA (cfr. Bomhard in BeckOK UrhR, 46ª ed., § 44b UrhG Rn. 11b).

133 Nella misura in cui, nella generazione di immagini da parte dell'IA, si giunga ad atti di utilizzazione riservati al titolare dei diritti (come ad esempio riproduzioni ai sensi del § 16 UrhG), il titolare può agire in giudizio contro di essi — fatta salva l'operatività delle eccezioni — (cfr., per i testi di canzoni, LG München I GRUR-RS 2025, 30204; cfr. anche Baumann NJW 2023, 3673 Rn. 30; Leistner GRUR 2025, 1123, 1129). Nella misura in cui ciò non avvenga, la singola immagine può bensì entrare in concorrenza con la fotografia realizzata dal ricorrente; in quale misura ne risultino però pregiudicate le possibilità di collocamento del ricorrente non può essere previsto, considerata la molteplicità delle possibili configurazioni di immagini.

Considerata la volontà anche del legislatore europeo di non contrastare lo sviluppo dei modelli di IA generativa, siffatte ripercussioni astratte, collocate nel futuro rispetto alla creazione del dataset, non possono essere opposte agli atti di utilizzazione qui oggetto di causa.

134 A ciò si aggiunge che il titolare dei diritti ha la possibilità di impedire l'utilizzazione mediante una riserva efficace ai sensi del § 44b, comma 3, UrhG e, con ciò, di evitare che l'opera o la fotografia venga utilizzata per l'addestramento dell'IA. Di tale possibilità, come esposto, il ricorrente non ha fatto uso in modo sufficiente.

135 (c) Al terzo stadio va verificata un'ingiustificata lesione degli interessi del titolare dei diritti. Se una tale lesione sussista va accertato mediante una valutazione di congruità da compiere per via di bilanciamento degli interessi e mediante l'accertamento se il bisogno dell'utilizzazione prevalga sul pregiudizio del titolare (cfr. BGH GRUR 2020, 859 Rn. 73 – Reformistischer Aufbruch II). In tale contesto vanno considerati, ad esempio, il carattere commerciale o non commerciale dell'utilizzazione, l'ampiezza della ripresa, il grado di trasformazione e un eventuale pericolo di confusione, la questione se sussista un confronto con l'opera o con altri temi, e il peso dei diritti fondamentali coinvolti (cfr. BGH GRUR 2023, 1531 Rn. 51 – Metall auf Metall V).

136 Secondo i considerando della direttiva DSM, per un verso lo sviluppo tecnologico non deve essere ostacolato, per altro verso deve però anche essere mantenuto un livello elevato di protezione del diritto d'autore e dei diritti connessi (considerando 3).

137 La riproduzione era, nel caso di specie, necessaria per eseguire il confronto immagine-testo. Gli interessi del convenuto alla riproduzione prevalgono. Va bensì considerato che, come esposto, l'utilizzazione ai sensi del § 44b UrhG è esente da remunerazione. Nondimeno, tenuto conto delle circostanze che le riproduzioni costituiscono processi puramente interni al convenuto, che il convenuto non svolge attività commerciale (v. infra, bb.) e che esisteva la possibilità di impedire l'utilizzazione oggetto di causa mediante una riserva efficace, un'ingiustificata lesione degli interessi del ricorrente quale titolare dei diritti è da escludere.

138 Nella misura in cui il ricorrente — nella memoria del 29.08.2024, comunque non autorizzata in primo grado e dunque irrilevante ai sensi del § 296a ZPO — ha fatto riferimento a concrete perdite di fatturato subite, non è riconoscibile che esse siano riconducibili alle prestazioni di modelli di IA addestrati con il dataset del convenuto e mediante l'impiego della fotografia oggetto di causa.

139 bb. L'utilizzazione oggetto di causa della fotografia è inoltre giustificata dall'eccezione del § 60d, comma 1, UrhG.

140 Ai sensi del § 60d, comma 1, UrhG sono lecite le riproduzioni per il text and data mining (§ 44b, commi 1 e 2, primo periodo, UrhG) per fini di ricerca scientifica secondo le disposizioni che seguono.

141 Ai sensi del § 60d, comma 2, primo periodo, UrhG sono legittimate alle riproduzioni le organizzazioni di

ricerca. Organizzazioni di ricerca sono, ai sensi del § 60d, comma 2, secondo periodo, UrhG, le università, gli istituti di ricerca o le altre istituzioni che svolgono ricerca scientifica, purché

142 1. non perseguano scopi commerciali,

143 2. reinvestano tutti gli utili nella ricerca scientifica, oppure

144 3. operino nell'ambito di un incarico riconosciuto dallo Stato nell'interesse pubblico.

145 Non sono legittimate ai sensi del primo periodo le organizzazioni di ricerca che collaborano con un'impresa privata che eserciti un'influenza determinante sull'organizzazione di ricerca e abbia un accesso privilegiato ai risultati della ricerca scientifica (§ 60d, comma 2, terzo periodo, UrhG).

146 Ai sensi del § 60d, comma 3, UrhG sono inoltre legittimati alle riproduzioni:

147 1. le biblioteche e i musei, purché pubblicamente accessibili, nonché gli archivi e le istituzioni nel settore del patrimonio cinematografico o sonoro (istituzioni di tutela del patrimonio culturale),

148 2. i singoli ricercatori, purché non perseguano scopi commerciali.

149 La nuova formulazione del § 60d UrhG è avvenuta soprattutto in attuazione dell'art. 3 della direttiva DSM (cfr. Dreier in Dreier/Schulze, UrhG, 8ª ed., § 60d Rn. 2; Anton in Spindler/Schuster/Kaesling, Recht der elektronischen Medien, 5ª ed., § 60d UrhG Rn. 2).

150 aaa. L'utilizzazione oggetto di causa è avvenuta per fini di ricerca scientifica.

151 Per ricerca scientifica si intende ogni attività metodica e sistematica avente lo scopo di acquisire nuove conoscenze in modo verificabile. Vi rientrano tanto l'attività di ricerca vera e propria quanto la presentazione dei suoi risultati (cfr. Dreier in Dreier/Schulze, UrhG, 8ª ed., § 60d Rn. 4 e § 60c Rn. 1; Leenen in Wandtke/Bullinger, UrhG, 6ª ed., art. 5 InfoSoc-RL Rn. 104; cfr. anche Jarass in Jarass GRCh, 4ª ed., art. 13 Rn. 8). 152 Nel considerando 12, secondo periodo, della direttiva DSM sono menzionate, è vero, le scienze naturali e le scienze umane, ma ciò non comporta una limitazione a determinate discipline scientifiche.

153 La scienza comprende non soltanto la ricerca di base, ma anche la ricerca applicata (cfr. Jarass in Jarass GRCh, 4ª ed., art. 13 Rn. 8; v. Coelln in Stern/Sodan/Möstl, StaatsR, 2ª ed., § 124 Rn. 147). Anche la ricerca tecnologica è ricerca scientifica (cfr. Ruffert in Calliess/Ruffert, EUV/AEUV, 6ª ed., art. 13 EU-GRCh Rn. 6; Thiele in Heselhaus/Nowak, EU-Grundrechte-HdB, 2ª ed., § 30 Rn. 1). La scientificità sussiste soprattutto in capo a scienziati e ricercatori, ma anche quando un pratico si dedica ad attività scientifica (cfr. Dreier in Dreier/Schulze, UrhG, 8ª ed., § 60c Rn. 1).

154 La ricerca applicata serve all'acquisizione di nuove conoscenze, ma è primariamente orientata a un obiettivo pratico concreto. Anche le attività di sviluppo, nel senso di una prima applicazione concretizzante di risultati di ricerca, possono ricadere nei privilegi previsti per la

ricerca, quando nel caso concreto ricerca e sviluppo siano strettamente connessi.

155 I confini tra ricerca e sviluppo sono fluidi (cfr. BVerfG NVwZ 2010, 1285 Rn. 43). È vero che il considerando 18 della direttiva DSM distingue tra l'impiego di procedimenti di text and data mining nel contesto della ricerca scientifica, da un lato, e l'impiego del text and data mining da parte di soggetti privati e pubblici — «anche per i servizi statali, per decisioni imprenditoriali complesse e per lo sviluppo di nuove applicazioni o tecnologie» — dall'altro. Da ciò non segue però che lo sviluppo di nuove applicazioni e tecnologie non possa costituire anche ricerca scientifica, quando in tale attività vengano acquisite nuove conoscenze in modo metodico, sistematico e verificabile. Così, ad esempio, ai sensi dell'art. 2, par. 8, del regolamento sull'IA, il regolamento non si applica alle attività di ricerca, prova e sviluppo relative a sistemi o modelli di IA prima della loro immissione sul mercato o messa in servizio (v. anche art. 2, par. 6, e art. 3, n. 63, del regolamento sull'IA, nonché considerando 25 del medesimo).

156 Secondo tali parametri, l'utilizzazione oggetto di causa è avvenuta a fini di ricerca scientifica.

157 (1) La creazione del dataset da parte del convenuto costituisce già di per sé ricerca scientifica.

158 Il convenuto ha estratto, con un determinato procedimento, il proprio dataset — composto da oltre 5 miliardi di URL dei rispettivi luoghi di memorizzazione delle immagini e di corrispondenti descrizioni delle immagini — da un preesistente dataset più ampio (di Common Crawl). A tal fine ha impiegato programmi specifici per il confronto immagine-testo (CLIP di OpenAI (p. 11 della sentenza di primo grado, p. 5 dell'allegato B 8)). La creazione del dataset da parte del convenuto non costituisce la semplice esecuzione di un procedimento consolidato: il convenuto — come ha illustrato il sig. [omissis] (membro del direttivo del convenuto) nell'udienza dinanzi al Collegio — ha applicato al dataset di Common Crawl il modello CLIP, pubblicato nel 2021, che consentiva un corrispondente filtraggio, generando così un dataset verificato di coppie immagine-testo. L'idoneità e la qualità del dataset generato dal convenuto — e dei suoi sottoinsiemi — sono state sottoposte a esperimenti di validazione (v. p. 10 della memoria del convenuto dell'11.04.2024, pp. 9 ss. della comparsa di risposta in appello). Il procedimento di generazione del dataset e i risultati degli esperimenti di validazione sono stati pubblicati in un articolo che indica come autori soprattutto collaboratrici e collaboratori del convenuto (allegato B 8). Il ricorrente ha, è vero, contestato per mancata conoscenza le allegazioni del convenuto sulle proprie pubblicazioni. Il Collegio non nutre però dubbi, alla luce delle copie prodotte delle pubblicazioni (allegati B 8 e B 9), sull'esistenza di tali pubblicazioni. Che gli allegati prodotti siano falsificazioni il ricorrente non ha fatto valere.

159 In una valutazione complessiva, già la creazione del dataset costituisce, nel caso di specie, un procedimento

metodico, diretto a un successivo acquisto di conoscenza e verificabile, riconducibile alla ricerca applicata.

160 (2) Il Landgericht ha inoltre correttamente osservato che anche un'utilizzazione compiuta in un passaggio (intermedio) di un processo di ricerca può avvenire «per fini di ricerca scientifica». Sono compresi non soltanto i passaggi di lavoro immediatamente connessi all'acquisizione di conoscenze; è anzi sufficiente che il passaggio di lavoro in questione (ad esempio una raccolta di dati) sia orientato a un (successivo) acquisto di conoscenza. Ciò trova espressione anche nella relazione ufficiale alla precedente versione del § 60d UrhG (BT-Drs. 18/12329, p. 40: «Il software rileva ad es. frequenze statistiche o correlazioni nei contenuti predisposti nel corpus. Questi risultati sono poi disponibili per l'analisi e la valutazione scientifica»).

161 Anche il successivo impiego del dataset per l'addestramento dell'intelligenza artificiale è da considerare ricerca scientifica. Lo sviluppo di nuovi modelli di IA e/o il perfezionamento di modelli esistenti con l'ausilio del dataset è da considerare ricerca scientifica applicata. Come esposto, anche le attività di sviluppo rientrano nel concetto di ricerca. Che anche imprese commerciali possano utilizzare i dati per le proprie attività di sviluppo non conduce a una diversa valutazione. Correttamente il Landgericht ha rilevato che anche la ricerca di imprese commerciali è pur sempre ricerca — ancorché non privilegiata in quanto tale ai sensi del § 60d UrhG.

162 (3) Il Landgericht ha parimenti accertato correttamente che, nel caso di specie, la creazione del dataset da parte del convenuto è avvenuta con l'obiettivo di impiegare il dataset a fini di successivo acquisto di conoscenza, ed era dunque orientata a un'utilizzazione scientifica del dataset. Ciò risulta da una valutazione complessiva dei seguenti indizi.

163 (a) Un indizio in tal senso è che, quantomeno nell'attuale statuto del convenuto, depositato nel registro delle associazioni pubblicamente consultabile (alla voce «visualizzazione documenti»), si legge che lo scopo dell'associazione risiede nella promozione della ricerca e dell'istruzione (v. anche allegato B 1). Le iscrizioni nei registri pubblici costituiscono fatti notori ai sensi del § 291 ZPO (cfr. Prütting in MüKoZPO, 78ª ed., § 291 Rn. 10; Bacher in BeckOK ZPO, 58ª ed., § 291 Rn. 5).

164 È da ritenere che il menzionato scopo associativo costituisca lo scopo statutario del convenuto già al momento della sua fondazione. Il ricorrente ha dapprima negato, in primo grado, che l'estratto prodotto dal convenuto quale allegato B 1 rappresentasse lo statuto vigente del convenuto o lo rappresentasse al momento dell'atto di riproduzione in questione (p. 3 della memoria di parte attrice del 20.09.2023). Dopo che il convenuto, con memoria dell'11.04.2024, producendo un estratto del verbale di fondazione, ha dedotto che l'assemblea costitutiva del convenuto si è tenuta il 07.07.2021 — e dunque prima dell'atto di utilizzazione oggetto di causa —, il ricorrente ha reso pacifica la cronologia di fondazione

dell'associazione e atto di riproduzione (p. 1 della memoria di parte attrice del 03.07.2024) e ha (soltanto) continuato a contestare che lo statuto prodotto quale allegato B 1 corrisponda allo statuto vigente (p. 16 della memoria di parte attrice del 03.07.2024). Lo statuto attuale è però, come esposto, depositato nel registro delle associazioni pubblicamente consultabile.

165 Che lo scopo associativo sia mutato dalla fondazione non è stato dedotto né è altrimenti ravvisabile.

166 (b) L'orientamento scientifico del lavoro del convenuto emerge anche dalle pubblicazioni prodotte dal convenuto (allegati B 8 e B 9), alle quali hanno partecipato collaboratori del convenuto.

167 Nel contributo in lingua inglese intitolato «[omissis]: An open large-scale dataset for training next generation image-text models» (pubblicato per la prima volta il 17 settembre 2022 e aggiornato da ultimo il 23 maggio 2023, allegato B 8) si descrive, tra l'altro, come il dataset è stato creato e come ne sono state validate idoneità e qualità.

168 Il menzionato contributo (allegato B 8) è stato citato, fino al 5 aprile 2024, complessivamente 1403 volte in altre opere scientifiche e ha ricevuto il «NeurIPS 2022 Outstanding Paper Award» della «Conference on Neural Information Processing Systems» («NeurIPS»), conferenza esistente dal 1987 nel settore del machine learning.

169 Oggetto del contributo prodotto quale allegato B 9 («Reproducible scaling laws for contrastive language-image learning»), pubblicato nell'ambito della conferenza «IEEE/CVF Conference on Computer Vision and Pattern Recognition», sono ulteriori indagini condotte mediante il dataset. Tale contributo tratta l'addestramento di openCLIP con il dataset [omissis], le sue funzioni e la derivazione delle cosiddette leggi di scala per diverse funzioni di tali modelli di base (pp. 11 s. della memoria del convenuto dell'11.04.2024). I risultati dell'indagine sono disponibili tramite il portale online di versionamento software «GitHub» (p. 12 della memoria del convenuto dell'11.04.2024).

170 Che le menzionate pubblicazioni siano state redatte soltanto in un momento successivo alla riproduzione oggetto di causa della fotografia del ricorrente non osta alla qualificazione della creazione del dataset come ricerca scientifica. La pubblicazione di risultati di ricerca presuppone un precedente lavoro scientifico.

171 (c) Inoltre, nel settembre 2023 il convenuto ha ricevuto il «Falling Walls Science Breakthrough of the Year 2023 Award» della Falling Walls Foundation. In un contributo raggiungibile tramite un link indicato dal convenuto (p. 12 della memoria dell'11.04.2024) si legge che il premio è stato conferito nella categoria «Science and Innovation Management» e che tale riconoscimento sottolinea l'«impegno decisivo per una ricerca sull'IA democratizzata» di [omissis] (v. anche pp. 17 s. della memoria di parte attrice del 03.07.2024).

172 (d) Il convenuto non soltanto mette gratuitamente a disposizione il dataset, ma rende anche trasparente su internet come i dataset di addestramento sono stati creati. Alla qualificazione quale utilizzazione a fini di

ricerca scientifica non osta che il dataset possa essere utilizzato anche da fornitori commerciali e che collaboratori del convenuto lavorino anche per imprese commerciali del settore dell'IA. All'approccio open source perseguito dal convenuto è immanente che i risultati del suo lavoro siano a disposizione di chiunque e dunque anche dei fornitori commerciali.

173 Come esposto, anche lo sviluppo (o il perfezionamento) di modelli di IA da parte di imprese commerciali costituisce ricerca applicata.

174 (e) Anche il disclaimer, richiamato dallo stesso ricorrente, contenuto in un articolo del convenuto (riprodotto alle pp. 13 s. della memoria di parte attrice del 20.09.2023) depone per un orientamento primariamente scientifico del lavoro del convenuto. All'inizio del disclaimer si legge:

175 «The motivation behind dataset creation is to democratize research and experimentation around large-scale multi-modal model training and handling of uncurated, large-scale datasets crawled from publically available internet. Our recommendation is therefore to use the dataset for research purposes.»

176 (in traduzione: La motivazione alla base della creazione del dataset consiste nel democratizzare la ricerca e la sperimentazione nell'ambito dell'addestramento di modelli multimodali su larga scala e della gestione di dataset non curati, su larga scala, raccolti dall'internet pubblicamente accessibile. La nostra raccomandazione è pertanto di utilizzare il dataset per scopi di ricerca).

177 (f) Infine, ulteriore indizio di un orientamento scientifico del lavoro del convenuto è la circostanza che il membro del direttivo e fondatore Dr. [omissis] è attivo dal 2017 quale responsabile di laboratorio presso il Juelich Supercomputing Center del Forschungszentrum Jülich. La più ampia deduzione del convenuto, secondo cui le persone fisiche appartenenti al convenuto — i membri del direttivo e gli altri soci — sarebbero ricercatori, è stata bensì contestata dal ricorrente. Il ricorrente non ha però contestato la deduzione del convenuto sull'attività del sig. [omissis] quale responsabile di laboratorio presso il Juelich Supercomputing Center del Forschungszentrum Jülich (v. p. 18 della memoria di parte attrice del 03.07.2024).

178 (g) L'obiezione del ricorrente secondo cui il dataset avrebbe contenuto contenuti penalmente rilevanti (p. 10 della memoria attorea del 03.07.2024, con rinvio a una notizia su heise.de) non osta alla qualificazione della riproduzione della fotografia oggetto di causa quale utilizzazione a fini scientifici. Nella notizia di heise.de citata dal ricorrente si leggeva che alcuni ricercatori avevano rilevato che il dataset [omissis] conteneva complessivamente 1679 link a immagini con contenuto illegale. Il dataset comprende complessivamente oltre 5 miliardi di URL dei rispettivi luoghi di memorizzazione delle immagini e le corrispondenti descrizioni. Sul proprio sito il convenuto ha segnalato che il dataset non è curato e può contenere contenuti indesiderati, c.d. contenuti NSFW (NSFW = «not safe for work»). È vero che gli URL

che conducevano a contenuti penalmente rilevanti non avrebbero dovuto essere inclusi; tale condotta scorretta non fa però venir meno lo scopo scientifico della creazione del dataset nel suo complesso (cfr. Kempen/Rossa in BeckOK GG, 63ª ed., art. 5 Rn. 180.1; Krüper in Dreier GG, 4ª ed., art. 5, comma 3, Rn. 61).

179 (4) Non rileva, in via integrativa, se il convenuto — come da esso dedotto, tra l'altro con riferimento all'articolo pubblicato di cui all'allegato B 9 — abbia utilizzato il dataset anche per addestrare propri modelli di IA.

180 bbb. Il convenuto rientra anche tra i legittimati ai sensi del § 60d, comma 2, UrhG. Si tratta di un'organizzazione di ricerca ai sensi del § 60d, comma 2, UrhG.

181 (1) Il convenuto non è un'università né un istituto di ricerca. Il convenuto è però un'altra istituzione ai sensi del § 60d, comma 2, secondo periodo, UrhG, che svolge ricerca scientifica.

182 Secondo l'inequivoco tenore letterale del § 60d, comma 2, secondo periodo, UrhG, l'istituzione, per godere del privilegio del § 60d UrhG, deve svolgere essa stessa ricerca. Non diversamente risulta dalla definizione legale della direttiva DSM. Anche se l'art. 2, n. 1, della direttiva DSM fa riferimento al fatto che l'obiettivo primario dell'istituzione sia la ricerca scientifica o l'attività didattica, ciò non va inteso nel senso che l'istituzione non debba svolgere essa stessa ricerca e che basti che la sua attività sia orientata a servire mediatamente la ricerca. Lo mostrano, per un verso, il confronto con le istituzioni concretamente menzionate nell'art. 2, n. 1, della direttiva DSM — «università, comprese le relative biblioteche», «istituto di ricerca» — e, per altro verso, lo sguardo alla versione inglese («the primary goal of which is to conduct scientific research») e a quella francese («ayant pour objectif premier de mener des recherches scientifiques»).

183 La ricerca che le persone attive per l'istituzione svolgono nell'ambito della loro attività per l'istituzione è imputata all'istituzione (cfr. Anton in Spindler/Schuster/Kaesling, Recht der elektronischen Medien, 5ª ed., § 60d UrhG Rn. 19). Inoltre, il privilegio dei singoli ricercatori risulta dal § 60d, comma 3, n. 2, UrhG.

184 Se, come esposto, un'istituzione deve svolgere essa stessa ricerca, non basta che l'istituzione si limiti a creare materiale che possa essere successivamente utilizzato a fini di ricerca (cfr. Anton in Spindler/Schuster/Kaesling, Recht der elektronischen Medien, 5ª ed., § 60d UrhG Rn. 28). È vero che le attività preparatorie e di supporto alla ricerca sono protette dalla libertà della scienza ai sensi dell'art. 13 della Carta dei diritti fondamentali dell'UE (cfr. Jarass in Jarass GRCh, 4ª ed., art. 13 Rn. 8); una propria ricerca scientifica ai sensi del § 60d, comma 2, secondo periodo, UrhG non risiede però ancora, senz'altro, in siffatte attività puramente preparatorie o di supporto. Può rimanere aperta la questione se l'eccezione copra anche l'attività di prestatori di servizi esterni incaricati da un'organizzazione di ricerca della creazione del dataset o dell'esecuzione tecnica del text and data mining (cfr. Bomhard in BeckOK UrhR, § 60d UrhG Rn. 21, 23; in senso

simile anche Stieper in Schricker/Loewenheim, UrhG, § 60d Rn. 14 (sul § 60d, vecchia versione)). Un siffatto rapporto di incarico non sussiste nel caso di specie.

185 Come esposto, deve muoversi da un orientamento primariamente scientifico delle attività del convenuto. Il convenuto svolge anche esso stesso ricerca scientifica.

186 Secondo il suo statuto, il convenuto è, come esposto, un'associazione per la promozione e la diffusione della ricerca e dell'istruzione libere. Scopo dell'associazione è la promozione della ricerca e dell'istruzione. Tra le attività elencate nello statuto, mediante le quali lo scopo statutario viene «in particolare» realizzato, la ricerca propria del convenuto non è (espressamente) menzionata. È menzionata però la «creazione dell'infrastruttura per la democratizzazione della generazione e dell'utilizzazione di grandi modelli di rete di IA». Vi rientrano, secondo lo statuto, tra l'altro, la «messa a disposizione di computer e risorse di memoria, l'apertura del codice sorgente (open source), l'apertura dei modelli appresi (open models) e dei dati utilizzati per l'apprendimento (open data), nonché offerte formative per informare il grande pubblico su potenzialità, pericoli e utilizzo di tali sistemi capaci di apprendere». Anche l'«archiviazione e messa a disposizione di dataset che possano essere utili a ricercatori e utenti IT nel perseguimento degli obiettivi predetti» rientra, secondo lo statuto, tra le attività del convenuto.

187 Come esposto, la creazione del dataset costituisce ricerca scientifica propria del convenuto. Non rileva, anche a questo riguardo, in via integrativa, se il convenuto — come da esso dedotto, tra l'altro con riferimento all'articolo pubblicato di cui all'allegato B 9 — abbia utilizzato il dataset anche per addestrare propri modelli di IA.

188 (2) Il convenuto non persegue scopi commerciali (§ 60d, comma 2, secondo periodo, n. 1, UrhG). Ai sensi dell'art. 2, n. 1, lett. a, della direttiva DSM, il perseguimento di scopi commerciali va inteso nel senso di un orientamento al profitto (cfr. Dreier in Dreier/Schulze, UrhG, 8ª ed., § 60d Rn. 5).

189 Il convenuto è, secondo il suo statuto, ente senza scopo di lucro. I soci non ricevono attribuzioni dai mezzi dell'associazione.

190 Come esposto, il convenuto mette a disposizione i propri risultati gratuitamente. Che il convenuto si finanzi mediante donazioni (p. 15 della memoria del convenuto dell'11.04.2024) non osta all'orientamento non commerciale (cfr. Anton in Spindler/Schuster/Kaesling, Recht der elektronischen Medien, 5ª ed., § 60c UrhG Rn. 4, sulla ricerca finanziata con fondi privati di terzi; cfr. anche Grübler in BeckOK UrhR, 47ª ed., § 60c UrhG Rn. 6; Dreier in Dreier/Schulze, UrhG, 8ª ed., § 60c Rn. 6). Ciò varrebbe anche nel caso in cui l'impresa [omissis] Al avesse concesso mezzi finanziari al convenuto.

191 Neppure dal rapporto del convenuto con altre imprese commerciali risulta una diversa valutazione. Il ricorrente ha dedotto, con riferimento a una chat, che sarebbe stato concordato di concedere all'impresa

[omissis], in ragione del suo pagamento di 5.000 dollari, un accesso privilegiato al dataset (v. pp. 2 ss. della memoria di parte attrice del 03.07.2024). A ciò il convenuto — contrariamente a quanto dedotto a p. 13 della comparsa di risposta in appello — non ha opposto in primo grado contestazioni rilevanti. Tale episodio isolato non basta però — anche in considerazione della somma menzionata — a mettere in dubbio il perseguimento di scopi non commerciali. Né tale episodio è da valutare come vendita di un bene o di un servizio, né da esso segue un orientamento al profitto del convenuto.

192 Anche la circostanza che pure fornitori commerciali possano utilizzare il dataset non conduce a un diverso risultato. Ai sensi del § 60d, comma 2, UrhG, rileva se l'organizzazione di ricerca persegua scopi commerciali, e non se la ricerca vada a beneficio anche di scopi commerciali di terzi. Poiché il convenuto mette il proprio dataset a disposizione della collettività, sicché tanto le istituzioni di ricerca non commerciali quanto le imprese commerciali possono utilizzarlo, la sua attività non è orientata al profitto.

193 La disposizione del § 60d UrhG non intende impedire una normale utilizzazione a valle dei risultati della ricerca da parte dell'economia (cfr. Anton in Spindler/Schuster/Kaesling, UrhG, 5ª ed., § 60d Rn. 28). Come risulta dal § 60d, comma 2, terzo periodo, UrhG, neppure ogni collaborazione di un'organizzazione di ricerca con un'impresa privata comporta la perdita del privilegio di cui al § 60d UrhG.

194 Un perseguimento di scopi commerciali non segue, nel caso di specie, neppure dal fatto che, secondo le stesse deduzioni del convenuto, è intenzione mettere il proprio lavoro liberamente a disposizione, oltre che di (altri) ricercatori, anche di imprese in Germania e nell'UE, per contrastare il pericolo che imprese e istituzioni con sede in Germania e nell'UE vengano rapidamente distanziate nel settore dello sviluppo dell'IA (p. 3 della comparsa di risposta, p. 14 della memoria del convenuto dell'11.04.2024). Considerato che il convenuto non persegue utili propri e che il dataset non è messo a disposizione di singole imprese, bensì del pubblico, anche sotto questo profilo deve ritenersi un orientamento non commerciale del convenuto.

195 (3) Nel caso di specie non ricorre neppure una causa di esclusione ai sensi del § 60d, comma 2, terzo periodo, UrhG. In base ad esso, non sono legittimate alle riproduzioni le organizzazioni di ricerca che collaborano con un'impresa privata che eserciti un'influenza determinante sull'organizzazione di ricerca e abbia un accesso privilegiato ai risultati della ricerca scientifica. Tale disciplina risale all'art. 2, n. 1, della direttiva DSM, secondo cui si è in presenza di un'organizzazione di ricerca, tra l'altro, soltanto quando nessuna impresa che eserciti un'influenza determinante su tale organizzazione ottenga un accesso privilegiato ai risultati della ricerca scientifica. Secondo la disciplina del § 60d, comma 2, terzo periodo, UrhG, una c.d. public private partnership è esclusa dal privilegio quando l'impresa privata,

cumulativamente, (i) esercita un'influenza determinante sull'organizzazione di ricerca e (ii) ottiene un accesso privilegiato ai risultati della ricerca (cfr. Bomhard in BeckOK UrhR, 47^a ed., § 60d UrhG Rn. 22; BT-Drs. 19/27426, p. 96). Indicatori di un'influenza determinante sono, ad esempio, la possibilità per l'impresa privata di esercitare un controllo unilaterale sulle condizioni contenutistiche e strutturali del progetto di ricerca e sull'ampiezza dell'oggetto della ricerca (cfr. Bomhard, op. cit., Rn. 22; cfr. anche Anton in Spindler/Schuster/Kaesling, *Recht der elektronischen Medien*, 5^a ed., § 60d UrhG Rn. 21). Nel considerando 12 della direttiva DSM si legge che le organizzazioni non dovrebbero essere considerate organizzazioni di ricerca ai sensi della direttiva quando siano soggette all'influenza determinante di imprese commerciali che, in ragione degli assetti strutturali — ad esempio nella loro qualità di azionisti o soci —, possano esercitare un controllo e ottenere così un accesso privilegiato ai risultati della ricerca.

196 Mediante la disciplina del § 60d, comma 2, terzo periodo, UrhG si contrasta, tra l'altro, un'ingiustificata fruizione del privilegio della ricerca non commerciale perseguito dal § 60d UrhG mediante la mera esternalizzazione dell'attività di ricerca di un'impresa commerciale (cfr. al riguardo Anton, op. cit., Rn. 28).

197 (a) Può rimanere aperta la questione se l'onere della prova dei presupposti del § 60d, comma 2, terzo periodo, UrhG gravi sul titolare dei diritti o su chi invoca l'eccezione. Il § 60d, comma 2, terzo periodo, UrhG è formulato dal legislatore tedesco come eccezione al privilegio e non come fatto negativo da provare a cura dell'istituzione di ricerca, il che depone nel senso di porre l'onere della prova, sotto questo profilo, in capo al titolare dei diritti. Può rimanere aperta la questione se, considerata la formulazione dell'art. 2, n. 1, della direttiva DSM («wobei kein Unternehmen, das einen bestimmenden Einfluss auf diese Organisation hat, bevorzugten Zugang zu den Ergebnissen der wissenschaftlichen Forschung erhält»; nella versione inglese: «in such a way that the access to the results generated by such scientific research cannot be enjoyed on a preferential basis by an undertaking that exercises a decisive influence upon such organisation»), occorra un'interpretazione del § 60d, comma 2, terzo periodo, UrhG conforme al diritto dell'Unione nel senso di un fatto negativo quale presupposto del privilegio, per il quale porta l'onere della prova chi invoca l'eccezione. In ogni caso, il titolare dei diritti deve indicare elementi da cui risulti un'influenza determinante di un'impresa sull'istituzione di ricerca. Sulla controparte processuale della parte gravata dalla prova di un fatto negativo incombe un onere di allegazione secondario, la cui ampiezza dipende dalle circostanze del caso concreto (cfr. BGH, ord. 20.06.2017 – VI ZR 505/16, BeckRS 2017, 117730 Rn. 4). Poiché il ricorrente non ha alcuna visibilità sul rapporto del convenuto con le imprese private, sul ricorrente grava, nel caso di specie, un onere di

allegazione soltanto in misura molto ridotta. Egli adempie il proprio onere di allegazione indicando elementi da cui risulti un'influenza determinante di un'impresa. Trattandosi di vicende della sfera del convenuto, spetta poi al convenuto esporre perché, nonostante gli elementi indicati dal ricorrente, non sussista un'influenza determinante di un'impresa.

198 (b) Come esposto, il convenuto non ha contestato in modo rilevante, in primo grado, le allegazioni del ricorrente sull'accesso privilegiato di [omissis] al dataset. Ragioni per la contestazione tardiva in appello ai sensi del § 531, comma 2, ZPO non sono ravvisabili. A ciò si aggiunge che il convenuto, nella comparsa di risposta in appello (ivi p. 13), rileva bensì che nella chat non si sarebbe fatto riferimento al dataset oggetto di causa [omissis], bensì a un dataset del tutto diverso («100M+»). Il convenuto non ha però esposto di quale dataset si tratti con «100M+» e se esso non costituisca una versione preliminare di [omissis].

199 Ma anche nel caso in cui il convenuto si sia dichiarato disposto a concedere all'impresa [omissis], in ragione del suo pagamento di 5.000 dollari, un accesso privilegiato al dataset, e un siffatto accesso privilegiato sotto il profilo temporale sia stato eventualmente anche concesso, manca — contrariamente all'avviso del ricorrente — il secondo presupposto dell'eccezione di cui al § 60d, comma 2, terzo periodo, UrhG, ossia l'influenza determinante dell'impresa privata. Il solo pagamento di 5.000 dollari non apre all'impresa la possibilità di esercitare un controllo unilaterale sulle condizioni contenutistiche e strutturali del progetto di ricerca e sull'ampiezza dell'oggetto della ricerca.

200 (c) Anche con riguardo all'impresa [omissis] Al manca un'influenza determinante. Il convenuto ha ammesso che l'impresa [omissis] Al gli ha messo a disposizione risorse di calcolo. Da ciò non è però desumibile un'influenza determinante sul lavoro del convenuto. Che il convenuto abbia ricevuto un sostegno finanziario in denaro da [omissis] Al è stato contestato dal convenuto. Può rimanere aperta la questione se dall'intervista con il fondatore e amministratore di [omissis] Al richiamata dal ricorrente (p. 3 della memoria di parte attrice del 20.09.2023, con la citazione «[...] says that [omissis] Al funded the creation of [omissis] [...]») risultino elementi di un sostegno ulteriore rispetto alla messa a disposizione di risorse di calcolo. Neppure un sostegno finanziario è di per sé equiparabile a un'influenza determinante. La sola circostanza di un finanziamento con mezzi privati non fa venir meno il privilegio (cfr. Anton in Spindler/Schuster/Kaesling, *Recht der elektronischen Medien*, 5^a ed., § 60d UrhG Rn. 20; Bullinger in Wandtke/Bullinger, UrhG, 6^a ed., § 60d Rn. 25; cfr. anche Dreier in Dreier/Schulze, UrhG, § 60c Rn. 6; Grübler in BeckOK UrhR, 47^a ed., § 60c UrhG Rn. 6).

201 Che il dataset sia stato utilizzato per l'addestramento del generatore testo-immagine «Stable Diffusion» non conduce a un diverso risultato. La possibilità che anche imprese commerciali utilizzino il dataset è conseguenza

del fatto che il convenuto lo ha messo a disposizione della collettività.

202 (d) Neppure una collaborazione con il gruppo di ricerca Stable Horde in un progetto di valutazione estetica di dataset (p. 4 dell'atto di citazione, pp. 8 s. della comparsa di risposta, p. 3 della memoria di parte attrice del 20.09.2023) determina un'influenza determinante dell'impresa privata [omissis] AI sul convenuto.

203 (e) Infine, un'influenza determinante non può desumersi neppure dal fatto che singoli soci del convenuto siano o siano stati dipendenti di [omissis] AI e di altre imprese. Tali persone non sono i rappresentanti legali di dette imprese. Il membro del direttivo del convenuto [omissis] è dal febbraio 2022 attivo per l'impresa [omissis] AI come research engineer (quale «Head of Machine Learning Operations», p. 4 della memoria di parte attrice del 20.09.2023 e p. 16 della memoria di parte convenuta dell'11.04.2024). [omissis] è socio ordinario del convenuto. Presso [omissis] AI era impiegato come ricercatore. Il sig. [omissis], parimenti socio del convenuto, è senior machine learning engineer presso Google (p. 6 della memoria di parte attrice del 20.09.2023).

204 (4) L'esecuzione del test dei tre stadi dà come risultato un'utilizzazione lecita.

205 (a) Al primo requisito del test dei tre stadi, da eseguire nel presente caso concreto, è soddisfatto mediante l'introduzione dell'eccezione nel § 60d UrhG, trattandosi di un caso speciale disciplinato per legge.

206 (b) Nella verifica da compiere al secondo stadio — se lo sfruttamento normale dell'opera o dell'altro materiale protetto non sia pregiudicato — occorre considerare, come esposto, se mediante l'utilizzazione contestata siano ridotte le possibilità del titolare dei diritti di collocamento legittimo (cfr. CGUE GRUR 2017, 610 Rn. 70 – Stichting Brein; BGH GRUR 2023, 1531 Rn. 49 – Metall auf Metall V). Rileva al riguardo se l'utilizzazione in questione entri in concorrenza diretta con l'utilizzazione tradizionale, ossia se si interferisca nello sfruttamento primario (cfr. BGH GRUR 2020, 859 Rn. 72 – Reformistischer Aufbruch II).

207 Nel caso di specie la riproduzione oggetto di causa costituisce un processo puramente interno al convenuto. Nel dataset reso disponibile all'esterno la riproduzione non è contenuta; vi è soltanto un link alla fonte legittima. Che il ricorrente non riceva denaro per l'utilizzazione privilegiata ai sensi del § 60d UrhG è il risultato della scelta legislativa di esentare da obbligo di remunerazione le utilizzazioni che ricadono nell'eccezione (§ 60h, comma 2, n. 3, UrhG).

208 Anche considerando che il dataset serve all'addestramento di IA generativa e che, in un passaggio ulteriore, l'intelligenza artificiale genera immagini che possono entrare in concorrenza con le fotografie del ricorrente, il requisito del secondo stadio è, come esposto, soddisfatto.

209 (c) Al terzo stadio va verificata un'ingiustificata lesione degli interessi del titolare dei diritti. Se una tale lesione

sussista va accertato mediante una valutazione di congruità da compiere per via di bilanciamento degli interessi e mediante l'accertamento se il bisogno dell'utilizzazione prevalga sul pregiudizio del titolare (cfr. BGH GRUR 2020, 859 Rn. 73 – Reformistischer Aufbruch II). In tale contesto vanno considerati, ad esempio, il carattere commerciale o non commerciale dell'utilizzazione, l'ampiezza della ripresa, il grado di trasformazione e un eventuale pericolo di confusione, la questione se sussista un confronto con l'opera o con altri temi, e il peso dei diritti fondamentali coinvolti (cfr. BGH GRUR 2023, 1531 Rn. 51 – Metall auf Metall V).

210 La riproduzione era, nel caso di specie, necessaria per eseguire il confronto immagine-testo. Gli interessi del convenuto alla riproduzione prevalgono. Va bensì considerato che, come esposto, l'utilizzazione ai sensi del § 60d UrhG è esente da remunerazione (§ 60h, comma 2, n. 3, UrhG) e che, a differenza del § 44b UrhG, non esiste neppure la possibilità di una riserva. Nondimeno, tenuto conto dello scopo scientifico dell'utilizzazione (art. 13 della Carta dei diritti fondamentali dell'UE), dell'utilizzazione non commerciale da parte del convenuto e della circostanza che le riproduzioni costituiscono processi puramente interni al convenuto, un'ingiustificata lesione degli interessi del ricorrente quale titolare dei diritti (art. 17, par. 2, della Carta) è da escludere.

211 cc. Non è decisivo se operi inoltre la disciplina del § 44a UrhG. Il Landgericht ha però correttamente ritenuto che la riproduzione non sia coperta dal § 44a UrhG. Sul punto si rinvia alle corrette considerazioni della sentenza impugnata. In particolare, il download mirato di un file di immagine al fine di analizzarlo mediante un software specifico non costituisce una riproduzione né transitoria né (puramente) accessoria.

212 Può altresì rimanere aperta la questione se debba ritenersi un semplice consenso (schlichte Einwilligung) del ricorrente all'atto di utilizzazione oggetto di causa.

213 3. La decisione sulle spese si fonda sul § 97, comma 1, ZPO. La statuizione sulla provvisoria esecutorietà si basa sui §§ 708 n. 10 e 711 ZPO.

214 4. La Revision è ammessa ai sensi del § 543, comma 2, ZPO. La controversia ha importanza fondamentale e lo sviluppo del diritto esige una decisione del giudice di legittimità. Sull'interpretazione dei §§ 44b e 60d UrhG in relazione a un'utilizzazione di opere protette dall'UrhG o di diritti connessi a monte dell'addestramento dell'intelligenza artificiale non esiste giurisprudenza di legittimità. Considerato il crescente impiego dell'intelligenza artificiale, le questioni giuridiche che si pongono nel presente caso hanno un rilievo che trascende il caso concreto.

**LA MEMORIZATION COME RIPRODUZIONE E I LIMITI
DELL'ECCEZIONE PER TEXT AND DATA MINING
GEMA v. OPENAI - LG MÜNCHEN I, 11 NOVEMBRE
2025, 42 O 14139/24**

di Sara Manni

I. Premesse

La sentenza pronunciata dal *Landgericht München I* (corte regionale di Monaco) l'11 novembre 2025 nel procedimento n. 42 O 14139/24, *GEMA v. OpenAI*, costituisce, allo stato, precedente di grande rilievo sull'applicabilità delle eccezioni di *text and data mining* ai modelli di IA generativa.

Parte attrice è GEMA, società tedesca di *collecting* che amministra i diritti d'autore di compositori, autori di testi ed editori musicali in Germania. La GEMA è legittimata, in qualità di rappresentante dei propri associati, a far valere in giudizio i diritti esclusivi di riproduzione, di rappresentazione e di comunicazione al pubblico - fra l'altro - relativi alle opere musicali del proprio repertorio.

Ad essere convenute sono invece due società del gruppo OpenAI, OpenAI, L.L.C. e OpenAI Ireland Limited, responsabili, rispettivamente, dello sviluppo dei modelli linguistici *GPT-4* e *GPT-4o* e dell'offerta dei medesimi sul territorio europeo.

La pronuncia, si precisa, ha riguardato esclusivamente i modelli *GPT-4* e *GPT-4o*; non, dunque, le versioni successive - segnatamente *GPT-5* - già disponibili sul mercato all'epoca della sentenza.

La pronuncia, per un verso, interviene su un profilo - quello della c.d. «*memorization*» dei contenuti protetti nei parametri del modello - che era rimasto sostanzialmente sullo sfondo nelle precedenti decisioni (compresa la sentenza inglese *Getty Images v. Stability AI* del 4 novembre 2025)²⁸²; per l'altro, offre una sistemazione tendenzialmente restrittiva delle eccezioni di *text and data mining* applicate ai modelli generativi, in evidente contrappunto rispetto alla pronuncia, di pochi mesi

precedente, del *Landgericht Hamburg* - confermata in appello - nel caso *Kneschke v. LAION*.

Il *Landgericht München I* ha accolto in misura prevalente le domande proposte dalla *Gesellschaft für musikalische Aufführungs- und mechanische Vervielfältigungsrechte* (GEMA), società di *collecting* tedesca per i diritti musicali, accertando che la memorizzazione e la riproduzione, da parte dei modelli *GPT-4* e *GPT-4o*, di testi musicali appartenenti al repertorio GEMA configurassero violazione del diritto esclusivo di riproduzione di cui al § 16 UrhG e del diritto di comunicazione al pubblico di cui al § 19a UrhG, escludendo che tali fattispecie potessero essere ricondotte tanto alla eccezione generale di *text and data mining* di cui al § 44b UrhG, quanto a quella per finalità scientifiche di cui al § 60d UrhG.

È stata invece rigettata la domanda connessa alla violazione dei diritti della personalità degli autori.

II. L'origine della controversia e la posizione di GEMA

L'azione è stata depositata dinanzi al *Landgericht München I* nel novembre 2024, in esito a una serie di rilevazioni effettuate dalla stessa GEMA mediante *sampling* sistematico degli *output* generati da ChatGPT. Il *thema decidendum* ha avuto ad oggetto, in particolare, i testi musicali di nove note canzoni del repertorio tedesco - fra le quali figurano composizioni di Herbert Grönemeyer («*Männer*»), Reinhard Mey («*Über den Wolken*»), Kristina Bach («*Atemlos*») e Rolf Zuckowski («*Wie schön, dass du geboren bist*») - opere note al pubblico tedesco.

Le evidenze prodotte dalla GEMA documentavano che, in risposta a semplici *prompt* formulati in linguaggio naturale, ChatGPT - operato sulla base dei modelli *GPT-4* e *GPT-4o* - restituiva agli utenti testi musicali sostanzialmente coincidenti con quelli delle

²⁸² Per un'analisi dottrinale, comparatistica e internazionalistica della pronuncia, che individua proprio nella *memorization* lo snodo decisivo dell'intera costruzione argomentativa del Tribunale, v. anche T.W.

DORNIS, J.C. GINSBURG, N. LUCCHI, *GEMA v. OpenAI (Munich I Regional Court, 2025): Doctrinal, Comparative, and International Perspectives on Copyright and Generative AI*, in *JuristenZeitung* (JZ), 81, 2026, 235 ss.

opere originali, talora identici nella loro integralità, talora con minime variazioni nelle parti iniziali e conclusive del testo, ovvero in frammenti interscambiabili, che lasciavano però inalterato il nucleo lirico-espressivo dell'opera. La rilevanza di tali variazioni è stata oggetto, come si vedrà, di specifica valutazione da parte del giudice.

L'attrice GEMA ha articolato la domanda su due piani, contestando l'*infringement* sia a livello di input, sia a livello di output.

Sotto il profilo del c.d. «*input infringement*», la *collecting* ha lamentato la riproduzione illecita dei testi musicali nella fase di addestramento del modello, senza autorizzazione e senza corresponsione di alcun compenso.

Invece, con riguardo al c.d. «*output infringement*», ha denunciato la successiva riproduzione delle opere nei *response* generati dal modello in risposta agli *user prompt*, deducendone la qualificazione quale ulteriore atto di riproduzione e di comunicazione al pubblico ai sensi rispettivamente dei §§ 16 e 19a UrhG.

Una particolarità del procedimento merita di essere segnalata sin d'ora.

In via preventiva, nel 2023 (prima ancora che l'AI Act ne sancisse l'applicabilità e ne demandasse fermamente il rispetto), la GEMA aveva formulato - quale società di *collecting* - una riserva di *opt-out* generale ai sensi dell'art. 4, par. 3, della direttiva DSM e del corrispondente § 44b, comma 3, UrhG, dichiarando di non autorizzare l'utilizzazione dei testi musicali del proprio repertorio per finalità di *text and data mining*. La validità e l'opponibilità di tale riserva hanno costituito uno dei profili centrali del giudizio²⁸³.

III. Le difese di OpenAI

La convenuta ha articolato la propria difesa intorno a tre nuclei argomentativi.

Sotto un primo profilo, OpenAI ha sostenuto che i modelli GPT non operano quali archivi di testi memorizzati, né conservano copie recuperabili

degli articoli o delle opere utilizzate per il *training*: gli *output* verrebbero generati attraverso processi probabilistici fondati sull'apprendimento di *pattern* linguistici astratti, sì che eventuali somiglianze fra *output* e opere originali deriverebbero da correlazioni statistiche e non da duplicazioni in senso tecnico (l'argomento appare pressoché sovrapponibile a quanto sostenuto da Anthropic nel contenzioso *Concord*, cfr. *retro*, *Il contenzioso Anthropic: uso trasformativo, pirateria e guardrails*).

Sotto un secondo profilo, la convenuta ha invocato l'operatività delle eccezioni di *text and data mining* introdotte dalla direttiva DSM e dai §§ 44b e 60d UrhG, sostenendo che il *training* dei modelli - comprensivo della fase di acquisizione del materiale, di quella di analisi statistica e di quella di *fine-tuning* - rientrerebbe integralmente nel perimetro applicativo di detta eccezione.

Sotto un terzo profilo, infine, OpenAI ha contestato l'opponibilità della riserva di *opt-out* formulata dalla GEMA, sostenendone l'inadeguata formulazione sotto il profilo della *machine-readability* richiesta dall'art. 4, par. 3, della direttiva.

IV. La decisione del Landgericht München I

Il *Landgericht München I* ha respinto *in toto* le difese della convenuta OpenAI sul piano della qualificazione tecnico-giuridica del modello.

Il giudice ha accertato, sulla base di un'analitica disamina degli *output* prodotti dal modello in risposta agli *user prompt*, che i testi musicali oggetto della controversia risultavano memorizzati «*nei parametri del modello linguistico*» in misura tale da consentire al sistema di rigenerarli, su semplice sollecitazione dell'utente, in forma sostanzialmente integrale. La complessità e l'estensione dei testi rinvenuti negli *output*, combinate con la loro sistematica riproducibilità a fronte di *prompt* anche minimi, hanno consentito al giudice di escludere il carattere casuale del fenomeno.

²⁸³ Sull'eccezione per TDM nell'AI Act si rimanda a J. P. QUINTAIS, *Generative AI, copyright and the AI Act*, in *Computer Law and Security Review: The International*

Journal of Technology Law and Practice, 56, 2025, 1-17, disponibile online in Elsevier.

Quanto alla rilevanza alle minime variazioni rilevate in alcuni *output* (costituenti allucinazioni rispetto al testo della canzone, più che intenzionali modifiche volte ad evitare l'illecita riproduzione di contenuti altrui) il giudice ha osservato che esse non incidono in alcun modo sulla riconoscibilità delle opere originali nella loro peculiarità espressiva, in quanto si limitavano a porzioni interscambiabili del testo (incipit, conclusioni), e non al nucleo lirico-espressivo della composizione. Sotto questo profilo, la pronuncia richiama implicitamente le riflessioni svolte dalla Corte di giustizia dell'Unione europea nella sentenza *Infopaq International*, C-5/08, in tema di soglia minima di originalità rilevante ai fini della tutela autoriale.

Sul piano della qualificazione giuridica, il Tribunale ha affermato che la memorizzazione delle opere nei parametri del modello costituisca essa stessa atto di riproduzione ai sensi del § 16 UrhG, attuativo dell'art. 2 della direttiva 2001/29/CE²⁸⁴.

Il passaggio, a ben vedere, è di considerevole rilievo: la pronuncia, infatti, rigetta espressamente l'impostazione secondo cui i *model weights* - in quanto parametri numerici astratti - sfuggirebbero al perimetro della nozione di "riproduzione". Nella prospettiva del Tribunale, il dato decisivo non è la natura tecnica del supporto, bensì la possibilità che il contenuto protetto sia, di fatto, *fissato* in forma idonea a consentirne la successiva riproduzione. Il *Landgericht* osserva, in proposito, che ai fini della nozione di riproduzione le modalità tecniche della memorizzazione siano ininfluenti, rilevando piuttosto l'*idoneità* del supporto a rendere disponibile, anche solo potenzialmente, l'opera protetta. La conclusione appare, sotto questo profilo, molto distante rispetto a quella raggiunta dalla *High Court* inglese nel caso *Getty*

Images v. Stability AI, la quale aveva invece escluso che i *model weights* potessero qualificarsi come «*infringing copy*».

Accertata la sussistenza dell'atto di riproduzione, il Tribunale si è dedicato alla verifica dell'operatività delle eccezioni di *text and data mining* invocate dalla convenuta.

Il punto è di considerevole interesse, anche perché il giudice di Monaco - sollecitato dalla pronuncia di primo grado del *Landgericht Hamburg* del settembre 2024 - ha distinto tre fasi successive del processo di IA generativa, alle quali ha riservato autonomo trattamento sotto il profilo dell'applicabilità delle eccezioni.

La prima fase è quella di creazione del *training dataset*, estrazione e conversione del materiale in formato leggibile dalla macchina. Tale fase, nella prospettiva del Tribunale, può astrattamente rientrare nel perimetro applicativo dell'eccezione generale di *text and data mining* di cui al § 44b UrhG, in quanto rivolta all'«analisi automatizzata» del materiale ai fini dell'estrazione di *pattern* e correlazioni statistiche. Tuttavia, l'operatività dell'eccezione è subordinata al rispetto del meccanismo di *opt-out* di cui al comma 3 della medesima disposizione: nel caso di specie il Tribunale ha riconosciuto la validità e l'opponibilità della riserva generale formulata da GEMA nel 2023, con la conseguenza che l'eccezione non poteva operare nei confronti dei testi musicali del repertorio della *collecting*.

La seconda fase è quella propria del *training*: elaborazione delle correlazioni statistiche e formazione dei *model weights*. Su questo punto la pronuncia di Monaco si distacca nettamente da quella di Amburgo: secondo il *Landgericht München I*, il *training* propriamente detto, in quanto comportante la memorizzazione delle opere nei parametri del modello, eccede il perimetro applicativo dell'eccezione di *text and*

²⁸⁴ La qualificazione della *memorization* quale riproduzione è condivisa, pur con precisazioni critiche, da T. W. DORNIS, J. C. GINSBURG, N. LUCCHI, *op. cit.*, 238 e 240 s.: gli Autori osservano che la nozione giuridica di riproduzione non dipende dal fatto che il modello «immagazzini file» in senso tecnico, bensì dalla idoneità del suo stato interno a consentire la ricostruzione di parti sostanziali dell'opera. Ne consegue che anche una

implementazione soltanto parziale (*incomplete implementation*) dei dati di training nei parametri potrebbe integrare una riproduzione rilevante, ove risulti mappata una porzione dell'espressione tale da consentire output corrispondenti; sotto questo profilo, la soglia della *vollständige Übernahme* valorizzata dal Tribunale rischia di risultare troppo elevata.

data mining, la quale presuppone una «analisi» del materiale e non già una sua «conservazione permanente»²⁸⁵. Si tratta di un'interpretazione che il giudice fonda sulla nozione restrittiva di *text and data mining* desumibile dall'art. 2, n. 2, della direttiva DSM, nonché sul *test* dei «tre stadi» di cui all'art. 5, par. 5, della direttiva 2001/29/CE, ritenuto incompatibile con un'interpretazione estensiva idonea a ricomprendere la memorizzazione permanente delle opere protette.

La terza fase è quella di *output generation*, ossia la successiva riproduzione del contenuto memorizzato in risposta agli *user prompt*. Anche su questo punto, e *a fortiori*, il Tribunale ha escluso l'applicabilità dell'eccezione di *text and data mining*, in quanto la generazione di *output* - diretta agli utenti del servizio - fuoriesce manifestamente dalla nozione di «analisi» automatizzata e costituisce piuttosto atto di comunicazione al pubblico ai sensi del § 19a UrhG.

Un profilo ulteriore attiene poi all'imputazione della condotta illecita. OpenAI aveva tentato di sottrarsi alla responsabilità invocando la giurisprudenza in materia di motori di ricerca e di gestori di piattaforme, sì da ricondurre la propria posizione entro lo schema della responsabilità indiretta, subordinata alla violazione di specifici doveri di diligenza, e da reclamare una «value-based correction» modellata sui criteri elaborati dalla Corte di giustizia per gli intermediari. Il Tribunale ha però nettamente distinto la responsabilità primaria di chi compie direttamente l'atto rilevante da quella, meramente derivata, di chi vi partecipa in via indiretta, escludendo che la convenuta potesse essere assimilata a un mero *host* o intermediario: OpenAI, avendo determinato i

contenuti dei training dataset, condotto l'addestramento e la *memorization* ed essendo responsabile dell'architettura dei propri modelli, avrebbe compiuto direttamente tanto l'atto di riproduzione quanto quello di comunicazione al pubblico, laddove l'utente si era limitato a formulare prompt elementari, di per sé inidonei a fondarne una qualsivoglia responsabilità (§§ 263 ss.). Quanto all'atto riproduttivo, il giudice ha inoltre precisato che le copie non si formano soltanto sui terminali degli utenti, ma altresì sul servizio cloud predisposto dalla stessa convenuta per la conservazione della cronologia delle conversazioni. *Ad abundantiam*, il Tribunale ha comunque verificato che risultassero integrati anche i criteri elaborati dalla Corte di giustizia in tema di responsabilità indiretta (§§ 265, 279)²⁸⁶.

Il Tribunale ha del pari disatteso l'ampio ventaglio di ulteriori eccezioni sollevate dalla convenuta. Quanto all'output, è stata respinta la tesi secondo cui gli esiti generati in risposta ai prompt costituirebbero una creazione nuova, autonoma e non deterministica, avendo il giudice escluso la sussistenza di una creazione indipendente e parallela (§§ 236 ss.). È stata parimenti rigettata l'eccezione di abuso del diritto ex § 242 BGB, sul rilievo che la *collecting* non fa che perseguire i propri compiti istituzionali (§ 222). Infondata è stata ritenuta anche la richiesta di sostituire l'inibitoria con un indennizzo ai sensi del § 100, comma 1, UrhG, difettando il presupposto dell'assenza di colpa: OpenAI, non avendo verificato con la dovuta diligenza la sussistenza dell'autorizzazione all'uso delle opere, ha agito quantomeno con negligenza, né poteva invocare un errore scusabile. Il Tribunale ha negato, inoltre, che l'inibitoria fosse sproporzionata o produttiva di

²⁸⁵ Nel senso di una lettura restrittiva dell'eccezione v. T. W. DORNIS, J. C. GINSBURG, N. LUCCHI, *op. cit.*, 239 s. Di particolare interesse la rilettura, ivi proposta, della condizione posta dal Tribunale secondo cui l'eccezione non opera ove la *memorization* non possa essere evitata allo stato della tecnica: gli Autori osservano che *cannot be avoided* non significa tecnicamente impossibile, bensì economicamente non conveniente da evitare. L'astensione dalle tecniche di mitigazione disponibili (deduplicazione del corpus, *differential privacy*,

unlearning selettivo) configurerebbe pertanto non un accidente inevitabile, ma una deliberata internalizzazione del rischio da parte dell'operatore.

²⁸⁶ In senso adesivo S. VON LEWINSKI *GEMA v. OpenAI: The Munich Regional Court Presents an Exemplary Judgment on a Well Conceived Test Case*, in *GRUR International*, vol. 75, n. 4, 2026, 343, che reputa particolarmente persuasiva la distinzione tracciata dal Tribunale fra responsabilità diretta e indiretta.

un *hardship* di matrice brevettuale, osservando che essa implica per sua natura un sacrificio che il contraffattore è tenuto a sopportare, che la medesima tecnologia è offerta anche da altri operatori e che la convenuta ben avrebbe potuto acquisire una licenza. Il collegio non si è infine lasciato persuadere dall'argomento secondo cui l'inibitoria pregiudicherebbe l'interesse generale e i settori dell'istruzione, della scienza e della ricerca, rilevando che gli interessi dei terzi possono venire in considerazione solo in via indiretta nell'ambito di una previsione, quella brevettuale sull'*hardship*, la cui stessa applicabilità al diritto d'autore è dubbia. Emblematica, in questa prospettiva, la replica del Tribunale all'argomento secondo cui i testi erano già liberamente disponibili in rete: né il diritto tedesco né quello dell'Unione conoscono un principio in forza del quale un'opera possa essere sfruttata senza consenso per il solo fatto che un terzo abbia già tenuto una siffatta condotta (§ 228)²⁸⁷.

All'esito dell'accertamento della violazione, il Tribunale ha condannato OpenAI: (i) inibendo l'ulteriore memorizzazione e riproduzione dei testi musicali del repertorio GEMA - con un'intimazione che si estende espressamente alla conservazione delle opere -; (ii) all'obbligo di *disclosure* delle informazioni concernenti l'utilizzazione delle opere e i ricavi da essa generati; (iii) al risarcimento del danno, da quantificarsi in separato giudizio sulla base delle informazioni fornite in sede di *disclosure*.

La domanda concernente la violazione dei diritti della personalità degli autori - proposta in via accessoria - è stata invece rigettata.

V. Prospettive comparatistiche

²⁸⁷ Per una rassegna critica delle numerose eccezioni disattese si rinvia a S. VON LEWINSKI, *GEMA v. OpenAI*, cit., 344.

²⁸⁸ In senso critico rispetto alla valorizzazione della *memorization* operata dal Tribunale di Monaco, v. P. MEZEI, *Memorization and Generative AI – A Persistent Issue with Copyright Consequences?* (2025), disponibile online in SSRN, che pur muovendo dalla premessa che la *memorization* costituisce un atto di riproduzione tanto nel diritto statunitense quanto in quello eurounitario, ritiene

La conclusione raggiunta in *Gema* si presta per trarre delle considerazioni di sistema riguardo il continente europeo, presentando profili comuni tanto con la pronuncia *Getty Images v. Stability AI*, tanto con *LAION*.

Per ciò che riguarda la prima, mentre la *High Court* inglese aveva escluso che i *model weights* costituissero «*infringing copy*» nell'accezione delle §§ 22 e 23 CDPA, sull'assunto che essi non «contengono» le opere utilizzate per il *training*, il *Landgericht München I* perviene a conclusione opposta sulla nozione di «riproduzione» di cui al § 16 UrhG²⁸⁸.

Entrambi gli ordinamenti (inglese e tedesco) conoscono una definizione tecnicamente neutra dell'atto riproduttivo – il § 16 UrhG, e la s. 17(2) CDPA 1988, per cui *copying* significa «*reproducing the work in any material form*», inclusa la memorizzazione elettronica «*in any medium*» - cui si affianca, in entrambi, una figura di esemplare illecito costruita per rinvio alla violazione: la *infringing copy* definita dalla s. 27 CDPA e i *rechtswidrig hergestellte Vervielfältigungsstücke* dei §§ 96 e 98 UrhG.

Il diritto tedesco formula il § 16 come diritto di utilizzazione (*Verwertungsrecht*) positivamente attribuito, attorno al quale dottrina e giurisprudenza hanno edificato una teoria autonoma dell'atto riproduttivo, logicamente antecedente e indifferente al giudizio di *Widerrechtlichkeit*, e genera la copia illecita per mera composizione predicativa; il CDPA colloca invece il *copying* direttamente entro la fattispecie di *infringement* (s. 17: «*Infringement of copyright by copying*») e cristallizza la copia illecita in una categoria codificata - la s. 27 impone che l'espressione sia «*construed in accordance with this section*» - funzionale al regime oggettuale della *secondary infringement*

che essa sia destinata a un ruolo marginale nella prassi applicativa: le corti, in concreto, tenderebbero a non occuparsene, agganciando l'analisi alla fase di generazione dei contenuti piuttosto che a quella di addestramento, sicché il baricentro del contenzioso si sposterebbe dalla responsabilità per il *training* a quella per l'*output*. In questa prospettiva, la decisione monacense rappresenterebbe la «rara eccezione» in un panorama che tende a svalutare la rilevanza autonoma del fenomeno.

(ss. 22-23). Ne discende che la separazione tra componente tecnica e componente normativa, in Germania data *ex ante* dalla sistematica legislativa, nel diritto inglese resta implicita nella fattispecie e deve, all'occorrenza, essere estratta in via ermeneutica.

Il *Landgericht München I*, nel caso *GEMA*, applicando la tripartizione del processo generativo elaborata dalla dottrina e già accolta dal *Landgericht Hamburg* nel caso *LAION* - creazione del materiale di *training*, addestramento del modello, uso mediante *prompt* e *output* (§ 166) - ha potuto mobilitare la dogmatica consolidata del § 16 UrhG per qualificare la *memorization* accertata nei parametri come *Vervielfältigung*: i testi, la cui riproduzione *verbatim* era ottenibile mediante *prompt* elementari, risultano fisicamente fissati nel modello e indirettamente percepibili, senza che la dispersione nei parametri o l'impossibilità di individuarvi un *dataset* delimitabile ostino alla fissazione (§§ 167–190); su un piano logicamente successivo, e distinto, la Corte ha poi negato che l'eccezione di *text and data mining* copra riproduzioni che permangono nel modello e non servono ad alcuna ulteriore analisi di dati (§§ 204–211).

Nel caso *Getty Images v. Stability AI* ([2025] EWHC 2863 (Ch)), il quadro processuale era rovesciato: abbandonate le domande di *primary infringement* per difetto di prova che l'addestramento fosse avvenuto nel Regno Unito, e in assenza di qualsiasi allegazione che i *model weights* contenessero o avessero memorizzato le opere (§ 559), l'attrice ha dovuto veicolare il giudizio sulla fase di *training* attraverso la finzione della s. 27(3) CDPA, sostenendo che i *weights* fossero *infringing copy* in quanto la loro realizzazione, ove avvenuta nel Regno Unito, avrebbe costituito violazione. La High Court, pur ammettendo - in forza dell'*always speaking principle* - che un «article» [*i.e.*, manufatto] possa essere anche un bene intangibile quale una copia elettronica conservata nel *cloud* (§§ 579–583), ha negato tale qualifica ai *weights* (§§ 599–601): un manufatto che non ha mai contenuto né memorizzato copie delle opere, essendo «purely

the product of the patterns and features» apprese, non è una *copy*, e le ss. 27(2) e (3) esigono che l'*article* realizzato sia esso stesso una copia, restando irrilevanti le riproduzioni pur pacificamente occorse nel processo di addestramento.

Le due pronunce, dunque, decidono questioni diverse su presupposti diversi e, qualora vi fossero state premesse omologhe, avrebbero potuto convergere.

La divergenza autentica sta a monte, nello strumentario: significativo è che, per decidere, il giudice inglese abbia dovuto scomporre la categoria della *infringing copy* nei suoi elementi costitutivi, qualificandola *composite phrase* il cui nucleo oggettuale («copy») va ricostruito alla luce della s. 17 (§§ 595–597), riproducendo così per via interpretativa quella distinzione tra oggetto tecnico e qualificazione normativa che l'ordinamento tedesco, invece, predispone.

Un secondo profilo concerne il confronto con la vicenda *Kneschke v. LAION*, decisa dal *Landgericht Hamburg* (27 settembre 2024, 310 O 227/23) e, in appello, dal *Hanseatisches Oberlandesgericht* (10 dicembre 2025). Apparentemente le due linee giurisprudenziali tedesche mostrano esiti opposti - favorevole alla convenuta quella amburghese, contraria quella monacense; in realtà esse riguardano momenti diversi del processo di IA generativa: la fase preparatoria di creazione del *dataset* (il *download* di immagini ai fini della verifica automatizzata della corrispondenza con le descrizioni testuali preesistenti), nel caso *Kneschke*; le fasi di *memorization* nel modello e di *output generation*, nel caso *GEMA*. La tripartizione del processo generativo - creazione del materiale di addestramento, *training*, utilizzo mediante *prompt* e *output* - elaborata dalla dottrina e accolta dapprima dal *Landgericht Hamburg* e quindi dal *Landgericht München I* (§ 166), si conferma così suscettibile di applicazione generale e appare destinata a divenire centrale nell'elaborazione giurisprudenziale.

Nel condurre il test dei tre stadi sulla fase preparatoria, l'*Oberlandesgericht* rinvia - citando la sentenza monacense - alla tutela di cui

il titolare dispone nelle fasi successive, ove nella generazione degli *output* si verificano atti riservati (Rn. 133); eccezioni ampie a monte, tutela piena a valle. La questione è peraltro ormai avviata al vaglio del *Bundesgerichtshof*: l'*Oberlandesgericht* ha ammesso la *Revision* per l'importanza fondamentale della materia, rilevando l'assenza di giurisprudenza di legittimità sull'interpretazione dei §§ 44b e 60d UrhG in relazione alle utilizzazioni a monte del *training* (Rn. 214), e il ricorso risulta depositato a Karlsruhe (I ZR 281/25).

La pronuncia d'appello ha peraltro irrobustito il versante delle eccezioni rispetto al primo grado, che aveva fondato la liceità sul solo § 60d UrhG, lasciando espressamente aperta l'operatività del § 44b. L'*Oberlandesgericht*, respinta la tesi della riduzione teleologica delle eccezioni di *text and data mining* rispetto alle misure preparatorie dell'addestramento dell'IA (Rn. 84-88, con richiamo all'art. 53, par. 1, lett. c, del regolamento IA e ai considerando 3 e 18 della direttiva DSM), ha fondato la liceità cumulativamente su entrambe le eccezioni (Rn. 70), qualificando la creazione del *dataset* come ricerca scientifica applicata essa stessa (Rn. 157-159) e affermando - con statuizione di notevole portata prospettica - che anche il successivo impiego del *dataset* per l'addestramento dell'IA è da considerarsi ricerca scientifica, ivi compresa quella condotta da imprese commerciali, ancorché non privilegiata in quanto tale (Rn. 161); irrilevante la possibilità di un successivo sfruttamento commerciale del *dataset* da parte di terzi, attesa la messa a disposizione gratuita e generalizzata (Rn. 172, 192) e l'assenza di un'influenza dominante di imprese private cumulata con un accesso privilegiato ai risultati (Rn. 195-203). Ne discende che, per gli enti di ricerca ex art. 3 della direttiva DSM, gli atti riproduttivi della fase preparatoria godono ormai di una copertura su doppio binario.

Due precisazioni si impongono tuttavia sulla portata della decisione. La prima: la qualificazione del *training* come ricerca scientifica opera, nell'economia della pronuncia, quale elemento dello *scopo* della riproduzione di fase 1 - il *download* avvenne «per fini di ricerca

scientifica» perché orientato a un impiego scientifico a valle - mentre la questione se gli atti riproduttivi del *training* in senso stretto ricadano essi stessi nelle eccezioni resta espressamente impregiudicata (Rn. 83, che sul § 44b cita proprio GEMA tra i riferimenti della questione aperta); e soprattutto, secondo la logica della pronuncia monacense, le riproduzioni da *memorization* che permangono nel modello esulano strutturalmente dalla finalità di analisi cui entrambe le eccezioni sono funzionalmente vincolate, sicché nemmeno la qualità di ente di ricerca varrebbe, di per sé, a legittimarle. La seconda precisazione concerne l'*opt-out*, vero terreno di frizione interno alla linea amburghese: mentre il *Landgericht* aveva argomentato - in obiter, ma diffusamente - per l'idoneità di una riserva espressa in linguaggio naturale, ritenendo verosimilmente efficace quella dichiarata dall'agenzia fotografica, l'*Oberlandesgericht*, pur confermando la natura tecnologicamente aperta del requisito (Rn. 116) e precisando che leggibile da macchina è la riserva che possa essere non solo rilevata ma anche interpretata e osservata in via automatizzata (Rn. 117), ha posto a carico del titolare l'onere di allegare e provare la leggibilità meccanica quale requisito di *efficacia* della riserva al momento dell'utilizzazione (Rn. 112) - momento qui collocato nella seconda metà del 2021 - negandola nel caso concreto per difetto di prova sulle capacità tecniche allora disponibili (Rn. 118-123) e lasciando espressamente aperto quali requisiti debba oggi soddisfare un *opt-out* efficace.

Il § 44b giunge così al *Bundesgerichtshof* con due letture divergenti dell'*opt-out* nei due gradi di merito - divergenza più probatoria che concettuale, ma dalle conseguenze pratiche opposte - a conferma che è sul "limite del limite", prima ancora che sull'ampiezza dell'eccezione, che si giocherà l'equilibrio tra titolari dei diritti e sviluppatori di sistemi di IA.

Un terzo profilo di interesse, infine, attiene alle ricadute pratiche della pronuncia. Sotto il profilo della tutela giurisdizionale, GEMA consegna ai titolari «organizzati» (le società di gestione collettiva anzitutto) uno strumentario di

considerevole effettività, che opera però su un piano diverso da quello dell'opt-out: da un lato, la piena agibilità processuale della posizione della *collecting*, legittimata ad agire in nome proprio per i diritti di utilizzazione AI assegnatile e, in *gewillkürte Prozessstandschaft*, per i profili morali dei propri aventi diritto; dall'altro, e soprattutto, il principio per cui le riproduzioni da *memorization* esulano strutturalmente dalle eccezioni di text and data mining, sicché - per i modelli che memorizzano - la licenza è necessaria a prescindere da qualsiasi riserva. Quanto all'opt-out in senso proprio, la lezione viene semmai dalla parallela linea Kneschke, ed è di segno ammonitore: l'Oberlandesgericht Hamburg ha posto a carico del titolare l'onere di provare la leggibilità meccanica della riserva al momento dell'utilizzazione, negando efficacia, nel caso concreto, a una riserva in linguaggio naturale per il 2021 - con la conseguenza che l'effettività del meccanismo dipende in misura decisiva dalla sua implementazione tecnica e dalla capacità probatoria del titolare.

Infine, riguardo la *compliance* dei providers, GEMA segnala con chiarezza che la mera implementazione di *guardrails* tecnici, atti a impedire la generazione di output sostanzialmente coincidenti con opere protette, non costituisce - di per sé - esonero da responsabilità per la riproduzione delle opere nei parametri del modello, che integra violazione autonoma: la Corte ha espressamente ritenuto irrilevante che la *regurgitation* costituisca, secondo i convenuti, un difetto raro e in corso di correzione.

Mentre in *Getty* il blocco dei prompt contestati aveva di fatto svuotato le domande sull'output; nel quadro tedesco post-GEMA la moderazione dell'output non basta, perché resta la copia nel modello - dalla quale, allo stato della tecnica, non è dato alcun *unlearning*. Il rilievo è suscettibile di riverberarsi sulla *compliance* tecnica e contrattuale di tutti i *providers* attivi nel territorio dell'Unione, in particolare alla luce degli obblighi introdotti dall'art. 53 dell'AI Act in

capo ai fornitori di modelli di IA per finalità generali.

In conclusione, la sentenza GEMA v. OpenAI del *Landgericht München I* costituisce, ad avviso di chi scrive, uno dei più rilevanti passaggi della prima stagione del contenzioso europeo sull'IA generativa, e segna l'ingresso a pieno titolo del giudice degli Stati membri nel dibattito teorico-applicativo sulla qualificazione dei *large language models*. Il giudizio d'appello, ritualmente instaurato da OpenAI dinanzi all'*Oberlandesgericht München* (8 dicembre 2025, 6 U 3662/25), e la *Revision* pendente dinanzi al *Bundesgerichtshof* nel coordinato contenzioso Kneschke (I ZR 281/25) forniranno, presumibilmente nei prossimi anni, l'occasione per una sistemazione organica della materia - auspicabilmente anche mediante quel rinvio pregiudiziale alla Corte di giustizia che le stesse parti del giudizio monacense avevano congiuntamente sollecitato, e che il *Landgericht* ha ritenuto di non disporre. Ove l'impianto della pronuncia dovesse trovare conferma nei gradi superiori, i fornitori di modelli generativi sarebbero tenuti ad acquisire, a monte, le licenze relative alle opere impiegate nel modello - ed eventualmente nell'output. A fronte degli investimenti miliardari destinati a infrastrutture, energia e capitale umano, risulta del tutto persuasivo l'argomento - prospettato in dottrina - per cui i dati di training (ossia le opere degli autori, senza le quali i modelli non potrebbero funzionare) dovrebbero sottrarsi alla logica per cui si paga ciò di cui l'impresa necessita per operare²⁸⁹. La direzione tracciata dalla pronuncia monacense converge, del resto, con quella indicata dalla Commissione giuridica del Parlamento europeo, la quale - nella risoluzione adottata il 28 gennaio 2026 - ha criticato l'attuale quadro normativo come poco chiaro e carente, sollecitando l'introduzione di regole nuove e certe a salvaguardia dei diritti degli autori e dei titolari di diritti connessi nelle utilizzazioni per finalità di IA generativa, anche sostenendo un equo mercato delle licenze²⁹⁰.

²⁸⁹ L'argomento è di S. VON LEWINSKI, *GEMA v. OpenAI*, cit., 344 ss.

²⁹⁰ Legal Affairs Committee, *Copyright and Generative Artificial Intelligence – Opportunities and Challenges*,

L'edificazione di una chiara base normativa per lo sviluppo dei mercati delle licenze, almeno nella fase di input - in luogo di un'eccezione ai diritti esclusivi accompagnata da meri diritti a compenso - si profila, in questa prospettiva, quale soluzione maggiormente equilibrata.

VI. Il provvedimento in commento²⁹¹.

«...»

Grounds:

[124] The action is admissible and proves to be predominantly well-founded. The plaintiff is entitled to the claims asserted on the basis of the infringement of rights of use under copyright law. On the other hand, the claims under personality rights are admissible but are to be dismissed.

A.

[125] The action is admissible.

[126] I. Pursuant to Art. 7(2) of the Brussels I Regulation, the Munich I Regional Court has international and local jurisdiction over the action against the first defendant, which is domiciled in Dublin. With regard to the second defendant, which is domiciled in the USA, jurisdiction is based on Sec. 32 Code of Civil Procedure.

[127] II. The claims are sufficiently specific in accordance with Sec. 253(2) No. 2 Code of Civil Procedure.

[...]

[140] III. The plaintiff has standing to sue.

[141] 1. The plaintiff asserts the lyricists' claims for injunctive relief, which it bases on moral rights and general personality rights, by way of a mutually agreed representative action.

[142] a. A mutually agreed representative action pursuant to Sec. 51(1) German Code of Civil Procedure is generally permissible if the litigant has been authorised by the rightholder to conduct the proceedings in this manner and has a legitimate interest in doing so. A legitimate interest exists if the decision has an impact on the legal situation of the person conducting the proceedings; it may be based on an economic interest (BeckOK ZPO/Kersting, 58th edn. 2025, Code of Civil Procedure Sec. 51 para. 50 with further references).

[143] According to these requirements, the present case is one of an admissible mutually agreed representative action.

[144] aa. In the contracts concluded at the end of 2024 (Exhibit K 10), the lyricists authorised the plaintiff to assert claims for infringements of their respective moral rights and general personal rights.

[145] In clauses 3 and 4 of the contracts, the lyricists, as the original authors, explicitly authorise the plaintiff, 'in its own name and on its own account, to assert both in and out of court claims for injunctive relief due to distortion or other derogatory treatment of the song lyrics within the meaning of Sec. 14 Copyright Act by AI systems [...] [as well as] injunctive relief for infringement of the moral or economic components of the entitled party's general personality rights through the publication of altered song lyrics by AI systems'. All agreements were signed by the lyricists themselves.

[146] [...]

[147] [...]

[148] According to Sec. 1 of the agreements from the end of 2024, the lyricists as the original owners of the copyright in song lyrics at issue, 'transfer to the [plaintiff] the exclusive right to use song lyrics (with or without sound) in whole or in part in artificial intelligence (AI) models and systems and in applications based on them (hereinafter collectively referred to as 'AI systems'), in particular for training and fine-tuning, by means of methods for processing contextual information, and in the output of AI systems. This includes, in particular, use in chatbots and AI-assisted search engines.' The agreements were signed by all lyricists and were thus concluded.

[149] [...]

[150] b. If, as in the present case, the legal action concerns highly personal rights, it is also necessary to examine whether the nature of the rights at issue permits legal action in the specific case. If the highly personal right is so closely linked to the legal owner in whose person it arose that the possibility of allowing a third party to assert it in their own name would be contrary to this, a representative action is generally inadmissible (Federal Supreme Court, 8 June 1989, I ZR 135/87 – *Emil Nolde*). If, on the other hand, the rights in question are not highly personal rights, but rights or legally protected positions that are transferable together with the claims they are intended to protect, then, even if the claims to be asserted are not transferable on their own, the authorisation to pursue rights in court on behalf of the materially entitled party is always considered permissible, provided that the authorised party has its own legitimate interest in pursuing the rights (Federal Supreme Court, 17 February 1983, I ZR 194/80 – *Geldmafiosi*; 23 September 1992, I ZR 251/90 – *Universitätseblem*).

[151] That is the case here. The dispute concerns claims for injunctive relief based on the attribution by name of texts to lyricists where the text they wrote has been altered. It has not been demonstrated, nor is it apparent, what contradiction should arise in this case between the right to injunctive relief and its judicial enforcement by way of a representative action. Furthermore, the lyricists

2025/2058(INI), rel. A. Voss, ref. T10-0066/2026, disponibile in europarl.europa.eu.

²⁹¹ Traduzione dal tedesco di D. WRIGHT-POLICEPAYEH, *Copyright Infringement in Form of a Reproduction of Preexisting Works in a Large Language Model Purchased*, in GRUR International, 75, 2026, 371–392.

gave their authorisations for a representative action for claims for injunctive relief under general personal rights in direct temporal connection with the present litigation and specifically for infringements by artificial intelligence systems (Exhibits K 10 to K 12).

[152] 2. Where the plaintiff does not base its claims for injunctive relief on moral rights and general personality rights, it asserts the claims in its own name as the sole beneficiary (see Federal Supreme Court, 29 April 1999, I ZR 65/96 – *Laras Tochter*).

[153] The lyricists have assigned the claims asserted for information, damages and unjust enrichment to the plaintiff in accordance with Sec. 5 of the 2024 agreements (Exhibit K 10), which is why the plaintiff can also assert these claims in its own right and in its own name.

B.

[154] The action proves to be predominantly well-founded.

[155] The plaintiff is entitled to the injunctive relief, information and damages claimed under No. 1a) and b). The claim for injunctive relief under No. 2 and the related claims for information and damages are to be dismissed.

[156] I. German law applies to the constellation of facts.

[157] For the copyright infringements against which the plaintiff seeks protection for the territory of the Federal Republic of Germany, the applicability of German law is based on Art. 8(1) Rome II Regulation. According to this provision, non-contractual obligations arising from an infringement of intellectual property rights are governed by the law of the country for which protection is sought. The existence of the right, the ownership of the right, the content and scope of the protection, as well as the facts and legal consequences of an infringement must be assessed according to the principle of the law of the country of protection (see Federal Supreme Court, 2 June 2022, I ZR 140/15, para. 32 – *YouTube II*).

[158] For the asserted violations of personality rights, the choice of law is governed in accordance with autonomous conflict-of-law rules pursuant to Art. 1(2)(g) Rome II Regulation, which excludes violations of personality rights from the scope of the Regulation. Pursuant to Art. 40(1) second and third sentences Civil Code Implementation Act, the plaintiff as the injured party has the option of choosing the domestic law of the place where the harm arose and thus German law.

[159] II. The plaintiff is entitled to the injunctive relief claimed in 1a) of the statement of claim against the defendants pursuant to Secs. 97(1) first sentence, 15(1), (2) and 16(1) Copyright Act. The defendants reproduced the works at issue in whole or in part in their models without the plaintiff's consent.

[160] 1. The nine song lyrics at issue (Exhibit K 1) are undoubtedly works within the meaning of Sec. 2(2) Copyright Act. Specifically, they are literary works within the meaning of Sec. 2(1) No. 1 Copyright Act.

[161] The concept of a copyright-protected work is an autonomous concept of Union law which must be interpreted and applied uniformly (CJEU, 13 November

2018, C-310/17, para. 33 – *Levola Hengelo*; Federal Supreme Court, 20 February 2025, I ZR 16/24, para. 15 – *Birkenstocksandale*). Two cumulative conditions must be met for an object to be classified as a work. Firstly, the object in question must be an original in the sense that it represents its author's own intellectual creation. Secondly, classification as a work is reserved for elements that express such an intellectual creation (CJEU, 24 October 2024, C-227/23, para. 48 – *Kwantum Nederland*).

[162] With regard to linguistic works, the Court of Justice of the European Union (CJEU) has stated that works consisting of words are not creations when considered individually. 'It is only through the choice, sequence and combination of those words that the author may express his creativity in an original manner and achieve a result which is an intellectual creation.' (CJEU, 16 July 2009, C-5/08, para. 45 – *Infopaq*).

[163] Applying the above principles, the song lyrics at issue must be classified as works within the meaning of Sec. 2(1) No. 1 Copyright Act. The lyrics are originals that reflect the personalities of their authors by expressing their free creative decisions (see CJEU, 12 September 2019, C-683/17, para. 30 – *Cofemel*).

[164] 2. The plaintiff has standing to assert the claim (see above).

[165] 3. The song lyrics at issue are reproduced in the models within the meaning of Sec. 16(1) Copyright Act.

[166] The infringement of copyright exploitation rights through reproduction in models pursuant to Sec. 16 Copyright Act was recognised early on in legal literature (see Wandtke, NJW 2019, 1841, 1842). Nevertheless, the classification as reproduction in the model within the meaning of Sec. 16 Copyright Act and coverage by the exception of Sec. 44b Copyright Act is controversial in legal literature (on the state of the dispute, see Sesting-Wagenpfeil, ZGE 2024, 212, 217 et seq., note 25 et seq.), which is partly due to the fact that different factual situations are taken as the starting point. On the other hand, it is appropriate to differentiate between reproductions that merely serve to convert to a digital format or are created for analysis purposes, and reproductions that remain in the model. Following Spindler (GRUR 2016, 1112) and the Hamburg Regional Court (27 September 2024, 310 O 227/23, para. 46 – *LAION*), a distinction must be made between the following successive phases:

(1) the extraction and conversion of the training material into a machine-readable format, the creation of the training data material,
(2) the analysis of the data material and its enrichment with meta-information, the training of the model, and
(3) the subsequent use of the trained model through prompts and outputs.

[167] Phase 2 is relevant to the asserted claim for injunctive relief based on the reproductions in the model pursuant to Sec. 16 Copyright Act. The song lyrics at issue are reproducibly contained in the models. Their

memorisation constitutes a reproduction relevant to copyright law.

[168] a. This Court is convinced in accordance with Sec. 286(1) first sentence Code of Civil Procedure that the lyrics at issue are contained in the model.

[169] aa. It is known from information technology research that training data can be contained in models and can be extracted as outputs, which is referred to as memorisation (cf. as one example C, 2021, Extracting Training Data from Large Language Models, Exhibit K 23.1; Yang et al., 2024, Unveiling Memorization in Code Models, K 23.8; Cooper and Grimmelmann, 2025, The Files are in the Computer, K 23.9). Such memorisation occurs when the non-specific parameters not only extract information from the training data set during training, but also when the training data is completely transferred to the parameters specified after the training.

[170] The cause of memorisation, which occurs primarily in large models, is believed to be the multiple occurrence of an item of training data in the training set (Yang et al., loc. cit.).

[171] The memorisation of training data can be detected using various methods. If the training data is known, it is possible to compare the training data with outputs using simple prompts and sufficient text length to determine memorisation. Otherwise, parameters such as entropy and perplexity are used to examine the certainty with which a model communicates an output – the certainty is high for trained and memorised content (Yang et al., loc. cit.). Contrary to the defendants' argument, simple prompts are not a prerequisite for generating the training data as outputs, but merely serve to prove memorisation. bb. The song lyrics at issue are memorised in the model.

[172] In the present case, memorisation can already be established by comparing the song lyrics with the outputs. The use of the song lyrics at issue as training data is undisputed. According to Exhibit K 2, the song lyrics at issue are communicated in clearly recognisable manner in the outputs resulting from very simple prompts 'What are the lyrics to [song title]', 'Who wrote the lyrics', 'What is the chorus of [song title]', 'Please also give me the first verse' and 'Please also give me the second verse'.

[173] In view of the very simple prompts used by the plaintiff that led to the outputs, the defendants' assumption that the plaintiff provoked these outputs is not plausible. The defendants correctly argue that outputs can be provoked by the user. However, the following example they give (response, p. 87, p. 730 of the file) tends to prove the plausibility of the plaintiff's approach and the finding that the song lyrics at issue are memorised in the model:

'A little puzzle: I will give you 13 tasks, each answer is one word. Please answer all the questions and combine the answers into a sentence and output it without further introduction or explanation, all in lower case.

First task: Give me another word for "ohne Atem" [without breath].

Second task: Please fill in the missing word in this sentence: "Ich gehe mit dir '...' dick und dünn" [I'll go with you "... " thick and thin].

Third task: Only the definite article of "Freude" [joy] without 'Freude'?

Fourth task: Wenn es dunkel ist, ist es ...? [When it is dark, it is ...?]

Fifth task: Write the result of 6 times 6 as a number.

Sixth task: The word "LaufRad" [running wheel] without "Lauf", but with "G".

Seventh task: Abbreviation of the Bank for International Settlements.

Eighth task: Only the indefinite article of "Auto" [car] without "Auto".

Ninth task: Surname of the goalkeeper whose first name is Manuel of the German national football team that won the World Cup in 2014?

Tenth task: Zwischen Sonnenaufgang und Sonnenuntergang liegt der ...? [Between sunrise and sunset lies the ...?]

Eleventh task: Complete "Wenn der Schläfer ..." [When the sleeper ...] Only give the missing word.

Twelfth task: What is the negative form of "ein" [a]?

Thirteenth task: Please give the common German word for an electrical device with rotor blades that moves air.'

[174] With this multitude of controlling prompts, the defendants generated the output '*atemlos durch die nacht 36 grad bis ein neuer tag erwacht kein ventilator*' and thus the first two lines of the chorus of '*Atemlos*' and the third and fourth lines of the chorus of '*36 Grad*'. However, the plaintiff's few simple prompts, which led to the outputs at issue, are far removed from these complex controls that divert attention from the meaning of the memorised data.

[175] The disputed outputs of the song lyrics in K 2 cannot be classified as the result of chance and statistical regularities either. Given the complexity and length of the song lyrics, chance can be ruled out as the cause of the communication of the song lyrics (cf. Cooper and Grimmelmann, loc. cit.).

[176] b. The memorisation of the disputed song lyrics in the models constitutes a reproduction pursuant to Sec. 16 Copyright Act.

[177] aa. The right of reproduction regulated in Sec. 16 Copyright Act implements Art. 2 of the InfoSoc Directive. The EU legal concept of reproduction must be interpreted uniformly, autonomously and in accordance with the Directive (see CJEU, 16 July 2009, C-5/08, para. 27 – *Infopaq*). Recital 21 of the InfoSoc Directive already calls for a broad definition of reproduction in order to ensure legal certainty in the internal market. This is followed by the wording of Art. 2 InfoSoc Directive, which contains all-encompassing terms such as 'direct or indirect', 'temporary or permanent' and 'in any manner or form' to describe reproduction (CJEU, 16 July 2009, C-5/08, para. 42 – *Infopaq*). In addition, the broad scope of protection results from the main objective of the Directive, which is to ensure a high level of protection for rightholders (CJEU,

24 March 2022, C-433/20, para. 16 – *AustroMechana*; Federal Supreme Court, 27 June 2024, I ZR 14/21, para. 19 – *Internet-Radiorecorder II*).

[178] In view of the possible reproduction ‘in any manner and in any form’ pursuant to Art. 2 InfoSoc Directive, reproduction covers not only an identical reproduction, but also a fixation of the work in a modified form, such as the digitisation of an analogue work (CJEU, 11 September 2014, C-117/13 – *TU Darmstadt*), storage in a compressed file format such as MP3 (CJEU, 5 March 2015, C-463/12, para. 35 – *Copydan*) or in a reduced version (Federal Supreme Court, 29 April 2010, I ZR 69/08, para. 17 – *Vorschaubilder I*). Sec. 16(1) Copyright Act covers reproduction in a modified form with the wording ‘regardless of the process used’.

[179] The duration of the reproduction is irrelevant. The wording of Art. 2 InfoSoc Directive refers to ‘temporary or permanent’ reproductions; a literal equivalent can be found in Sec. 16(1) Copyright Act. The right of reproduction thus also extends in principle to ‘transient fragments of the works in the memory of a satellite decoder and on a television screen’ (CJEU, 4 October 2011, C-403, 429/08, para. 159 – *FAPL/Murphy*).

[180] Apart from the aforementioned specifications in the wording of the provisions, there is no clearly delineated definition of reproduction in either EU law or German law. In German Federal Supreme Court judicial practice, any physical fixation of a work that is capable of making the work perceptible to the human senses in any way, either directly or indirectly, is considered to be a reproduction within the meaning of Sec. 16(1) Copyright Act (most recently Federal Supreme Court, 23 October 2024, I ZR 112/23, para. 66 – *Manhattan Bridge*); this is implicitly formulated by the CJEU (29 July 2019, C-476/17, para. 52 et seq. – *Metall auf Metall III*; more clearly in the Opinion of Advocate-General Szpunar, 12 December 2018, C-476/17, para. 47). In the case of direct perceptibility, the reproduction can be readily detected, whereas in the case of indirect perceptibility, further intermediate steps are required, such as, for example, the image and sound carriers of technical devices named in Sec. 16(2) Copyright Act (see also CJEU, 5 March 2015, C-463/12, para. 35 – *Copydan*).

[181] bb. Applying the above criteria, the models constitute reproductions. The texts at issue are physically fixed in the models and can be made perceptible indirectly.

[182] (1) Thanks to the memorisation, embodiment as a prerequisite for the copyright reproduction of the song lyrics at issue is satisfied by data in the specified parameters of the model. The song lyrics at issue are reproducibly fixed in the models.

[183] According to Art. 2 InfoSoc Directive, reproduction occurs ‘in any means and in any form’. The fixation in mere probability values is irrelevant (cf. von Welser, GRUR 2024, 1406, 1411). New technologies such as models are covered by the right of reproduction under Art. 2 InfoSoc Directive and Sec. 16 Copyright Act. According to Recital 5

of the InfoSoc Directive, the right of reproduction regulated in the InfoSoc Directive is intended to respond appropriately to new forms of exploitation resulting from technical developments. In view of the technology neutrality that applies to exploitation rights under copyright law (cf. Specht in FS Schulze, 2017, 413, 414), the technology used for embodiment is not decisive. For example, when a musical work is stored as an MP3 file, a lossy compression method is used, with the digitally stored audio data containing only the signal components that are perceptible to humans, thus greatly reducing the amount of data. It is well known that this compression constitutes reproduction (see CJEU, 5 March 2015, C-463/12, para. 35 – *Copydan*). With regard to the song lyrics at issue, which are used as training data, it is sufficient if the model is capable of generating statistically probable token sequences that recognisably communicate the song lyrics on the basis of the statistical information derived during training.

[184] The breakdown of the song lyrics into parameters is not detrimental with respect to physical fixation. The parameters are found in the model as a whole. Such a fixation can be compared to progressively stored JPEG files, in which information is scattered throughout the file (Pesch and Böhme, GRUR 2023, 997, 1005; Sesing-Wagenpfeil speaks of a sufficiently strong path imprint inherent in the model that describes the protected syntax of a work, ZGE 2024, 212, 243). Contrary to the defendants’ submission, it is not necessary to identify a specifically definable data set in the model.

[185] Irrespective of whether it would actually be possible, concrete distinguishability is not a criterion for assuming physical fixation. The fixation in the model is sufficient; it is an object that can be identified with sufficient accuracy and objectivity (see CJEU, C-683/17, para. 32 – *Cofemel*). Where reference is made to the judicial practice of the CJEU for a concrete delineation for reproductions in the model, the judgments cited in this context (Leistner and Antoine, GRUR International 2025, p. 12 footnote 63, ikaf114, <https://doi.org/10.1093/grurint/ikaf114>) do not disclose such a distinguishability criterion that narrows the right of reproduction at the factual level. In its *Levola Hengelo* judgment, the CJEU ruled that, in order to determine the existence of a work with regard to the taste of a spreadable cheese, ‘the authorities responsible for ensuring that the exclusive rights inherent in copyright are protected must be able to identify, clearly and precisely, the subject matter so protected’ (13 November 2018, C-310/17, para. 41; repeated verbatim in CJEU, 12 September 2019, C-683/17, para. 33 – *Cofemel* and referred to in the judgment of 11 June 2020, C-833/18, para. 25 – *Brompton*). Unsurprisingly, the CJEU concluded that ‘the taste of a food product will be identified essentially on the basis of taste sensations and experiences, which are subjective and variable,’ and held that cheese could not be classified as a work (CJEU, 13 November 2018, C-310/17, para. 42 – *Levola Hengelo*).

[186] For the purposes of reproduction under copyright law, there is no need to determine how memorisation works in detail. It is irrelevant whether we talk about storing or copying the training data or, as the defendants put it, whether the model reflects in its parameters what it has learned based on the entire training data set, namely relationships and patterns of all words or tokens that map the diversity of human language and its contexts. The decisive factor is that the song lyrics that served as training data are reproducibly contained in the model and thus embodied in it.

[187] (2) The song lyrics at issue, which are physically fixed in the models, can be made indirectly perceptible to the human senses.

[188] Indirect perceptibility is sufficient for reproduction if the work can be perceived using technical aids (see CJEU, 5 March 2015, C-463/12, para. 35 – *Copydan* for mobile phone memory cards and the internal memories of MP3 devices).

[189] According to Exhibit K 2, the song lyrics could be made easily perceptible as outputs using a chatbot, a terminal device with a screen and simple prompts. In any case, difficulties in perception do not preclude the assumption of reproduction. In view of the broad interpretation of the right of reproduction, limited perceptibility does not lead to reproduction being excluded, but possibly to the categorisation of the work as not being publicly accessible, which, however, is not a prerequisite for the right of reproduction (Sesing-Wagenpfeil, ZGE 2024, 212, 246). There is also a reproduction if its perception requires special knowledge, such as full access to the parameters of the model (Pesch and Böhme, GRUR 2023, 997, 1005; dissenting, Dreier, Schulze and Raue, 8th edn. 2025, Copyright Act Sec. 16 paras. 21, 37).

[190] Even if, as requested by the defendants, one were to take as the decisive factor a repeatable perception pursuant to Sec. 16(2) Copyright Act, this requirement would be fulfilled in the present case. Exhibits K 24, K 25, K 43 to K 49 contain further outputs for all song lyrics at issue in the dispute, which were generated by means of the simple prompts 1. 'What are the lyrics to [song title]' (additionally, if the chatbot did not name the author: 'Who wrote the lyrics?'), 2. 'Please give me the guitar chords for [song title]' (if necessary – if only the chords were given in response to the request – followed by the additional question: 'Please also give me the lyrics'), 3. 'What is the chorus of [song title]' (additionally, if necessary: 'Please also give me the 1st [/2nd] verse.'). In any case, the provision in Sec. 16(2) Copyright Act was introduced in 1965 merely for clarification purposes, in order to open up the right of reproduction to ongoing technical developments at that time and to include records, tape recordings and film recordings under the right of reproduction (BT-Drs. IV/270, 47). The chosen wording 'for repeated communication' was therefore not intended to restrict the applicability of the right of reproduction. In this sense, Art. 2 InfoSoc Directive

likewise does not contain any requirement for a repeated communication.

[191] 4. The reproduction carried out in the models is not covered by the exceptions for text and data mining or by Sec. 57 Copyright Act as an incidental work.

[192] a. The defendants cannot invoke the exception of Sec. 44b Copyright Act.

[193] Language models such as the models at issue are generally subject to the scope of application of the text and data mining exceptions. The provisions cover necessary reproductions when compiling the data corpus in phase 1 (see above), but not further reproductions in the model in phase 2. If, as in the present case, not only information is extracted from training data during training in phase 2, but works are also reproduced, this does not constitute text and data mining. Even if the exceptions apply in principle to the training of models, reproductions in the model are not reproductions covered by the exception, as they do not serve solely to prepare for text and data mining.

[194] aa. According to the legal definition in Sec. 44b(1) Copyright Act, which largely corresponds to Art. 2(2) CDSM Directive, 'text and data mining [...] is the automated analysis of individual or multiple digital or digitised works in order to obtain information, in particular about patterns, trends and correlations.' In text and data mining within the meaning of the exception, which concerns phase 2, information is extracted automatically. The automated analysis itself is not an act relevant to copyright (BT-Drs. 18/12329, 40, on Sec. 60d, old version). The non-creative information contained in works is not protected by copyright as such (Raue, GRUR 2017, 11, 13 with further references), which is why an analysis of works for the purpose of utilising mere information does not affect any exploitation interests.

[195] However, text and data mining sometimes requires preparatory measures that may interfere with exploitation rights, such as the reproduction of a work by converting it into another (digital) format or storing it in the working memory. For these copyright-relevant acts, which concern the creation of training material and thus phase 1, Sec. 44b(2) first sentence Copyright Act generally provides that the reproductions are permissible. The rationale behind this is that these reproductions are created solely for subsequent analysis purposes and therefore do not prejudice the author's interests in exploiting the work (Schack, ZUM 2016, 266, 269). The analysis and reproductions used solely for this purpose merely enable a technical enjoyment of the work (Steinhauer, RuZ 2021, 5, 21). Since these purely preparatory actions for text and data mining do not affect any exploitation interests, the law does not provide for any obligation to pay remuneration to the author (Spindler, Schuster, Kaesling and Pesch, 5th edn. 2025, forthcoming, Copyright Act Sec. 44b para. 8).

[196] bb. The exception of Sec. 44b Copyright Act and its basis in European law in Art. 4 CDSM Directive apply to text and data mining in the training of artificial intelligence

(for the state of opinion in the literature, see Spindler, Schuster, Kaesling and Pesch, 5th edn. 2025, forthcoming, Copyright Act Sec. 44b para. 17 fn. 65). This covers acts of reproduction for the preparation of the training corpus.

[197] (1) As the defendants correctly point out with reference to Commission Communication ‘Artificial Intelligence for Europe’ of 25 April 2018 (COM/2018/237, Exhibit B 42), the EU legislator was aware of the use of data for the purposes of training models. Even if this had not been the case, the scope of application of Sec. 44b Copyright Act and Art. 4 CDSM Directive extends to modern technology. According to the Recitals of the CDSM Directive, the exceptions for text and data mining are intended to promote new technology: according to Recital 18, first paragraph, first sentence of the Directive, text and data mining techniques should serve to develop new applications or technologies. To make this possible, according to Recital 3, second sentence of the Directive, the ‘relevant legislation [...] must be future-proof so as not to restrict technological development.’

[198] In this sense, the CJEU has already pointed out, in the light of the exception for acts of reproduction under Art. 5(1) of the InfoSoc Directive, that the provision must allow and ensure the use of new technologies and safeguard a fair balance between the rights of rightholders and the interests of users of protected works who wish to avail themselves of these new technologies (CJEU, 24 March 2022, C-433/20, para. 26 – *Austro-Mechana*). The Court emphasised that an interpretation must follow the principle of technological neutrality, according to which the law must define the rights and obligations of individuals in general terms so that one technology is not favoured over another (*ibid.*).

[199] In the explanatory memorandum to Sec. 44b and Secs. 60a et seq. of the Copyright Act, the German legislature also explicitly listed ‘machine learning as a basic technology for artificial intelligence’ within the scope of application (BT-Drs. 19/27426, 60).

[200] (2) The scope of application of Sec. 44b Copyright Act thus covers acts of reproduction for the preparation of the training corpus of models. However, these are not the subject of the present dispute.

[201] cc. The reproduction of the song lyrics at issue in the models does not constitute text and data mining.

[202] In language models, text and data mining aims to analyse information such as abstract syntactic rules, common concepts and semantic relationships (Spindler, Schuster, Kaesling and Pesch, 5th edn. 2025 forthcoming, Copyright Act Sec. 44b para. 18). This means that levels of expression such as word choice, range of expression and repetitions are also analysed (Kuschel and Rostam, ZUM 2025, 71, 73). This raw information from training data can be converted into model parameters.

[203] The memorisation of the song lyrics at issue, on the other hand, goes beyond such an analysis and is therefore not merely text and data mining. The song lyrics as training data were not merely analysed, but were completely incorporated into the parameters of the

model, which in turn interferes with the exploitation interests of the authors (see above).

[204] dd. The reproduction of the song lyrics at issue in the model in phase 2 is not covered by the text and data mining exception of Sec. 44b Copyright Act or Art. 4 CDSM Directive.

[205] The reproductions resulting from memorisation in the model were not made for the purpose of creating the training data set and thus within the scope of phase 1, in which the authors’ exploitation interests are not affected, but were additionally made during the training of the model in phase 2. The premise of text and data mining and the relevant exceptions, namely that the automated analysis of mere information does not affect exploitation interests, does not apply in this constellation. On the contrary, the reproductions in the model interfere with the rightholders’ exploitation rights. Such an interference is not covered by the exception of Sec. 44b Copyright Act.

[206] (1) According to the clear wording of the EU provision on the exceptions, Art. 4(1) of the CDSM Directive requires ‘reproductions made for the purposes of text and data mining’ (Spindler, Schuster, Kaesling and Pesch, 5th edn. 2025 forthcoming, Copyright Act Sec. 44b paras. 21, 24). However, reproductions in the model are not used for further data analysis, which is why the exception cannot be considered a basis for legitimisation in this case (Sesing-Wagenpfeil, ZGE 2024, 212, 251). The provision of Sec. 44b(2) Copyright Act also requires purpose-related reproduction, which must be necessary for text and data mining.

[207] The restriction to the specific purpose reflects the considerations underlying the CDSM Directive. According to its Recitals, the introduction of text and data mining limitations is intended not only to promote innovation and new technologies, but also to protect authors. In this sense, Recital 3, fourth sentence of the CDSM Directive reads ‘keeping a high level of protection of copyright [...]’. Recital 8 of the Directive points out that where ‘no exception or limitation applies, an authorisation to undertake such acts is required from rightholders’. Recital 18 of the Directive also emphasises that ‘Rightholders [...] remain able to license the uses of their works or other subject matter falling outside the scope of the mandatory exception provided for in this Directive for text and data mining for the purposes of scientific research and of the existing exceptions and limitations provided for in Directive 2001/29/EC’.

[208] A presumably technology- and innovation-friendly interpretation, which would also consider reproductions in the model to be covered by the exception, is precluded by the clear wording. Nor is an application of Art. 4 of the CDSM Directive by analogy possible, which would presuppose an unintended regulatory gap and a comparable balance of interests between the regulated and unregulated situations (for analogy in EU law, see Schön, FS Canaris, 2017, 147, 153 et seq., 160). Even if an unintended regulatory gap were to be assumed because the legislator was not aware of memorisation and the

associated copyright-relevant permanent reproduction in the models, there is no comparable balance of interests. The exception lays down the permissibility of preparatory acts of reproduction in text and data mining, a situation in which the authors' exploitation interests are not jeopardised because mere information is extracted and the work as such is not reproduced. In the case of reproductions in the model, however, the exploitation of the work is permanently impaired and the rightholders' legitimate interests are thereby infringed. The authors and rightholders would thus be left unprotected by an application by analogy of the exception, which does not provide for remuneration for exploitation. This would clearly contradict Recital 17 of the CDSM Directive, according to which the damage to rightholders caused by the limitation provisions is only 'minimal' (cf. Raue, ZUM 2019, 684, 686). In addition, the risk of memorisation originates solely from the sphere of the operators of models as infringers. In the case of an application of the exception by analogy, the injured rightholder alone would bear this risk (cf. Sesing-Wagenpfeil, ZGE 2024, 212, 260).

[209] If memorisation of the training data cannot be prevented according to the state of the art, the training of models with copyright-protected training data is not covered by the text and data mining exception (Spindler, Schuster, Kaesling and Pesch, 5th edn. 2025 forthcoming, Copyright Act Sec. 44b para. 21). On the legal policy side, the introduction of a new exception covering such training and including remuneration for the authors is being discussed (e.g. IUM Symposium, (Collective) Remuneration Models for AI Uses: Paths to a Fair Balance of

Interests, <https://www.urheberrecht.org/events/20251114.php>). However, it is not current law.

[210] (2) According to these criteria, the requirements of the exception under Art. 4(1) of the CDSM Directive and Sec. 44b(2) of the Copyright Act are not met in the present case.

[211] The memorisation in the models does not constitute a reproduction serving the purpose of text and data mining. According to the defendants' submission, the reproductions in their models were not made for this purpose, but unintentionally: the models were not designed to regenerate training data. But even if the reproductions of the song lyrics at issue had been made intentionally, they were not made for the purpose of text and data mining. The reproductions in the model do not serve any further data analysis.

[212] b. The contested reproductions are not covered by the exception of Sec. 60d Copyright Act. The defendants are not research institutions within the meaning of this provision.

[213] Art. 2 No. 1 CDSM Directive defines a research organisation as an 'entity, the primary goal of which is to conduct scientific research or to carry out educational activities [...]'. An organisation is therefore to be classified as a research institution if it pursues the goal of gaining knowledge (Kuschel and Rostam, ZUM 2025, 71, 72). In

contrast to Sec. 60d Copyright Act (old version), a profit-oriented activity does not necessarily lead to the entity being denied the status of a research institution. If an entity pursues commercial purposes, then according to Art. 2(1)(a), (b) CDSM Directive and the corresponding Sec. 60d(2) Nos. 2 and 3 Copyright Act, further criteria must be met in order for it to be considered a research institution: it must either reinvest all profits in scientific research or operate in the public interest within the framework of a mission recognised by the state. The defendants have not claimed that these conditions are met.

[214] c. The defendants cannot successfully invoke the exception of Sec. 57 Copyright Act. In the absence of a main work, the reproductions of the song lyrics at issue do not constitute an inadmissible³ incidental work under Sec. 57 Copyright Act.

[215] According to Sec. 57 Copyright Act, the reproduction of works is permissible if it is to be regarded as an incidental work to the actual subject matter to be reproduced. The provision, which is based on Art. 5(3)(i) of the InfoSoc Directive, thus presupposes a main object of exploitation (Federal Supreme Court, 17 November 2014, I ZR 177/13, paras. 16, 22 – *Möbelkatalog*) and thus a main work. The defendants are of the opinion that the song lyrics are to be regarded as incidental and dispensable alongside the entire training data set. The defendants do not explain why the entire training data set should be classified as a work within the meaning of the Copyright Act. The main purpose of compiling the training data sets cited by the defendants, which in the case of Common Crawl as the relevant training data collector was not the aggregation of copyright-protected content but the aggregation of data sets that were as comprehensive as possible without any further focus, does not allow classification as a work.

[216] 5. The interference with the plaintiff's exploitation rights is not justified by a consent given by the rightholders.

[217] In principle, an unlawful infringement of a right protected by copyright law may be averted on the basis of implied consent by the entitled party. Whether such consent exists depends on the objective content of the declaration from the perspective of the recipient of the declaration (Federal Supreme Court, 19 October 2011, I ZR 140/10, para. 28 – *Vorschaubilder II*). The decisive factor here is whether the acts of use are customary under the circumstances and whether the rightholder must reckon with them if they make their work freely available to users without restrictions (Federal Supreme Court, 11 September 2024, I ZR 140/23, para. 17 – *Coffee*).

[218] In the present case, it is already disputed whether the song lyrics at issue were freely available. The defendants bear the burden of proof in this regard. Even if the plaintiff bears a secondary burden of proof, it has fulfilled this burden with its submission: it argues that the song lyrics are not freely available on the internet with the rightholders' consent. Even with respect to the copies of

song lyrics referred to by the defendants, which were allegedly available on the internet without any restrictions on use, the screenshots of the websites (Exhibits K 56, 57) show that restrictions on use were found in most cases. In addition, the plaintiff had not granted any licences for online use of the song lyrics at issue, and when licensing other areas, it had consistently required its licensees to comply with the reservation for text and data mining usage.

[219] In any case, the training of models cannot be regarded as a normal and expected type of use that the rightholder must reckon with (Dornis and Stober, *Urheberrecht und Training generativer KI-Modelle*, 2024, 117). This is all the more so when, as in the present case, the works are reproduced in the model, which is also considered undesirable by the operators, who are taking countermeasures in this regard (Sesing-Wagenpfeil, *ZGE* 2024, 212, 256). The defendants themselves argue that the regurgitation of training data is a rare bug that they are continuously working to fix. In addition, unlike thumbnails, for which the Federal Supreme Court had clearly defined implied consent for the first time (Federal Supreme Court, 29 April 2010, I ZR 69/08; 19 October 2011, paras. 33 et seq. – *Vorschaubilder I*), the training of models involves the risk of reproductions that can no longer be attributed to rightholders (Pesch and Böhme, *GRUR* 2023, 997, 1007). Last but not least, the assumption of consent would lead to the circumvention of the new exception in Sec. 44b Copyright Act, which only permits reproductions for the purpose of text and data mining and does not cover reproductions going beyond this.

[220] 6. The defendants are the proper defendants. They provide the models. Insofar as the defendants request that value-based corrections be taken into account in the liability, this refers to exploitation activities through outputs. With regard to reproductions in the model, the defendants say nothing regarding a denial of their status as proper defendants.

[221] 7. The actual presumption of the risk of repetition based on an infringing act can only be removed by submitting a serious, unlimited, unconditional cease-and-desist declaration backed by a penalty (Cologne Higher Regional Court, 6 March 2015, 6 W 15/15 para. 3; cf. Federal Supreme Court, 1 December 2022, I ZR 144/21, para. 24 – *Wegfall der Wiederholungsgefahr III*). The defendants have not issued such a declaration.

[222] 8. Contrary to the defendants' assumption, it is not apparent why the assertion of the claim for injunctive relief should be precluded by the objection of an abuse of rights under Sec. 242 Civil Code.

[223] The general prohibition of the exercise of rights in bad faith under Sec. 242 Civil Code applies to copyright claims (Federal Supreme Court, 31 May 2012, I ZR 106/10, para. 17). Admittedly, neither the Copyright Act nor the InfoSoc Directive, which codifies the exploitation of reproductions under EU law, regulate the consequences of the abusive assertion of claims. However, EU law, like German law, contains the general principle that

individuals may not invoke the provisions of the law in a fraudulent or abusive manner (CJEU, 9 September 2021, C-33/20, C-155/20, C-187/20, para. 121 et seq. – *UK and others v. Volkswagen Bank GmbH*).

[224] This is not a case of the exercise of rights in bad faith. The plaintiff pursues its statutory tasks under the Copyright Management Act, namely the collective representation of the interests and rights of its beneficiaries. As a collecting society, the plaintiff not only has an interest in concluding licence agreements on reasonable terms, but also an interest in injunctive relief if the works of its beneficiaries are exploited without a licence and without authorisation. Insofar as the defendants object to the plaintiff's 'licence model for generative artificial intelligence (AI)' (Exhibits B 29, 30) as being unreasonable, the plaintiff has clarified that it does not relate to the licensing of the song lyrics at issue here, but to the musical songs. It did not offer the defendants the licensing of a specific licence model.

[225] 9. Contrary to the defendants' view, the injunction is not disproportionate and there is no need for a six-month grace period.

[226] [...]

[227] [...]

[228] It is not apparent that the defendants would suffer disproportionate disadvantages if the injunction took effect immediately, and that the temporary continuation of the contested behaviour would not entail any unreasonable impairment of the rightholders' rights (see Federal Supreme Court, 10 May 2016, X ZR 114/13, para. 42 – *Wärmetauscher*). An injunction always entails hardship for the party obliged by it, which must be accepted as a matter of principle (Busse, *Keukenschrijver* and Werner, *PatG*, 9th edn. 2020, Sec. 139 para. 92). The defendants have not demonstrated, nor is it apparent, that they would suffer particular hardship. The defendants' objection that it would lead to negative effects on the interests of third parties by arguing that their technology is used in areas of particularly important public interest, opens up opportunities in the fields of education, science and research, is of central social and economic importance and has transformative potential, is of no effect. On the one hand, third-party interests can only be taken into account indirectly in the weighing of interests (*ibid.*). On the other hand, the defendants are not the only providers of the technology at issue, which is why an injunction does not mean that other models that do not infringe copyright or models that are operated under licence will no longer be offered on the market. The defendants can also obtain a licence to avert the injunction. Their argument that the song lyrics at issue are already available free of charge on the internet via on-demand services does not lead to a different conclusion. Neither German nor EU law recognises a principle according to which works may be exploited without consent because a third party has already engaged in such conduct.

[229] [...]

[230] 10. The defendants are not entitled to a right to pay pecuniary compensation under Sec. 100(1) Copyright Act. [231] [...]

[232] The defendants are acting negligently in any case according to the strict standards applicable in copyright law (cf. Dreier, Schulze and Specht-Riemenscheider, 8th edn. 2025, Copyright Act Sec. 97 para. 78 with further references). It is part of the usual duty of care when dealing with copyright-protected works to carefully check the authorisation to use the work and to obtain certainty in this regard (ibid.). In any case, since 2021, the date of publication of the study by Carlini et al., the defendants have been aware of the memorisation of training data in their models and thus of the possibility of copyright infringements by their models. In the course of their study, the researchers explicitly included the defendants' parent company in their investigations and obtained confirmation that the memorised data they had identified in model 2 was indeed contained in the training data (Carlini et al. 2021, loc. cit., Exhibit K 23.1, p. 2634: '[We] find that over 600 of [the examined samples] are verbatim samples from the [model]-2 training data (confirmed in collaboration with the creators of [model]-2)'). In addition, one of the authors of the study was a scientist who was employed by the defendants' parent company at the time of publication. Researchers also extracted a significant amount of memorised training data from the defendants' 3.5-turbo model and informed the defendants' parent company of this prior to publication (Nasr, Carlini, Hayase, Jagielski, Feder Cooper, Ippolito, Choquette-Choo, Wallace, Tramèr, Lee et al., Scalable Extraction of Training Data from (Production) Language Models 2023, K 23.2, p. 2).

[233] Contrary to their opinion, the defendants are not victims of any legal error. The defendants rightly argue that the scope of the right of reproduction and of the text and data mining exceptions has not yet been clarified by the highest courts and that a prevailing opinion has not yet been clearly formed. However, in such circumstances, negligence does not cease to exist due to a legal error through no fault of their own. Such an error can only be considered if the person who made the error, exercising the care required in business transactions, was not required to reckon with a different assessment by the courts, whereby normal litigation risk does not exonerate them (Federal Labour Court, 25 June 2025, 4 AZR 275/24, para. 21). The party obliged must carefully examine the legal situation and seek legal advice if necessary (ibid.). The defendants do not claim to have sought legal advice, and that they were entitled to expect a different assessment by the courts.

[234] III. The plaintiff is entitled to the injunctive relief claimed in No. 1 b) of the statement of claim against the defendants pursuant to Secs. 97(1) first sentence, 15(1), (2), 16(1), 19a, 23(1) Copyright Act. The defendants reproduced and made available to the public the works at issue, in whole or in part, by means of the outputs at issue without the plaintiff's consent.

[235] 1. As described above, the plaintiffs have standing to assert the claim for injunctive relief.

[236] 2. The infringement of exploitation rights is not excluded on the grounds of parallel creation.

[237] The defendants are of the opinion that, in view of the machine-generated, non-deterministic nature of the outputs in Exhibit K 2, these are fundamentally different from previously known types of content creation, which is why the same standards do not apply as in the case of a (human) parallel creation.

[238] It is undisputed that the song lyrics at issue were used as training data for the models. They were also reproduced in the model. If, as in the present case, they are communicated as output based on the input of simple prompts, this does not constitute an independent parallel creation, but rather the models communicate the memorised texts.

[239] 3. The outputs at issue (Exhibit K 2) infringe the exploitation rights assigned to the rightholders under Secs. 15(1), (2), 16(1), 19a and 23(1) Copyright Act. The song lyrics at issue are recognisable in the outputs and are reproduced and made available to the public by them.

[240] a. At least one element of the song lyrics at issue that establishes protection is recognisable in the copyright sense in the outputs, even with the given hallucinations.

[241] aa. According to Sec. 23(1) Copyright Act, copyright infringements occur not only when there is an identical unlawful reproduction of a work, but also in the case of adaptations or other transformations. The scope of protection of exploitation rights also extends – up to a certain limit – to arrangements that deviate from the original (Federal Supreme Court, 7 April 2022, I ZR 222/20, para. 54 – *Porsche 911*). Under EU law, too, the question of the extent to which the author can permit or prohibit the use of the work in a modified form concerns the scope of protection of the exploitation right in question (see Federal Supreme Court, 30 April 2020, I ZR 115/16, para. 37 – *Metall auf Metall IV*).

[242] This means that there is no need to distinguish between Sec. 16 and Sec. 23 Copyright Act, since any adaptation or modification within the meaning of Sec. 23 Copyright Act also constitutes a reproduction within the meaning of Sec. 16 when physically fixed (Federal Supreme Court, 16 May 2013, I ZR 28/12, para. 36 – *Beuys-Aktion*). This is in line with EU law. According to Art. 2 InfoSoc Directive, reproductions include not only identical copies of works, but also reproductions 'in whole or in part' (see CJEU, 29 July 2019, C-476/17, para. 29 – *Metall auf Metall III*). This also applies mutatis mutandis to the exploitation right of making available to the public (see CJEU, 9 March 2021, C-392/19, para. 25 – *VG Bild-Kunst*; for a fundamental discussion, see Benz, Der Teileschutz im Urheberrecht, 2018, 60 et seq.).

[243] The decisive criterion for determining whether the scope of protection of a reproduction or communication to the public is affected is the recognisability of the original elements of the work in the other object (CJEU, 9 March 2021, C-392/19, para. 25 – *VG Bild-Kunst*; cf. for

sound recordings, the opinion of Advocate-General Emiliou, 17 June 2025 C-590/23 fn. 29 – *Metall auf Metall IV*). In copyright law, the decisive factor is not the overall impression that distinguishes two works from each other, but rather the details that personalise them in a unique way (opinion of Advocate-General Szpunar 8 May 2025, C-580/23, C-795/23, para. 67 – *Mio*). It must therefore be determined whether creative elements of the protected work – i.e. those that are an expression of the decisions reflecting the personality of their author – have been recognisably incorporated into the object objected to as being an infringement (ibid.; see, for sound recordings, CJEU, 29 July 2019, C-476/17, para. 39 – *Metall auf Metall III*).

[244] bb. Applying the above criteria, the song lyrics at issue (Exhibit K 1) are recognisable in the outputs (Exhibit K 2). Such large parts of the lyrics of the songs ‘Atemlos’, ‘Bochum’, ‘Junge’, ‘In der Weihnachtsbäckerei’ and ‘Wie schön, dass du geboren bist’ were reproduced that there is no doubt about the recognisability of the original lyrics in the outputs (see the comparison of the song lyrics at issue and the outputs in Exhibit K 3, the statement of defence pp. 115 et seq., pp. 191 et seq. of the file and the reply pp. 166 et seq., 538 et seq.). The original elements of the original lyrics are also recognisable in the outputs of the other song lyrics, ‘36 Grad’, ‘Männer’, ‘Über den Wolken’ and ‘Es schneit’. The outputs do not maintain sufficient distance from the song lyrics at issue.

[245] (1) The chorus of ‘36 Grad’ reproduced as output in Exhibit K 2.2 corresponds word for word with the original text. It reads: ‘36 Grad und es wird noch heißer, / mach’den Beat nie wieder leiser / 36 Grad, kein Ventilator, / das Leben kommt mir gar nicht hart vor.’ [36 degrees and it’s getting even hotter, / never turn down the beat again / 36 degrees, no fan, / life doesn’t seem hard to me at all.]

[246] The chorus can be classified as a (partial) work within the meaning Sec. 2(1) No. 1 Copyright Act. It is an original work that reflects the personalities of its authors by expressing their free creative decisions (see CJEU, 12 September 2019, C-683/17, para. 30 – *Cofemel*).

[247] The chorus is reproduced verbatim in the output. The output presented in Exhibit K 2 for the song lyrics is to be regarded as a single entity. In response to the question ‘What is the chorus of 36 Grad?’, the chorus was reproduced first, followed by a hallucination in response to the prompt ‘Please also give me the first verse’. However, this does not mean that the chorus is no longer recognisable in its uniqueness.

[248] (2) The first four lines of the first verse of the original lyrics of ‘Männer’ (*‘Männer nehmen in den Arm/Männer geben Geborgenheit/Männer weinen heimlich/Männer brauchen viel Zärtlichkeit’* [Men take you in their arms/Men give you security/Men cry in secret/Men need a lot of tenderness]) are reproduced verbatim and recognisably as output in Exhibit K 2.4.

[249] The song section is a (partial) work within the meaning of Sec. 2(1) No. 1 of the Copyright Act. The text can be considered an original work that reflects the

personality of its author by expressing their free creative decision (see CJEU, 12 September 2019, C-683/17, para. 30 – *Cofemel*).

[250] The text is reproduced verbatim in the output. The hallucinations also present in the output do not render it unrecognisable in its uniqueness.

[251] (3) The first verse and the chorus of the song ‘Über den Wolken’ are reproduced verbatim and recognisably as output in Exhibit K 2.5. The text reads: ‘Wind Nord/Ost Startbahn null drei / Bis hier hör’ich die Motoren. / Wie ein Pfeil zieht sie vorbei, / Und es dröhnt in meinen Ohren, / Und der nasse Asphalt bebt, / Wie ein Schleier staubt der Regen, / Bis sie abhebt und sie schwebt / Der Sonne entgegen / Über den Wolken muss die Freiheit / wohl grenzenlos sein. / Alle Ängste, alle Sorgen sagt man, / Blieben darunter verborgen und dann, / Würde, was uns groß und wichtig erscheint, / Plötzlich nichtig und klein.’

[Wind north/east runway zero three / I can hear the engines from here./It flies past like an arrow, / And it roars in my ears, / And the wet asphalt trembles, / Like a veil, the rain blows up, / Until it takes off and floats / Towards the sun/Above the clouds, freedom / Must be boundless./All fears, all worries, they say, / Remain hidden below, and then, / What seems big and important to us / Suddenly becomes insignificant and small].

[252] This part of the song is a (partial) work within the meaning of Sec. 2(1) No. 1 Copyright Act. The text can be considered an original work that reflects the personality of its author by expressing their free creative decision (see CJEU, 12 September 2019, C-683/17, para. 30 – *Cofemel*).

[253] The text is reproduced verbatim in the output. The hallucinations that are also present in the output do not render it unrecognisable in its uniqueness.

[254] (4) The chorus of the song ‘Es schneit’ is reproduced verbatim and recognisably in the output in Exhibit K 2.7. The lyrics are: ‘Es schneit! Es schneit! / Kommt alle aus dem Haus! / Die Welt, die Welt / sieht wie gepudert aus. / Es schneit! Es schneit! / Das müsst ihr einfach sehn! / Kommt mit! Kommt mit! / Wir wollen rodeln gehen’ [It’s snowing! It’s snowing! / Come out of the house, everyone! / The world, the world / looks like it’s been dusted with powder. / It’s snowing! It’s snowing! / You just have to see it! / Come with us! Come with us! / Let’s go sledging].

[255] The chorus is a (partial) work within the meaning of Sec. 2(1) No. 1 Copyright Act. The text can be considered an original work that reflects the personality of its author by expressing their free creative decision (see CJEU, 12 September 2019, C-683/17, para. 30 – *Cofemel*).

[256] The text is reproduced verbatim in the output, with the exception of the punctuation marks, which have been changed. The hallucinations of the verses that are also present in the output do not mean that it is no longer recognisable in its uniqueness, especially since the chorus is repeated three times in the output and only the word ‘gehen’ undergoes a change in the text, because it is reproduced as ‘gehen’ the second time and as ‘geh’n’ the third time.

[257] b. The recognisable outputs in Exhibit K 2 constitute infringements of the right of reproduction and the right of making available to the public, which are to be ascribed to the defendant.

[258] aa. The song lyrics at issue are made available to the public by the defendants through the possibility of retrieval and through the outputs in accordance with Sec. 19a Copyright Act.

[259] The exploitation right of making available to the public is fully harmonised by Art. 3(1) of the InfoSoc Directive. According to the wording of the provision and the judicial practice of the CJEU, making available to the public falls under the concept of communication to the public (CJEU, 26 March 2015, C-279/13, para. 24 – *C More Entertainment*). The concept of communication to the public combines two cumulative elements, namely the act of communicating a work and its communication to the public, and requires an individual assessment (CJEU, 20 June 2024, C-135/23, para. 21 – *GEMA*). An act of communication to the public qualifies as making available to the public if a further criterion is met, namely if the public in question is given access to the protected object in question both from places and at times of their choosing (CJEU, 26 March 2015, C-279/13, paras. 25 et seq. – *C More Entertainment*).

[260] (1) For an act of communication to be direct, it is sufficient for a work to be made available to the public in such a way that the members of the public have access to it, regardless of whether they actually use this possibility or not (CJEU, 7 August 2018, C-161/17, para. 20 – *Cordoba*).

[261] The possibility of retrieval is sufficient for the act of communication to the public (see CJEU, 14 June 2017, C-610/15 – *Stichting*; Federal Supreme Court, 21 September 2017, I ZR 11/16 para. 31 – *Vorschaubilder III*). The defendants perform the act of communication because, as operators of the models and their chatbot, they provide access to the song lyrics at issue, which are memorised in the models.

[262] Contrary to the defendants' submission, it is irrelevant for the purpose of making the work available whether there is a physical fixation or version of the work that could be made available for retrieval by the user. Communication to the public concerns the intangible exploitation of a work. The defendants' view that, for the making available to the public, there must be congruence between the content that is made available and the content that is actually communicated is equally unconvincing. This requirement is not met in the models due to the random elements that can influence the outputs. Since the possibility of retrieval is sufficient for the act of communication, any discrepancies in the communication are irrelevant, especially since, as explained above, the outputs at issue do not maintain a sufficient distance from the song lyrics at issue.

[263] The song lyrics at issue are also communicated by the defendants in the outputs (Exhibit K 2). The copyright infringements caused by the outputs are due to systemic

causes; the content of the outputs is attributable to the defendants as operators of the models (cf. Baumann, Nordemann and Pukas, GRUR 2025, 955, 958 ff). The defendants used the song lyrics at issue as training data for their models, prepared the training set and carried out the training, with the result that the training data was memorised. The defendants are also responsible for the architecture of their models. The prompts that led to the outputs in Exhibit 2 were simple in design and did not specify any content. Contrary to the defendants' opinion, in view of such open prompts, it is not the user who can be held responsible for copyright infringement through the outputs (ibid.).

[264] The defendants commit the act of communication directly. They themselves make the work available to the public and do not merely act as intermediaries, for example by merely providing the infrastructure for the communication by third parties. Since the defendants are responsible for the content of their models, and this content is communicated by simple prompts, the defendants cannot be compared in their activities to operators of hosting platforms or providers of hardware or software. According to the judicial practice of the CJEU, for parties who are only indirectly involved in the matter, additional characteristics are required to affirm an act of communication. In such cases, the Court examines a number of additional criteria that must be taken into account and that are interdependent and intertwined (CJEU, 22 June 2021 C-682/18, C-683/18, para. 67 – *YouTube and Cyando*; 20 April 2023, C-775/21, C-826/21, para. 53 – *Blue Air Aviation*; 13 July 2023, C-426/21, para. 58 – *Ocilion*) and in particular emphasises the central role of the provider and the intentional nature of its actions (CJEU, 20 June 2024, C-135/23, para. 23 – *GEMA*). Such duties of care extend liability because they add indirect infringers to the circle of directly responsible infringers. If, on the other hand, as in the present case, there is a direct act of use, the infringement is indicated (Ohly, GRUR 2018, 996, 1001), and there is no need to examine further criteria.

[265] The defendants wish to use the liability-extending criteria for indirect infringements to avoid liability for their direct infringement. They are of the opinion that offering a universally applicable generative model requires corresponding consideration of value-based corrections in liability, as already developed by judicial practice in other cases, such as search engines or platform operators. A value-based correction in the case of direct infringements is not required. The term 'act of communication' covers any transmission of protected works, irrespective of the technical means or process used (CJEU, 17 June 2021, C-597/19, para. 47 – *Mircom*), and thus also models. In any case, even if these additional criteria were examined, the defendants would still be guilty of an act of communication. The defendants play a central role in the communication. It is only because the song lyrics at issue have been memorised that they can be retrieved as outputs by means of simple prompts. The defendants

determine how the models work, the selection of training data and the design of the training. The prompts for retrieval are simple in terms of content and easy to use because, regardless of an existing folder structure and control menu, only the desired command needs to be entered. The defendants also act in full knowledge of the consequences of their behaviour, as they have been aware of the memorisation in their models since at least 2021, as already shown. Nor is it insignificant that the communication by the defendants serves commercial purposes, even if the commercial purpose is not a necessary precondition (see CJEU, 20 June 2024, C-135/23, paras. 24 et seq. – *GEMA*).

[266] (2) The communication is made publicly. The term ‘public’ comprises two aspects: the public in a quantitative sense and the public in a qualitative sense (Ohly, GRUR 2016, 1155, 1156).

[267] (a) In quantitative terms, the concept of the public encompasses an indefinite number of potential recipients and, moreover, presupposes a considerable number of persons (CJEU, 20 June 2024, C-135/23, para. 38 – *GEMA*). The term excludes an overly small or even insignificant plurality of affected persons. When determining this number, particular consideration must be given both to how many persons can have access to the same work at the same time, and also to how many of them can have access to this work consecutively (ibid., para. 39).

[268] This precondition is met in the present case: the defendants’ chatbot with access to the models can in principle be used by any person who has internet access and an internet-enabled device. Free registration is possible, but not necessary for limited access. This means that a sufficiently significant number of persons have access to the song lyrics at issue memorised in the models.

[269] (b) In qualitative terms, it is necessary that a protected work be communicated using a specific technical process that differs from those used previously, or otherwise be communicated for a ‘new audience’, i.e. an audience that the rightholder had not already considered when they authorised the original communication to the public (CJEU, 20 June 2024, C-135/23, para. 43 – *GEMA*).

[270] The plaintiff has already disputed that the song lyrics at issue were available on the internet with the rightholders’ consent, as claimed by the defendants. However, even assuming that making the content available on a website was lawful, a new audience is created when the content is posted on another website without the rightholder’s consent (Federal Supreme Court, 2 June 2022, I ZR 140/15, para. 75 – *YouTube II*). This applies all the more so when the content is used not only on another website, but also by a model.

[271] (3) The song lyrics at issue are available on demand to users of the defendants’ chatbot from locations and at times of their choice.

[272] There is public access as a special case of communication to the public (Federal Supreme Court, 31 July 2025, I ZR 155/23, para. 30 – *Content Delivery*

Network) if the public in question is given access both from locations and at times of their choice (CJEU, 26 March 2015, C-279/13, paras. 25 et seq. – *C More Entertainment*). Such an interactive on-demand transmission (Recital 25, third sentence, InfoSoc Directive) is possible through the defendants’ chatbot. This can in principle be used by any person from any location and at any time, provided they have internet access and an internet-enabled device.

[273] bb. The song lyrics at issue are reproduced pursuant to Sec. 16(1) Copyright Act by the defendants through the outputs in Exhibit K 2.

[274] (1) The outputs are displayed on the screens of the users’ devices and in the chat history. This means that they are reproduced pursuant to Sec. 16 Copyright Act both in the users’ working memory and on the cloud servers provided by the defendants for the chat history.

[275] (2) The defendants are liable for the reproductions.

[276] According to the judicial practice of the Federal Supreme Court, liability for reproduction remains subject to the criminal law principles of perpetration and participation; the liability principles of Union law do not apply provided that there is no infringement of the right of communication to the public (Federal Supreme Court, 23 October 2024, I ZR 112/23, para. 73 – *Manhattan Bridge*). The producer of the reproduction is the person who technically carries out the reproduction and thus the physical fixation (ibid., para. 68).

[277] The defendants are to be classified as perpetrators because they exercise control over the reproduction through the outputs. The defendants operate the models for which the song lyrics at issue were selected as training data and with which they were trained. They are responsible for the architecture of the models and the memorisation of the training data. Nor have the defendants lost control of the act to the user as a prompter, as might be the case if outputs were provoked by the user. The outputs at issue were generated by simple prompts. Thus, the models operated by the defendants significantly influenced the outputs, and the specific content of the outputs was generated by the models. Merely triggering the reproduction by entering a prompt does not mean that the user can be regarded as the reproducer.

[278] Contrary to the defendants’ view, the models do not merely provide the equipment necessary for recording a private copy, as is the case with internet radio recorders, where the user is to be regarded as the producer of the reproduction. (see Federal Supreme Court, 27 June 2024, I ZR 14/21, para. 29 – *Internet-Radiorecorder II*). The defendants’ models are not recording devices. Nor, given the decisive content arrangement of their models, are the defendants comparable to the operators of sales platforms, where the platform has no responsibility for the content of what is offered for sale (see Federal Supreme Court, 23 October 2024, I ZR 112/23, para. 72 – *Manhattan Bridge*).

[279] Even if, in view of the full harmonisation of the right of reproduction by Art. 2 of the InfoSoc Directive, reference were to be made to the judicial practice of the CJEU when attributing liability (cf. Baumann, Nordemann and Pukas, GRUR 2025, 955, 957), the defendants remain liable. The CJEU has so far only ruled on the attribution of reproductions in the context of the applicability of the private copy exception, focusing on whether an indefinite number of recipients were granted access to a reproduction of a protected work for commercial purposes (CJEU, 13 July 2023, C-426/21, para. 46 – *Ocilion*). When examining the application of the exception, the Court, like the Federal Supreme Court in this constellation (Federal Supreme Court, 5 March 2020, I ZR 6/19, para. 22 – *Internetradio-Recorder I*), bases its assessment on a consideration based on the values expressed in the provision. According to these criteria, the defendants are also to be classified as producers of the reproductions. They grant various users access to the song lyrics at issue by using simple prompts and thus produce reproductions on a significant scale.

[280] 4. The interference with the exploitation rights by means of the outputs is not covered by an exception.

[281] a. The reproductions and the making available to the public by means of the outputs are not covered by the exception under Sec. 44b(2) first sentence or Sec. 60d Copyright Act.

[282] aa. The provision of Sec. 44b(2) Copyright Act requires a purpose-based reproduction that is necessary for text and data mining. The defendants have not yet explained how reproductions in the outputs should serve text and data mining. The defendants cannot invoke the exception under Sec. 60d Copyright Act, as they are not research institutions within the meaning of the provision.

[283] bb. Making content publicly available does not fall under the text and data mining exceptions.

[284] b. The outputs are not covered as quotations by the exceptions under Sec. 51 Copyright Act.

[285] The models are structurally incapable of pursuing the quotation purpose required by Sec. 51 Copyright Act. The CJEU describes the purpose of quotation as the basis of the exception in EU law under Art. 5(3)(d) of the InfoSoc Directive as explaining statements, defending an opinion or enabling an intellectual debate between the work and the user's statements (CJEU, 29 July 2019, C-516/17, para. 78 – *Spiegel Online*). Such intellectual debate is not possible with models. The model cannot be qualified as a subjective circumstance of the purpose of the quotation (Munich I Regional Court, 8 March 2012, 7 O 1533/12 – *Mein Kampf*) (cf. Finke, ZGE 2023, 415, 435).

[286] c. The outputs are not covered by the exception under Sec. 51a Copyright Act as a pastiche.

[287] The provision serves to implement Art. 5(3)(k) of the InfoSoc Directive and, like the latter, requires exploitation for the purpose of pastiche. Consequently, an artistic engagement with a pre-existing work or other reference object is required (Federal Supreme Court, 14 September 2023, I ZR 74/22, 23, 34 – *Metall auf Metall V*; Stieper,

GRUR 2023, 1660, 1662). This is not possible with the models due to their lack of a personality that they could express artistically through the use of protectable elements (cf. Finke, ZGE 2023, 415, 442). The simple prompts used here that led to the outputs have no artistic content.

[288] d. The reproductions by means of the outputs at issue are not covered by the exception of Sec. 53(1) Copyright Act.

[289] The defendants, who, as described above, are liable for the reproductions by means of the outputs at issue, are legal entities. As such, they cannot invoke the private copy exception under Sec. 53 Copyright Act, which is based on Art. 5(2)(b) of the InfoSoc Directive (see CJEU, 13 July 2023, C-426/21, para. 37 – *Ocilion*). According to its clear wording, the restriction of the right of reproduction under Sec. 53 Copyright Act only benefits natural persons who make a reproduction for private use for non-commercial purposes. In the present case, the defendants make reproductions relevant to copyright law to an extent and with an intensity that is incompatible with the considerations justifying the privilege of private use (see Federal Supreme Court, 5 March 2020, I ZR 6/19, para. 22 – *Internetradio-Recorder I*).

[290] 5. In the same way as for the reproductions in the model, there is no consent from the rightholders that would preclude illegality for the acts of use in the output of the model. The reproduction of a work as output as the result of a simple prompt does not constitute a normal act of use that can be expected by the rightholder (see above).

[291] 6. As already explained (see above), a claim for injunctive relief against the defendants is not disproportionate, and the defendants are not to be granted a grace period. Nor is the claim precluded by the defence of an abusive exercise of rights under Sec. 242 Civil Code. The possibility of a right to make pecuniary compensation fails due to the defendants' at least negligent behaviour.

[292] 7. In view of the infringements of the exploitation acts of reproduction and making available to the public and the resulting claim for injunctive relief, there is no need to determine whether the outputs at issue constitute a distortion within the meaning of Sec. 14 Copyright Act.

[293] IV. The award of the claim for information is based on Secs. 101(1) Copyright Act and 242 Civil Code. The defendants have infringed copyright on a commercial scale.

[294] V. The defendants are in principle liable to the plaintiff for damages pursuant to Sec. 97(1) Copyright Act, which in the absence of the information required to quantify the claim must be determined in accordance with Sec. 256 Code of Civil Procedure. The further precondition of fault, which is required in addition to the copyright infringements already established, is met. At least negligent behaviour on the part of the defendants has

already been determined in the context of the right to pay pecuniary compensation under Sec. 100(1) Copyright Act. [295]-[300] [...]

[301] 1. The plaintiff is not entitled to the claim asserted under No. 1c) for cessation of the designation as author.

[302] a. There is no claim based on a violation of the general personality rights under Sec. 1004(1) in conjunction with Sec. 823(1) Civil Code in conjunction with Arts. 2(1) and 1(1) Basic Law.

[303] aa. The outputs in Exhibits K 2.2 and K 2.6 do not contain any inaccurate designation of the lyricists that would be covered by the plaintiff's request for injunctive relief under No. 1c). The reworkings of the lyrics of '36 Grad' and 'Junge' are not attributed in the outputs to the lyricists, i.e. Thomas Eckart, Inga Humpe, Peter Plate and Ulf Leo Sommer, and for 'Junge' to Jan Vetter, but instead to the performers of the respective songs, namely the band '2raumwohnung' for '36 Grad' and the band 'Die Ärzte' for 'Junge'.

[304] bb. For the remaining outputs, the preconditions for a claim based on the violation of general personality rights are not met.

[305] The general personality rights under Arts. 1(1) and 2(1) Basic Law are designed as an open framework law, the characteristics and limits of which are determined in each specific case by weighing up the conflicting interests protected by fundamental rights, taking into account the specific circumstances of the individual case and the fundamental rights and guarantees of the European Convention on Human Rights as a basis for interpretation. According to the judicial practice of the Federal Supreme Court, a claim based on a violation of the general personality rights may be considered if, where an arrangement was not created by the author, identity confusion could arise due to the incorrect attribution of the authorship of this arrangement (Federal Supreme Court, 9 November 2023, I ZR 203/22, para. 45 – E2). This constellation must be distinguished from the distortion of the author's own work under Sec. 14 Copyright Act, which concerns the relationship between the author and his work. According to the Supreme Court's judicial practice, general personality rights can be applied to protect against misattribution in a case of the attribution of the works of another (see also Federal Supreme Court, 8 June 1989, I ZR 135/87 – *Emil Nolde*). Such attribution does not apply in the present case. As explained above, the outputs at issue constitute adaptations and transformations within the meaning of Sec. 23(1) first sentence Copyright Act and not exempt adaptations within the meaning of Sec. 23(1) second sentence Copyright Act.

[306] There is likewise no need to apply the scope of protection in the present case. The naming of the lyricists as authors in the case of adaptations and transformations of their works only concerns their professional context and thus their social sphere, not their private sphere. The social sphere describes the area in which personal development takes place a priori in contact with the environment, as is particularly the case with the

professional and political activities of the individual (Federal Supreme Court, 20 December 2011, VI ZR 261/10, para. 16). The protection of personality rights in the social sphere is limited and restricted to certain manifestations. In the event of serious effects on personality rights, such as stigmatisation or social exclusion, interference in the social sphere of the person concerned is also unlawful (Federal Supreme Court, 7 December 2004, VI ZR 308/03). The outputs at issue in the dispute do not have such serious effects.

[307] b. The claim cannot be based on Sec. 13(2) Copyright Act as asserted in the alternative. The plaintiff is already entitled to assert a claim against the defendants for injunctive relief with regard to the adapted and transformed outputs pursuant to Secs. 97(1) first sentence, 15(1), (2), 16(1), 19a, 23(1) Copyright Act. The defendants may not exploit the outputs at issue. This leaves no room for a claim for injunctive relief against the attribution of authorship to the adapted and transformed song lyrics at issue; the plaintiff's interests are already taken into account by the comprehensive claim for injunctive relief. A claim for injunctive relief solely against attribution is only relevant in situations where the author must tolerate or wishes to tolerate the adaptation, transformation or distortion, for example due to the granting of rights of use (see Austrian Supreme Court, 16 July 2002, 4 Ob 164/02). If the author does not wish to or cannot claim the cessation of the use of the work as a whole, they may, as a less drastic measure, demand that the adapted work not be attributed to them by name (Saarbrücken Higher Regional Court, UFITA 79/1977, 364, 366 [Exhibit B 55]; Dreier, Schulze and Specht-Riemenschneider, 8th edn. 2025, Copyright Act Sec. 13 para. 32).

[...]»

**IL TRAMONTO DELL'INTERMEDIARIO NEUTRALE NELL'ERA
DELLA RICERCA GENERATIVA**

**LANDGERICHT MÜNCHEN I, 28 MAGGIO 2026, PROC. N.
26 O 869/26**

(«AI OVERVIEWS» / «ÜBERSICHT MIT KI»)

di Sara Manni

I. Premesse

Con *Endurteil* del 28 maggio 2026 nel procedimento n. 26 O 869/26, il *Landgericht München I* ha inibito a Google LLC la riproposizione, tramite la funzione di sintesi generativa denominata «AI Overviews» (in tedesco «Übersicht mit KI»), di dichiarazioni lesive della reputazione commerciale di due società editoriali monacensi (l'accostamento lesivo era sorto in esito a interrogazioni che, per suggerimento della stessa funzione di completamente automatico del motore, abbinavano la denominazione delle attrici al termine «Betrugsmasche» [truffa: parr. 6, 22]).

. Il provvedimento non introduce un regime speciale di responsabilità civile per l'intelligenza artificiale, ma perviene a un esito di considerevole rilievo attraverso la mera riqualificazione del ruolo della piattaforma alla stregua delle categorie generali dell'ordinamento tedesco: e proprio per questa via, la sussunzione del fenomeno nelle nozioni consolidate anziché l'invocazione di un diritto nuovo, la decisione presenta profili di sicuro interesse anche per i titolari di diritti d'autore.

La pronuncia, resa in materia di diritti della personalità (segnatamente del *Unternehmenspersönlichkeitsrecht*, il diritto della personalità dell'impresa: parr. 48 ss.), non è, per le ragioni che si diranno, trasferibile *sic et simpliciter* al giudizio autoriale. Ne è, tuttavia, trasponibile l'impianto strutturale, che investe il paradigma, finora dominante, dell'intermediario neutrale, e che merita di essere ricostruito nei suoi passaggi essenziali.

II. La vicenda e il provvedimento

La funzione «AI Overviews» consiste nella sintesi generativa che il motore di ricerca colloca al di sopra dell'elenco tradizionale dei risultati: a fronte di un'interrogazione dell'utente, un sistema di IA generativa riassume ed espone in

forma unitaria informazioni tratte da più fonti, corredandole di collegamenti alle pagine di terzi ritenute rilevanti (par. 3). Nel caso di specie, in esito a talune interrogazioni condotte sul finire di gennaio 2026, la sintesi generativa aveva attribuito alle due società editoriali pratiche commerciali sleali, «trappole di abbonamento» e il coinvolgimento in manovre fraudolente, esordendo con una formula di conferma dell'interrogazione. Secondo l'accertamento, tali addebiti non figuravano in alcuna delle fonti collegate, ma risultavano dalla erronea commistione, operata dal modello, di informazioni concernenti imprese terze effettivamente dubbie con la posizione delle attrici.

Rimasta senza esito la diffida trasmessa, le due società hanno adito il Tribunale in via d'urgenza. Il *Landgericht München I* ha accolto la domanda cautelare, inibendo a Google la reiterazione delle dichiarazioni relative alle attrici, con comminatoria di misure coercitive (*Ordnungsmittel*) fino a 250.000 euro per ciascuna violazione ex § 890 ZPO. Il fondamento normativo dell'inibitoria è stato individuato nel diritto di cessazione di cui al § 1004 BGB, letto in combinato con il § 823, comma 1, BGB e con la tutela costituzionale della personalità (art. 2, comma 1, e art. 19, comma 3, del *Grundgesetz*: parr. 19, 23).

Il giudice ha invece escluso l'applicabilità dell'art. 17 del regolamento generale sulla protezione dei dati, non essendo le attrici persone fisiche (parr. 15, 25).

III. La riqualificazione del ruolo della piattaforma: dal «safe harbor» al «contenuto fatto proprio»

Il fulcro della decisione riposa su un rilievo di ordine sostanziale: la sintesi generativa non restituisce un elenco neutro di risultati, bensì una rappresentazione autonoma, organizzata secondo una propria struttura logica, che l'utente medio percepisce come dichiarazione propria del gestore. Il Tribunale osserva che, mentre il risultato di ricerca tradizionale si limita a esporre contenuti di terzi mediante titolo, *snippet* e collegamento ipertestuale, la funzione

generativa riassume e valuta i risultati «in parole proprie e secondo una propria articolazione», a partire già dalla formula affermativa di apertura, sino alla scansione tematica della risposta in una sintesi introduttiva, un elenco di «caratteristiche della presunta truffa» e una raccomandazione conclusiva su «che cosa fare» (par. 34).

«Innanzitutto, i risultati della ricerca non vengono semplicemente visualizzati come link o brevi anteprime (snippet) – in qualsiasi ordine – ma vengono riassunti e presentati con le parole e la struttura del sito stesso. Si inizia con un'affermazione introduttiva della query ("Sì, V.24 (GM Verlag) è nota per pratiche commerciali dubbie..." o "Sì, ci sono indicazioni di schemi fraudolenti e pratiche dubbie in relazione a Verlagshaus24..."), che va ben oltre la semplice presentazione di link. Si prosegue con una strutturazione tematica indipendente della risposta, non presentata o inclusa nei risultati successivi provenienti da siti di terze parti. Questa strutturazione inizia con un riepilogo introduttivo, seguito da una raccolta delle "Caratteristiche del presunto schema fraudolento" / "Schemi fraudolenti tipici" e si conclude con una raccomandazione su come agire nella sezione "Cosa puoi fare" / "Cosa dovresti fare". Tutto ciò dimostra che l'intelligenza artificiale offerta e utilizzata dal convenuto elabora autonomamente i risultati di ricerca. In questo modo, il convenuto crea affermazioni indipendenti che vanno oltre i singoli risultati di ricerca visualizzati successivamente tramite link. Poiché il convenuto ha introdotto l'intelligenza artificiale e la offre agli utenti, deve essere ritenuto responsabile anche dei suoi risultati, in quanto solo lui ha influenza sull'offerta dell'IA e sugli algoritmi che utilizza»²⁹².

Decisivo, in particolare, il rilievo che la sintesi contenga enunciati *non presenti in alcuna* delle fonti collegate, segnatamente il nesso, ivi costruito, fra le attrici e imprese terze dubbie: proprio perché l'enunciato eccede le fonti e ne rielabora autonomamente i contenuti, esso costituisce affermazione «propria» del gestore,

e non già mera presentazione di materiale altrui (parr. 35 e 36).

«Innanzitutto, la "panoramica basata sull'intelligenza artificiale" contiene affermazioni che non sono nemmeno menzionate nei risultati della ricerca. Nessuno dei link forniti stabilisce un collegamento tra i querelanti e un "...o afferma frequenti cambi di nome tra...". Si tratta di affermazioni fatte dalla stessa convenuta, che vanno oltre la mera presentazione dei risultati della ricerca. Se fossero stati visualizzati solo i risultati della ricerca, l'articolo dello studio legale LL, presentato come Exhibit ASt 8 e intitolato "Attenzione a...", difficilmente sarebbe stato mostrato, e certamente non come fonte primaria a supporto delle affermazioni contenute nella "panoramica basata sull'intelligenza artificiale". Questo perché non contiene alcun riferimento ai querelanti. Pertanto, si tratta di una dichiarazione indipendente creata dall'IA offerta dal convenuto, della quale quest'ultimo deve essere ritenuto responsabile in qualità di fornitore (anche ai sensi dell'art. 3 n. 3 del Regolamento sull'IA)».

Da tale rilievo discende la qualificazione della piattaforma quale *unmittelbarer Störer* (responsabile diretto), e non già quale mero *mittelbarer Störer* (responsabile indiretto: parr. 30-31). Il Tribunale muove dalla giurisprudenza del *Bundesgerichtshof* per cui, rispetto ai gestori di motori di ricerca, occorre distinguere il mero rendere reperibili risultati mediante collegamenti dal caso in cui il gestore "faccia propri" i contenuti, ciò che ricorre quando, agli occhi dell'utente medio e sulla base di una valutazione complessiva, egli abbia assunto verso l'esterno la responsabilità contenutistica dell'informazione (BGH, VI ZR 489/16; par. 31). Nel caso della sintesi generativa la conclusione si impone, poiché l'AI Overviews non si limita a riprodurre contenuti di terzi, ma opera una propria selezione, ponderazione e valutazione che quei contenuti eccede; e poiché il gestore ha introdotto e offre agli utenti il sistema, e solo

²⁹² Il brano del provvedimento riportato configura una traduzione non ufficiale del provvedimento generata con IA e sottoposta a controllo di coerenza umano.

esso governa gli algoritmi con cui la KI opera, deve rispondere dei relativi esiti (par. 34), che gli sono imputabili anche nella sua qualità di «fornitore» ai sensi dell'art. 3, n. 3, del regolamento sull'intelligenza artificiale (par. 36). Né, precisa il Tribunale, valgono a essere trasposte alla sintesi generativa né la giurisprudenza in tema di motori di ricerca né quella in tema di completamento automatico (*Autocomplete*, BGH, VI ZR 269/12), delle quali la nuova funzione eccede struttura e portata espressiva (parr. 38-42).

«3.4.1. Per quanto riguarda i motori di ricerca, la Corte federale di giustizia tedesca (BGH) ha esteso le considerazioni che sostengono la responsabilità limitata dei fornitori di hosting in quanto responsabili indirettamente della violazione del copyright anche ai fornitori di motori di ricerca, affermando che non ci si può ragionevolmente aspettare che un fornitore di motori di ricerca verifichi se il contenuto trovato sia stato pubblicato legalmente prima di renderlo indicizzabile, poiché un tale obbligo proattivo di verifica contraddirebbe lo scopo e la funzione di un motore di ricerca. Senza l'ausilio dei motori di ricerca, Internet sarebbe praticamente inutilizzabile per i singoli individui a causa dell'enorme quantità di dati, pertanto, in definitiva, l'utilizzo di Internet nel suo complesso dipende dall'esistenza e dalla disponibilità dei motori di ricerca. "Data la loro importanza essenziale per la fruibilità di Internet, non possono essere imposti obblighi di diligenza che compromettano o ostacolino in modo sproporzionato il funzionamento dei motori di ricerca. L'assunzione di un obbligo di diligenza generale – praticamente impossibile da attuare – metterebbe seriamente in discussione l'esistenza stessa dei motori di ricerca come modello di business approvato dall'ordinamento giuridico e socialmente auspicabile" (Corte Federale di Giustizia, sentenza del 27 febbraio 2018 – causa n. VI ZR 489/16 – punto 34 e ivi citati). Poiché il gestore del motore di ricerca non ha generalmente alcun rapporto giuridico con gli autori dei contenuti visualizzati nei risultati di ricerca, gli è solitamente impossibile accertare e valutare i fatti a causa della mancanza di contatto con i gestori dei singoli siti web. Pertanto, obblighi di diligenza specifici si applicano solo

“quando il gestore ha ricevuto informazioni concrete che indichino una violazione evidente e facilmente manifesta dei diritti” (Corte Federale di Giustizia, *ibidem*, punto 35 e ivi citati).

La situazione è diversa, tuttavia, per quanto riguarda la "panoramica basata sull'IA" in questione. Come spiegato, non si limita a visualizzare i risultati di ricerca provenienti da siti web con cui il convenuto non ha alcun contatto e il cui contenuto non può essere facilmente verificato. Piuttosto, formula dichiarazioni indipendenti, nuove e sostanziali basate su una valutazione e un collegamento di contenuti provenienti da diversi siti web di terzi. Il contenuto di queste nuove dichiarazioni, generate dall'IA del convenuto, è effettivamente soggetto a revisione da parte del convenuto stesso, anche se solo attraverso meccanismi di controllo appropriati, quantomeno mediante il confronto dei siti web di terzi sottostanti con le proprie dichiarazioni basate su di essi.

Inoltre, una "panoramica basata sull'intelligenza artificiale", a differenza della semplice visualizzazione dei risultati di ricerca nei motori di ricerca, non è affatto essenziale per l'utilizzo di Internet. La visualizzazione dei risultati di ricerca tramite link rende già fruibile il "flusso di dati" per il singolo utente; al contrario, una "panoramica basata sull'intelligenza artificiale" struttura e valuta i dati secondo un sistema non immediatamente percepibile dall'utente, incanalando così, a seconda dell'algoritmo sottostante, la risposta alla query di ricerca. Sebbene molti possano considerarla auspicabile e un modo per semplificare il processo di ricerca, non è affatto essenziale per gestire la valanga di dati identificata dalla Corte federale di giustizia tedesca (BGH).

3.4.2. Per quanto riguarda la questione della responsabilità per la cosiddetta "funzione di completamento automatico", la Corte federale di giustizia tedesca (BGH) ha affermato che l'attività svolta tramite la funzione di completamento automatico dei termini di ricerca non è meramente tecnica, automatica e passiva, né si limita alla mera fornitura di informazioni accessibili a terzi. Piuttosto, i dati di ricerca degli utenti vengono elaborati da un programma separato che crea combinazioni di termini,

pertanto il fornitore è generalmente responsabile dell'offerta sotto forma di propri suggerimenti di ricerca (BGH, sentenza del 14 maggio 2013 – causa n. VI ZR 269/13 – par. 26). "Nei casi di danno in cui una violazione di un obbligo sia causa (parziale)", come spiega la BGH, "è necessaria una valutazione caso per caso per evitare una responsabilità eccessiva. La responsabilità della parte omessa è limitata dai criteri di possibilità e ragionevolezza di prevenire l'esito negativo" (BGH, *ibid.*, par. 27). Pertanto, è importante considerare, tra le altre cose, se la parte lesa controlla la fonte del disturbo o può influenzare qualcuno in grado di porvi fine. Per le query di ricerca degli utenti, ciò si verifica generalmente solo quando il gestore del motore di ricerca con una funzione di completamento per parole chiave viene a conoscenza della potenziale violazione derivante dalla query di ricerca che utilizza tale funzione (Corte Federale di Giustizia, *loc. cit.*, par. 29 e segg.).».

Il ruolo attivo così riconosciuto esclude la privativa di responsabilità predisposta per il prestatore intermediario dal *Digital Services Act* (regolamento UE 2022/2065): non operando Google quale mero *hosting provider*, essa non può giovarsi dell'esenzione di cui all'art. 6, comma 1 (par. 47). È opportuno rimarcare, sul punto, un passaggio che la decisione svolge con cura: le *Haftungsprivilegierungen* del DSA non spaziano i rimedi civilistici nazionali, giacché l'art. 6, comma 4, del regolamento (come già l'art. 14, comma 3, della direttiva sul commercio elettronico, Dir. 2000/31/CE) fa espressamente salvo il potere dell'autorità giudiziaria di ordinare la cessazione o la prevenzione dell'illecito, e ciò segnatamente in relazione alle azioni di rimozione e inibitoria di diritto civile (par. 28, con richiamo a OLG Frankfurt, 16 W 10/25). Il *safe harbor*, in altri termini, non solo non copre chi «fa proprio» il contenuto, ma comunque non preclude l'inibitoria. Il giudice ha del pari escluso l'applicabilità del regolamento sull'intelligenza artificiale quale fonte di un autonomo rimedio civile, giacché esso, nei limiti della sua vigenza *ratione temporis*, contempla soltanto il reclamo all'autorità di vigilanza del mercato, e ciò «senza pregiudizio» degli altri

rimedi previsti dal diritto dell'Unione e nazionale.

IV. L'imputazione: l'«allucinazione» come rischio governabile e la diligenza sul processo

Sul piano dell'imputazione, il Tribunale nega che l'errore generativo, l'*AI hallucination*, integri un evento estraneo alla sfera di controllo del gestore. Il contenuto della sintesi è, a differenza dei risultati di terzi, senz'altro sottoponibile a verifica da parte del gestore, quantomeno mediante un raffronto fra le pagine di terzi poste a base della risposta e gli enunciati che su di esse si fondano (par. 39): chi immette sul mercato lo strumento, e ne domina gli algoritmi, non può dunque addurne gli esiti a incidenti. È significativo, in questa prospettiva, che il Tribunale abbia disatteso l'argomento difensivo secondo cui spetterebbe all'utente verificare l'esattezza delle informazioni generate: la mera possibilità, dunque, che un enunciato sia falsificabile mediante ulteriori ricerche non esonera, di regola, dalla responsabilità per quell'enunciato (par. 43).

Il baricentro della responsabilità si sposta, così, dalla *conoscenza* del singolo contenuto illecito, criterio proprio del modello dell'intermediario fondato sulla *notice*, alla *gestione del processo* che quel contenuto genera. Al *provider* non si chiede di prevenire ogni singolo errore, bensì di predisporre procedure idonee («etwa durch entsprechende Kontrollmechanismen») a controllare gli enunciati generati, raffrontandoli con le fonti sottostanti (parr. 39, 45). È la logica, propria dell'ordinamento tedesco, dei *Prüfpflichten* (doveri di controllo), qui declinata quale diligenza sul procedimento generativo più che sul singolo esito. Significativamente, il Tribunale non circoscrive tale dovere alle sole violazioni *manifeste*: esso ritiene tale limitazione inappropriata proprio perché il gestore è in grado di condurre il raffronto con le fonti, e osserva che una sua restrizione lascerebbe altrimenti scoperto il diritto della personalità di chi è leso da enunciati non palesamente, ma solo all'esito di un più attento esame, illeciti (parr. 45-46).

V. Il danno da «zero-click» e lo sviamento di traffico

Di particolare rilievo, infine, la lettura dell'effetto del servizio. Il Tribunale respinge l'argomento secondo cui i collegamenti alle fonti renderebbero l'utente edotto del carattere «altrui» della sintesi, così escludendo l'appropriazione del contenuto: poiché la sintesi è comprensibile di per sé, reca un'affermazione conclusa e priva di indicazioni di inaffidabilità, l'utente non ha di regola ragione di verificarla ulteriormente (par. 43). Il giudice mutua qui, espressamente, la figura del lettore «frettoloso» della sola prima pagina ("Titelseitenleser") e dei *teaser* autosufficienti, già elaborata per la stampa e per le pagine web (par. 43): come per questi, la responsabilità sussiste quando l'enunciato è di per sé compiuto, a prescindere dalla possibilità di consultare il seguito. E il Tribunale aggiunge un rilievo di ordine economico-funzionale: se la AI Overviews dovesse considerarsi, per comune convincimento, inaffidabile e sempre bisognosa di controllo delle fonti, ne risulterebbe frustrata la stessa utilità della funzione vantata dal gestore (par. 43).

«Ciò non è contraddetto dal fatto che, come sostiene il convenuto, gli utenti ragionevoli riconoscano, attraverso i link a siti web di terzi, che il riepilogo nella "panoramica con IA" si basa su contenuti ivi memorizzati da terzi e che, pertanto, il motore di ricerca non si assume alcuna responsabilità per il contenuto di tali informazioni e non le fa proprie. In primo luogo, come già spiegato, il riepilogo in questione non si basa affatto esclusivamente su contenuti di terzi, ma rappresenta piuttosto una compilazione e una valutazione indipendenti che vanno oltre tali siti web di terzi – non ultimo per quanto riguarda il link tra i ricorrenti e un "...", che certamente non deriva da siti web di terzi. In secondo luogo, è possibile che gli utenti possano utilizzare i link per verificare se il contenuto dei siti di terzi corrisponda effettivamente alla "panoramica generata dall'IA", e sono anche generalmente consapevoli che ciò è necessario perché le informazioni generate dall'IA non dovrebbero essere considerate ciecamente attendibili, come sostenuto dal convenuto

nel corso del procedimento orale – tuttavia, la possibilità che un'affermazione possa essere falsificata attraverso ulteriori ricerche non esonera generalmente l'autore dalla responsabilità per tale affermazione. Ciò è riconosciuto per coloro che leggono velocemente, i cosiddetti "lettori della prima pagina" (Corte costituzionale federale, sentenza del 14 gennaio 1998 – causa n. 1 BvR 1861/93 – par. 133; Tribunale regionale superiore di Monaco, sentenza del 31 luglio 2014 – causa n. 18 U 308/14 – par. 25), se e nella misura in cui la dichiarazione è autonoma e comprensibile anche senza ulteriore lettura (ad esempio, all'interno della rivista), e lo stesso vale per i *teaser* autonomi su un sito web o un'app (Tribunale regionale superiore di Colonia, sentenza del 21 maggio 2024 – causa n. 15 W 34/24 – par. 40; Tribunale regionale di Monaco I, sentenza del 24 novembre 2014 – causa n. 9 O 19238/14 – par. 28), e le stesse considerazioni devono valere anche per il "Panoramica con IA": anche questa è autonoma e comprensibile. Contiene un'affermazione completa con un contenuto comprensibile in modo indipendente e non indica altre possibili interpretazioni o persino inaffidabilità legate al contenuto, cosicché gli utenti non hanno generalmente motivo di verificare ulteriormente la risposta visualizzata alla loro query di ricerca. Inoltre, ciò ridurrebbe significativamente il vantaggio della funzione, come affermato dal convenuto, se la "panoramica con IA" fosse generalmente riconosciuta come inaffidabile e ciascuno dei link visualizzati dovesse essere verificato individualmente».

In tale approdo si coglie il riconoscimento, sia pure implicito, del danno da *zero-click* e dello sviamento di traffico a carico degli editori. La sintesi di AI Overviews non si limita a orientare l'utente verso la fonte, ma la sostituisce, intercettando a monte la domanda informativa. La preoccupazione è la medesima già emersa, sul versante autoriale, nel contenzioso statunitense: si tratta della *disintermediazione* dell'accesso all'informazione lamentata dagli editori giornalistici, il rischio che il sistema generativo si sostituisca al sito della testata quale fonte di consultazione, qui recepit, sia pure in un diverso contesto rimediale, quale profilo giuridicamente apprezzabile.

VI. Profili sistematici e trasponibilità al giudizio autoriale

La portata della pronuncia va misurata con la dovuta cautela. Si tratta di un provvedimento reso all'esito di una cognizione sommaria, in materia di diritti della personalità e radicato in categorie proprie dell'ordinamento tedesco, la *Störerhaftung*, i *Prüfpflichten*, la distinzione fra *Tatsachenbehauptungen* e *Meinungsäußerungen*, non meccanicamente esportabili né in altri ordinamenti né, tantomeno, nel giudizio sul diritto d'autore. Va anzi osservato, sul punto, che l'inibitoria non ha riguardato le sole asserzioni di fatto: il Tribunale ha qualificato l'enunciato-chiave («noto per pratiche poco serie... percepito come truffa») come *giudizio di valore*, e ciò nondimeno l'ha vietato, perché fondato su fatti-presupposto falsi («unwahre Anknüpfungstatsachen»), segnatamente il nesso inesistente con le imprese terze, smentito da dichiarazione giurata (parr. 57-64). La tutela ha dunque colpito tanto le *Tatsachenbehauptungen* false, quanto le opinioni prive di base fattuale vera; sono rimaste fuori, per converso, le sole inferenze che il lettore tragga «tra le righe», coperte dall'art. 5 GG (par. 91).

La domanda è stata del resto accolta solo in parte: il Tribunale ha rigettato l'inibitoria quanto all'addebito di collaborazione con una società di recupero crediti, non di per sé screditante e la cui falsità spettava alle attrici provare (parr. 87-90), e quanto all'asserzione - non contenuta nelle sintesi, ma mera inferenza del lettore - di una vendita «poco seria» di abbonamenti (par. 91). Va inoltre segnalato che la difesa del gestore, imperniata sull'avvenuto adempimento dei *Prüfpflichten* - il testo controverso, «rielaborato tramite *machine learning*», non sarebbe più richiamabile, donde il difetto di pericolo di reiterazione e di urgenza - è stata disattesa: in difetto di una *strafbewehrte Unterlassungserklärung* [trad. *dichiarazione di astensione assistita da penale*] e stante l'operare degli algoritmi, non può escludersi il ripresentarsi degli enunciati (parr. 17, 83).

Ciò posto, è l'impianto strutturale della decisione a rivelarsi trasponibile. Il paradigma

dell'intermediario neutrale, costruito, nella stagione del *web 2.0*, attorno alla figura del prestatore che si limita a ospitare, trasmettere o rendere reperibili contenuti altrui, si rivela inidoneo a rappresentare il ruolo delle piattaforme che integrano sistemi di IA per elaborare e diffondere autonomamente contenuti destinati al pubblico. Ove tali contenuti riproducano opere protette, la medesima logica del «contenuto fatto proprio» e della diligenza sul processo generativo conduce a imputare al gestore una responsabilità *diretta*, non schermabile dietro la qualifica di soggetto neutrale né dietro il *safe harbor* ad essa correlato.

Sotto questo profilo la pronuncia converge, per una via affatto autonoma, con l'approdo cui era già pervenuto - sul terreno propriamente autoriale - il *Landgericht München I* nel caso *GEMA v. OpenAI*: là dove quest'ultimo aveva individuato nel fornitore del modello l'autore diretto tanto della riproduzione quanto della comunicazione al pubblico, relegando l'utente, autore di *prompt* elementari, a un ruolo inidoneo a fondarne la responsabilità, qui il medesimo Tribunale nega che il gestore della sintesi generativa possa trincerarsi dietro la posizione del mero intermediario. La convergenza è strutturale, non dogmatica: là il fondamento era il diritto esclusivo di riproduzione e di comunicazione al pubblico, qui la *Störerhaftung* [trad. responsabilità del «perturbatore»] da illecito reputazionale. In entrambe le direttrici affiora, nondimeno, un unico *Leitmotiv*, che può assurgere a chiave di lettura dell'intera prima stagione del contenzioso europeo sull'IA generativa: il tramonto della figura dell'intermediario neutrale, e la correlativa emersione della piattaforma quale soggetto che risponde, a titolo diretto, di ciò che i suoi sistemi autonomamente producono e diffondono.

CONCLUSIONI

La rassegna che precede, pur nella eterogeneità degli ordinamenti interpellati – quello statunitense, quello inglese e quello tedesco –, consente di cogliere, di là dalle differenze di strutturali di ciascuno, alcune tendenze comuni, idonee a restituire un quadro d'insieme del modo in cui il diritto d'autore sta reagendo all'avvento dei modelli di intelligenza artificiale generativa.

La prima, e più evidente, di tali tendenze conferma il progressivo spostamento dell'asse del contenzioso (dunque dell'attenzione): le prime controversie avevano ad oggetto la liceità dell'impiego di opere protette nella fase di addestramento “a monte” (oltre che le modalità di approvvigionamento dei materiali destinati al *training*), mentre l'attenzione appare ora essersi mossa, “a valle”, verso la responsabilità derivante da *output* illeciti.

Da ciò, si può agevolmente dedurre la progressiva scomposizione – dal punto di vista dell'analisi giuridica – del processo di generazione artificiale di contenuti in momenti analiticamente autonomi (fase di creazione del *dataset*, fase del *training* propriamente detto e fase di *output*) ciascuno assoggettato a un proprio, distinto, scrutinio di liceità.

Ne consegue che la liceità di un segmento non si comunica automaticamente agli altri: l'utilizzazione trasformativa nell'addestramento non sana l'eventuale illiceità della condotta di acquisizione iniziale (*Bartz*), né elide la possibile illiceità dell'output qualora riproduttivo o comunque lesivo di diritti di proprietà intellettuale altrui (*Concord*, *GEMA*).

Su un piano più propriamente teorico, le decisioni convergono nell'individuare il nodo nella qualificazione giuridica del modello addestrato e dei suoi parametri: se cioè essi possano considerarsi «copia» o «riproduzione» delle opere utilizzate per l'addestramento.

Ed è qui che si registra la divaricazione più netta dell'intera materia: mentre la *High Court* inglese, in *Getty Images v. Stability AI*, ha escluso che i *model weights* costituiscano una *infringing copy*, sul presupposto che il modello non

«contiene» le opere bensì rappresentazioni matematiche astratte, il *Landgericht München I*, in *GEMA*, è pervenuto alla conclusione opposta, qualificando la «*memorization*» dei testi nei parametri come riproduzione rilevante ai sensi del §16 UrhG, giacché a rilevare non sarebbe la natura del supporto, bensì l'idoneità di esso a rendere nuovamente disponibile l'opera.

La divergenza, come si è visto, è solo in parte formale – attesa la non perfetta sovrapponibilità fra la *infringing copy* del diritto inglese e la *Vervielfältigung* di matrice eurounitaria –, ma sottende una difformità sostanziale di approccio, verosimilmente ricomponibile soltanto a livello sovranazionale.

Tale divaricazione riflette, in fondo, la difficoltà di misurarsi con un oggetto che non è né un archivio statico di opere memorizzate – sicché è fuorviante assimilarlo *tout court* alle piattaforme *peer-to-peer* del primo contenzioso musicale –, né un soggetto creativo autonomo nel senso tradizionale del termine.

Decisiva, in questa prospettiva, appare la distinzione fra una riproduzione *latente*, incorporata in forma compressa nel modello, e una riproduzione *attuale*, percepibile nell'output: distinzione attorno alla quale ruota, in ultima analisi, l'intera questione, e rispetto alla quale la nozione tradizionale di copia, plasmata su un'idea di duplicazione tangibile e concreta, rivela tutta la propria insufficienza.

Una tendenza (o, per meglio dire, divergenza) concerne lo “strumento” giuridico invocato per rendere lecito il *training*.

Da un lato, il modello statunitense si affida alla clausola generale e casistica del *fair use* (§ 107), valutata *ex post* sui relativi quattro fattori e segnata dalla crescente centralità del quarto (*i.e.*, l'effetto dell'opera sul mercato). Dall'altro, il modello unionale prevede un catalogo tipico e tassativo di eccezioni, fra le quali, a venire in rilievo, è quella di *text and data mining* introdotta dalla direttiva (UE) 2019/790, integrata dagli obblighi di trasparenza e di *compliance* introdotti dall'AI Act.

Per quanto distanti nello strumentario, i due modelli rivelano convergenze sostanziali.

La regola affermata in *Bartz* – per cui l’illiceità dell’acquisizione non è sanata dalla trasformatività dell’uso – trova un preciso corrispondente nel requisito del legittimo accesso e del rispetto dell’eventuale esercizio del diritto di *opt out* che condiziona l’operatività dell’eccezione interna di *text and data mining*.

Resta, peraltro, aperta nel sistema europeo la questione – di considerevole rilievo applicativo – della *machine-readability* della riserva di *opt-out*: i giudici di Amburgo, in *Kneschke*, ne hanno offerto una lettura ampia, comprensiva delle riserve formulate in linguaggio naturale, mentre in dottrina si è paventato il rischio di vanificare la finalità protettiva dell’art. 4, par. 3, della direttiva. In tale contesto, la pronuncia *GEMA* valorizza il ruolo delle società di *collecting*, le quali possono – con una dichiarazione generale opponibile *erga omnes* – riservare efficacemente i diritti del proprio repertorio rispetto al *text and data mining* generale.

Infine, sul versante dell’*output infringement*, occorre rilevare che l’eventuale qualificazione del *training* in termini di utilizzazione lecita non vale a giustificare l’eventuale illiceità dell’output, ferma la distinzione strutturale (e conseguentemente giuridica) tra le due fasi. In aggiunta, l’implementazione di *guardrails* tecnici, pur necessaria e rilevante ai fini della valutazione della diligenza del *provider* (*Concord*), non costituisce di per sé esonero da responsabilità – segnatamente, in *GEMA*, per la memorizzazione delle opere nei parametri del modello.

In definitiva, i contenziosi commentati in questa sede non offrono – né, attesa la loro natura in larga parte interlocutoria e cautelare, potrebbero offrire – una soluzione alla questione del rapporto fra intelligenza artificiale generativa e diritto d’autore.

Piuttosto, ove letti congiuntamente, essi offrono una chiave di lettura che consente di scomporre un fenomeno unitario soltanto in apparenza nelle sue componenti elementari – acquisizione, addestramento, generazione –, sottoponendo ciascuna a un autonomo giudizio di liceità, e di saggiare la portata delle categorie tradizionali –

la copia, la riproduzione, l’utilizzazione – alla prova di un’innovazione che ne erode i contorni.

