

L'USO DI OPERE D'ARTE NEI SISTEMI DI AI GENERATIVA IN FASE DI INPUT E OUTPUT

LAURA CASTELLI
*Professore associato
nell'Università di Milano*

SOMMARIO: 1. L'utilizzo delle opere d'arte nei sistemi di AI generativa. – 2. Il quadro normativo europeo e italiano. – 3. L'uso di immagini in fase di *input* ai soli fini di addestramento della macchina. – 4. Possibili lesioni del diritto d'autore in assenza di *opt-out*. – 5. I criteri di liceità dell'*appropriation art* creata dall'intelligenza artificiale.

1. – Negli ultimi anni lo sviluppo delle intelligenze artificiali generative ha sollevato nuove questioni giuridiche, connesse tra l'altro all'uso di immagini protette dal diritto d'autore nei processi di addestramento delle macchine AI¹. Le AI generative, infatti, come quelle basate su reti neurali profonde, necessitano di enormi quantità di dati per apprendere schemi visivi e produrre nuove immagini². L'inserimento di opere d'arte, fotografie e altri contenuti visivi protetti dal *copyright* pongono interrogativi circa la compatibilità di tali pratiche con le normative vigenti. A livello internazionale, il dibattito si concentra sul difficile bilanciamento tra innovazione tecnologica e tutela dei diritti degli autori; la comunità scientifica si interroga, in particolare, sia sulla possibilità che la macchina possa utilizzare immagini protette dal diritto d'autore in virtù delle eccezioni, di cui

¹ Ai sensi dell'art. 3.1 AI ACT, Reg. 2021/0106 (COD), approvato dal Parlamento il 31 marzo 2024, si definisce sistema di intelligenza artificiale «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».

² Tra i diversi tipi di intelligenza artificiale, quella generativa è la più evoluta, ed è in grado di generare contenuti nuovi, creare testi o immagini grazie alle indicazioni (*prompt*) fornite dall'utente. Per funzionare ha bisogno di essere istruita e addestrata, mediante l'inserimento di un enorme quantitativo di dati che, combinati ad algoritmi, le permettono di fornire risposte e generare testo. Sul funzionamento dell'AI e i modelli di addestramento cfr. Y. LE CUN, Y., G. HINTON, Y. BENGIO, *Deep learning*, in *Nature*, 2015, 436 ss.; N. CRISTIANINI, *La scorciatoia*, Bologna, 2023; sul funzionamento dei sistemi che generano immagini cfr. L. TURINI, *A chi spettano i diritti sulle opere dell'intelligenza artificiale. Dall'addestramento su opere altrui alla generazione di opere nuove*, Santarcangelo di Romagna, 2024, 15 ss.

si dirà a breve, previste dalla legge per il *text and data mining* (TDM)³ - intesa come modalità di reperimento di dati e di estrazione degli stessi, che sovente trae appunto informazioni da opere dell'ingegno - in funzione dell'operatività della macchina, sia sulle conseguenze economiche e culturali della generazione automatica di contenuti derivati.

Nel frattempo gli artisti "tradizionali" denunciano i rischi di violazione dei diritti d'autore insiti nell'utilizzo di immagini di opere d'arte per addestrare le macchine AI: è di questi giorni la notizia di una petizione, che ha raccolto le firme di più di seimila artisti, con cui si è chiesto a Christie's di sospendere l'asta *Augmented Intelligence*, la prima dedicata solamente a opere generate tramite AI e *Machine Learning* (dipinti, sculture, stampe, opere digitali e interattive, firmate da pionieri dell'intelligenza artificiale degli anni '60 e da artisti contemporanei). La Casa d'aste non ha dato seguito a tale richiesta, manifestando al contrario entusiasmo per l'iniziativa e, più in generale, per i nuovi sistemi di AI generativa – da intendere non tanto come un sostituto della creatività umana, quanto come uno strumento per migliorarla - e la vendita si è conclusa con un ricavo totale di circa 730.000 dollari.

In attesa di conoscere quali linee guida in materia di diritto d'autore verranno emanate dallo U.S. Copyright Office oltreoceano⁴, è lecito

³ L'intelligenza artificiale è un *software* e, per funzionare, utilizza gli algoritmi, ovvero una serie di istruzioni che, dopo essere state eseguite, permettono alla macchina di ottenere un certo risultato. L'AI si basa su sistemi di *machine learning*, ovvero su algoritmi che individuano le connessioni tra dati e danno esecuzione ai compiti, attraverso l'interferenza, ossia processi di deduzione logica dei dati che hanno appreso. Il grado di performatività è strettamente connesso alla quantità e qualità di dati impiegati in sede di addestramento. Le reti neurali hanno più strati: uno di *input*, uno di *output* e uno intermedio, che elabora i dati. Per funzionare, una macchina deve essere addestrata, mediante l'assimilazione del maggior numero possibile di dati. Tale addestramento è suddiviso in tre fasi: la prima è quella di "codificazione del testo"; la seconda è definita "codificazione dell'immagine"; la terza è quella in cui viene realizzata una tabella che unisce i testi e le immagini codificate, mediante una griglia che ha gradi di probabilità di corrispondenza tra definizioni e figure. Come rileva L. TURINI, *op. cit.*, 12, simile lavoro non è fatto sulla totalità dell'immagine, ma su ogni elemento di cui essa è composto, così che all'interno di un sistema vi sono punti che corrispondono a piccoli pezzi (*token*) che, presi in considerazione singolarmente, non sono percepibili.

⁴ Il 31 luglio 2024 l'U.S. Copyright Office ha pubblicato la prima parte del rapporto sulle questioni legali e politiche relative al diritto d'autore e all'intelligenza artificiale, contenente una Raccomandazione per l'adozione di una legislazione federale per contrastare la distribuzione non autorizzata di *deepfakes*. Il 28 gennaio 2025 è stata pubblicata la seconda parte, riguardante la tutela, mediante le norme in materia di diritto d'autore, delle opere create mediante AI. Si attende ora la terza e ultima parte,

domandarsi se e quali regole del nostro Paese siano in grado di fornire una risposta al quesito posto da simili tecniche: è legittimo inserire immagini di opere d'arte protette dal diritto d'autore per allenare e addestrare sistemi di AI generativa?

2. – Nel nostro Paese, l'uso di immagini tutelate dal diritto d'autore per l'addestramento di modelli di intelligenza artificiale generativa si inserisce in un contesto normativo complesso, che vede l'intersezione tra il diritto d'autore tradizionale e le più recenti regolamentazioni sull'uso dei dati.

Con particolare riferimento a queste ultime, in Europa la principale normativa di riferimento è la direttiva (UE) 2019/790 sul diritto d'autore nel mercato unico digitale - recepita nel nostro ordinamento con il d. lgs. 177/2021 - che ha introdotto specifiche regole per il c.d. *text and data mining* (TDM) definito, all'art. 2, come «qualsiasi tecnica di analisi automatizzata volta ad analizzare testi e dati in formato digitale avente lo scopo di generare informazioni, inclusi, a titolo non esaustivo, modelli, tendenze e correlazioni».

In dottrina vi è chi ha dubitato che il legislatore europeo, all'epoca dell'emanazione della direttiva, fosse realmente consapevole dello sviluppo che avrebbero avuto i sistemi di AI generativi e intendesse riferirsi anche alle modalità di costruzione di un modello di apprendimento automatizzato⁵. A tal riguardo si è osservato come nel *memorandum* di accompagnamento della bozza polacca di recepimento della direttiva *Digital Copyright* sia stata messa in luce l'estraneità della disciplina europea rispetto alle attività di *machine learning*, considerato che quando è stata emanata la direttiva non si potevano ancora prevedere le potenzialità dei sistemi di AI generativi⁶.

che affronterà le questioni giuridiche derivanti proprio dall'addestramento dei modelli AI mediante opere protette dal diritto d'autore.

⁵ A. PEUKERT, *Copyright in the Artificial Intelligence Act – A primer*, in *GRUR International*, vol. 73, VI, 2024, 503 s.; G. GEIER, V. IAIA, *Towards and independent EU Regulator for copyright issues of generative AI: what role for the AI office (but more importantly: what's next?)*, in *Auteurs et medias*, II, 2024, 188 s.; T. DORNIS, W. TIM, S. STÖBER, *Urheberrecht und training generativer KI-Modelle – technologische und juristische Grundlagen*, 2024, in www.papers.ssrn.com; contra P. KELLER, *TDM: Poland challenges the rule of EU law*, in *Kluwer Copyright Blog*, 20 febbraio 2024.

⁶ V. IAIA, *La proprietà intellettuale nell'era dell'intelligenza artificiale*, Milano, 2025, 84. Secondo l'A. inoltre, l'area di libero esercizio di estrazione di testo e di dati disegnata dagli artt. 3 e 4 della direttiva ha fatto supporre che «l'intercettazione delle attività di *machine learning* sia perlopiù il frutto di una (s)fortunata coincidenza, anziché di encomiabile lungimiranza». Si veda anche V. IAIA, *Intelligenza artificiale generativa e diritto d'autore*, in *manuale sull'intelligenza artificiale*, a cura di R. RAZZANTE, Torino, 2024, 209.

Indipendentemente dall'esistenza o meno di questa consapevolezza, ciò che emerge dalla lettera della legge è che dopo aver definito l'attività di *mining* – intesa come analisi automatizzata di dati – sono previste, agli artt. 3 e 4, due distinte eccezioni al diritto esclusivo di riproduzione e di estrazione⁷. L'articolo 3, in particolare, prevede un'eccezione al potere escludente del diritto d'autore in favore degli istituti di ricerca e agli enti di patrimonio culturale, rispetto alle attività di estrazione di testo e di dati utili per scopi di ricerca scientifica. *Ratio* della norma è quella di favorire la ricerca basata su *big data* e intelligenza artificiale, permettendo agli enti di ricerca di analizzare notevoli quantità di testi e dati senza ostacoli giuridici.

Laddove ad accedere a queste informazioni siano privati e aziende, tale necessità è regolata dall'articolo successivo, che prevede un'eccezione più ampia, ma soggetta al diritto di *opt-out* da parte dei titolari dei diritti d'autore. L'art. 4 si riferisce in particolare alle estrazioni e riproduzioni di opere dell'ingegno per finalità commerciali e introduce un'eccezione più estesa, consentendo il *text and data mining* per qualsiasi finalità, ma prevedendo al contempo ai titolari dei diritti d'autore di esercitare il c.d. *opt-out*, vietando l'estrazione di testi e dati di proprie opere.

La disposizione in esame pone una serie di problemi sia applicativi, sia di interpretazione.

In primo luogo, allo stato le modalità operative di esercizio dell' *opt-out* appaiono tutt'altro che semplici: il titolare di opere protette dal diritto d'autore potrà senz'altro manifestare il proprio rifiuto attraverso misure tecnologiche di protezione, prevedendo ad esempio restrizioni all'accesso di dati; ove però si dovesse trattare di opere pubblicamente disponibili *online*, la legge impone l'utilizzo di «strumenti che consentano lettura automatizzata», senza specificare come debba essere espresso tale rifiuto a seconda dei dati che possono essere oggetto di *TDM*. Occorre considerare che attualmente non esistono *standard* univoci adottati da tutti i sistemi e ciò rende complesso l'esercizio del diritto di *opt-out*, essendo elevato il rischio che esso non sia leggibile dalla macchina, tutte le volte in cui sia espresso in un formato che i sistemi informativi non sono in grado di riconoscere automaticamente senza necessità di interpretazione manuale.

⁷ L'estrazione di dati da parte di una macchina di *AI generativa* si riferisce al processo mediante il quale un sistema di *AI* è in grado, sulla base delle indicazioni e delle informazioni fornite dall'utente (c.d. *prompt*), di generare qualcosa di nuovo, creare testi o immagini e fornire risposte. Funziona sulla base di calcoli probabilistici e necessita di essere addestrata su un enorme quantitativo di dati che, abbinati a algoritmi, le permettono di mettere in relazione tali dati e generare testo. Si veda in materia N. CRISTIANINI, *op. cit.*

La norma in esame pone anche dubbi interpretativi: occorre infatti individuare il campo di applicazione delle eccezioni appena esaminate, volte a contemperare l'interesse all'innovazione con la tutela dei diritti d'autore.

Da un esame d'insieme della disciplina, l'intenzione del legislatore sembra essere quella di regolare non tanto l'attività di analisi automatizzata dei dati, ma quella di estrazione degli stessi. La distinzione tra analisi dei dati ed estrazione, che emerge nel testo originale della direttiva, sfuma nel testo italiano della stessa, ove i termini *mining* ed *extraction* vengono tradotti nel medesimo modo, utilizzando l'espressione "estrazione". Questa confusione potrebbe allora avallare l'ipotesi che, a parte nei casi eccezionali di cui agli artt. 3 e 4, l'attività di analisi dei dati sia sottoposta a restrizioni, laddove, a un attento esame della normativa vigente, allo stato non sembrano esserci norme che sottopongono ad autorizzazione l'analisi dei dati, che la macchina conduce in modo automatico e per di più in un linguaggio incomprensibile per l'uomo⁸.

Nel recepire la direttiva 2019/790, il nostro Paese ha disciplinato solo la fase di estrazione e riproduzione di opere, senza occuparsi della fase precedente, ossia quella di analisi dei dati⁹, così che, al momento, l'utilizzo delle immagini per il solo addestramento della macchina sembrerebbe consentito¹⁰.

Sul piano pratico, peraltro, in questa fase i problemi dipendono dal fatto che, ad oggi, risulta quasi impossibile conoscere il vero funzionamento delle macchine e verificare se esse, in fase di analisi dei dati, si limitino a studiarli e ad acquisire nozioni e idee, di per sé liberamente appropriabili, o se copino il materiale esaminato. A tal riguardo occorre peraltro considerare che la Commissione europea, in uno studio recente in materia di *copyright* e nuove

⁸ L. TURINI, *op cit.*, 115, la quale, nell'ipotizzare che l'attività di *text and data mining* sia libera e non soggetta ad autorizzazione, trova conferma nell'art.4, comma 1, della direttiva, in cui si legge che «gli Stati membri dispongono un'eccezione o una limitazione ai diritti di cui all'articolo 5, lettera a), e all'articolo 7, paragrafo 1, della direttiva 96/9/CE, all'articolo 2 della direttiva 2001/29/CE, all'articolo 4, paragrafo 1, lettere a) e b), della direttiva 2009/24/CE e all'articolo 15, paragrafo 1, della presente direttiva per le riproduzioni e le estrazioni effettuate da opere o altri materiali cui si abbia legalmente accesso ai fini dell'estrazione di testo e di dati», in cui dunque l'eccezione è riferita solo ai «diritti protetti dalle norme richiamate che non includono il diritto di *text and data mining*».

⁹ All'art. 70-*quater* l. aut. si legge: «Fermo 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».

¹⁰ Si veda sul punto L. TURINI, *op. cit.*, 116.

tecnologie, nell'esprimersi sull'ambito di applicazione dell' *opt-out* – che ha il fine di riservare i diritti autoriali ove sia possibile rinvenire una riproduzione - ha osservato come non necessariamente tutte le copie tecniche realizzate nel processo *TDM* siano da considerarsi riproduzioni¹¹.

3. – Si è detto di come dall'analisi delle regole dettate dalla direttiva 2019/790 sul diritto d'autore nel mercato unico digitale non emergano regole specifiche riferite al semplice inserimento di immagini da utilizzare per l'allenamento della macchina¹².

Il Parlamento europeo, dal canto suo, nonostante abbia integrato il testo dell' *AI Act* poco prima della sua emanazione¹³, proprio al fine di far fronte ai problemi che l'utilizzo di opere autoriali senza il consenso dei titolari potrebbe generare, ha dedicato alla fase di *input* soltanto la previsione di cui all'art. 53, par. 1, lett. d), ove si sancisce l'obbligo per i fornitori di *AI* per

¹¹ *Study on copyright and new technologies: copyright data management and Artificial Intelligence*, European Commission, February 2022, <https://op.europa.eu/en/publication-detail/-/publication/cc293085-a4da-11ec-83e1-01aa75ed71a1/language-en>. Di seguito il testo in inglese del passaggio citato nel testo: «Even if the exception can be aside by an opt out this does not entail that all technical copies made in tdm process are to considered as reproductions: the output is meant to reserve the right, where the right is applicable, i.e. where a reproduction can be found (in particular determining whether the work is recognizable and thus exploited in the course of the TDM process or afterwards)».

¹² Osserva G. GHIDINI, *Intelligenza artificiale: due questioni di proprietà intellettuale*, in *Riv. dir. ind.*, 2024, 145: «Ora, impedire l'accesso e l'uso, anche ultra-temporaneo, di dati per sistemi IA solo allo scopo di istruirli a produrre nuove opere, equivarrebbe, in termini umani, al divieto, per un autore, di 'archiviare' nella sua mente, e utilizzare, frammenti di opere di autori precedenti. Tale divieto sarebbe inconciliabile con la protezione del copyright non solo perché impraticabile, ma perché, precisamente, non vi può essere alcuna violazione autoriale nello studiare e memorizzare opere altrui allo scopo di produrre nuove opere — distinte, ripeto ancora, quanto a forma espressiva».

¹³ Reg. (U.E.) 2024/1689 del Parlamento Europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (C.E.) n. 300/2008, (U.E.) n. 167/2013, (U.E.) n. 168/2013, (U.E.) 2018/858, (U.E.) 2018/1139 e (U.E.) 2019/2144 e le dir. 2014/90/UE, (U.E.) 2016/797 e (U.E.) 2020/1828 (reg. sull'intelligenza artificiale). Scopo di tale atto normativo è quello di «migliorare il funzionamento del mercato interno, istituendo un quadro giuridico uniforme in particolare per quanto riguarda lo sviluppo, l'immissione sul mercato, la messa in servizio e l'uso di sistemi di intelligenza artificiale (sistemi di IA) nell'Unione". Si noti che esso si applicherà a decorrere dal 2 agosto 2026, ad eccezione dei capi I e II che si applicano a decorrere dal 2 febbraio 2025; dei capi III, sezione 4, V, VII, XII e dell'articolo 78 che si applicano a decorrere dal 2 agosto 2025, ad eccezione dell'articolo 101; dell'articolo 6, paragrafo 1 che si applica a decorrere dal 2 agosto 2027. Per un commento del Regolamento cfr. V. CALAPRICE, *Il Regolamento europeo (c.d. AI Act)*, in *Manuale sull'intelligenza artificiale*, a cura di R. RAZZANTE, Torino, 2024, 17 ss.

finalità generali¹⁴ di redigere e mettere a disposizione del pubblico «una sintesi sufficientemente dettagliata dei contenuti utilizzati per l'addestramento del modello di AI». Tale norma, invero, non impone particolari obblighi in fase di analisi, se non quello di rendere trasparente l'utilizzo del materiale impiegato per l'addestramento della macchina, ciò che del resto sembrerebbe necessario al fine di consentire agli autori delle opere utilizzate di venire a conoscenza delle modalità di sfruttamento delle stesse, per scegliere se e come far valere il diritto di *opt-out*, altrimenti verosimilmente impossibile¹⁵.

Peraltro, anche laddove il legislatore scegliesse in futuro di regolare e limitare espressamente l'utilizzo di immagini in fase di analisi, tale limitazione si troverebbe necessariamente a fare i conti con alcune criticità.

In primo luogo occorre considerare che, in sede di regolamentazione della fase di analisi dei dati, tra gli obblighi che il legislatore potrebbe imporre, molti sarebbero difficili da rispettare, con il rischio, giustamente rilevato dalla dottrina, di *overregulation*¹⁶: basti pensare alla possibilità che i fornitori di

¹⁴ L'art. 3, n. 63, *AI Act* definisce il modello di AI per finalità generali «un modello di IA, anche laddove tale modello di IA sia addestrato con grandi quantità di dati utilizzando l'autosupervisione su larga scala, che sia caratterizzato da una generalità significativa e sia in grado di svolgere con competenza un'ampia gamma di compiti distinti, indipendentemente dalle modalità con cui il modello è immesso sul mercato, e che può essere integrato in una varietà di sistemi o applicazioni a valle, ad eccezione dei modelli di IA utilizzati per attività di ricerca, sviluppo o prototipazione prima di essere immessi sul mercato».

¹⁵ Osserva V. IAIA, *op. ult. cit.*, 90, come «in assenza di alcuna *disclosure* sui contenuti autoriali utilizzati sarebbe estremamente difficile, se non impossibile, per gli autori verificare il rispetto della propria manifestazione di volontà e, in caso di violazione, disporre degli strumenti probatori per legittimare una pretesa risarcitoria». Si veda anche quanto previsto dal Considerando 107: «Al fine di aumentare la trasparenza sui dati utilizzati nelle fasi di pre-addestramento e addestramento dei modelli di IA per finalità generali, compresi testo e dati protetti dalla normativa sul diritto d'autore, è opportuno che i fornitori di tali modelli elaborino e mettano a disposizione del pubblico una sintesi sufficientemente dettagliata dei contenuti utilizzati per l'addestramento del modello di IA per finalità generali. Pur tenendo debitamente conto della necessità di proteggere i segreti commerciali e le informazioni commerciali riservate, la presente sintesi dovrebbe essere di respiro ampio e generale, anziché dettagliata sotto il profilo tecnico, al fine di agevolare le parti con interessi legittimi, compresi i titolari dei diritti d'autore, nell'esercitare e far rispettare i loro diritti ai sensi del diritto dell'Unione, ad esempio elencando le principali raccolte o serie di dati che sono state inserite nell'addestramento del modello, quali grandi banche dati o archivi di dati privati o pubblici, e fornendo una descrizione delle altre fonti di dati utilizzate. È opportuno che l'ufficio per l'IA fornisca un modello per la sintesi, che dovrebbe essere semplice ed efficace nonché consentire al fornitore di fornire la sintesi richiesta in forma descrittiva».

¹⁶ V. IAIA, *op. ult. cit.*, 101.

servizi di *AI*, proprio per eludere le disposizioni che impongono limitazioni a tutela del diritto d'autore, decidano di addestrare i propri algoritmi con banche dati disponibili in Paesi, come il Giappone, che possiede regole di tutto vantaggio per le attività di *machine learning*¹⁷.

Inoltre, non può trascurarsi che difficilmente un autore sarebbe in grado non solo di dimostrare, ma ancor prima di accorgersi dell'utilizzo di una sua opera ai soli fini di addestramento della macchina, laddove quest'ultima producesse poi una creazione per nulla riconducibile alle immagini utilizzate per allenarsi: i sistemi di *AI* non sono soliti divulgare informazioni circa il materiale utilizzato per l'allenamento della macchina, così che diventa difficile individuare le opere impiegate a tale scopo¹⁸.

Di contro, sarebbe evidente l'uso di un'immagine protetta dal diritto d'autore laddove la macchina ne creasse una copia, ai fini di asserite finalità di ricerca scientifica; o anche se elaborasse un'opera identica a una già esistente e tutelata, attribuendosene la paternità; o, ancora, se ne creasse una che fosse il frutto di elaborazione e trasformazione di altra protetta.

Nel primo caso, l'opera copiata dovrebbe considerarsi tutelata dalle norme autoriali, salvo che l'autorità giudiziaria eventualmente adita non dovesse ricondurre le copie realizzate dalla macchina a una delle eccezioni riconosciute dalla direttiva e sopra richiamate, come ad esempio è accaduto nella prima controversia instaurata in materia in Europa, che ha visto coinvolti il fotografo Robert Kneschke e LAION ("*Large-scale Artificial Intelligence Open Network*"): l'artista ha in particolare contestato a quest'ultima il fatto di aver costruito un *dataset* contenente un *link* per ogni immagine oggetto di possibile estrazione e copia, tra cui compariva anche una sua opera, nonostante l'immagine fosse reperibile su un sito internet i cui termini e condizioni prevedevano una riserva esplicita dei diritti d'autore. La Corte distrettuale di Amburgo ha in questo caso ritenuto applicabile l'eccezione

¹⁷ Per un'analisi della disciplina giapponese a sostegno delle attività tecnologiche che implicano l'accesso a dati coperti dai diritti autoriali cfr. P. SAMUELSON, *Text and Data Mining of In-Copyright works: is it legal?* in *Communications of ACM*, vol. 64, I, 2021, 20; T. UENO, *Flexible copyright exception for "Non-enjoyment purposes – Recent amendment in Japan and its implication*, in *GRUR International*, 2022, 3, 149.

¹⁸ Che vi sia una totale mancanza di informazioni sull'insieme dei dati di addestramento emerge anche da una recente vicenda giudiziaria in cui la Corte distrettuale della California ha rigettato l'azione proposta da alcuni artisti hollywoodiani nei confronti di Open AI, a causa dell'estrazione non autorizzata di propri libri, proprio per carenza di prove, non ritenendo sufficienti a tal fine i risultati, prodotti dagli attori, di una ricerca sul sito *haveibeentrained.com*, un servizio che sarebbe in grado di individuare i contenuti utilizzati tramite training algoritmici: Corte distrettuale della California 12 febbraio 2024, *Paul Trembaly et al. V. OpenAI, Silverman et v. OpenAI*, nn.23-cv-03223-AMO e 23-cv-03416-AMO.

connessa alle finalità di ricerca scientifica, che ha interpretato in senso lato, includendo anche il lavoro necessario all'acquisizione di conoscenza futura, come ad esempio la raccolta di enormi quantità di dati utili a giungere a conclusioni empiriche attraverso un metodo induttivo¹⁹.

Laddove invece la macchina producesse un'opera identica a un'altra protetta dal diritto d'autore, attribuendosene la paternità, ci si troverebbe di fronte a un plagio²⁰. Nonostante non ci sia nella legge italiana alcun esplicito riferimento a tale condotta, si è soliti ricondurre questo illecito, denominandolo "plagio-contraffazione", alla previsione di cui all'art. 156 l. aut., che disciplina l'ipotesi in cui venga violato un diritto di utilizzazione economica, cui peraltro normalmente si accompagna la falsa attribuzione della paternità in capo a un soggetto diverso dall'autore²¹.

Come è noto, con riferimento alle opere d'arte, la giurisprudenza è solita negare l'esistenza di un plagio allorché due lavori, pur prendendo spunto dalla stessa idea ispiratrice, si differenzino negli elementi essenziali, che ne caratterizzano la forma espressiva: in altre parole, a essere tutelata dalla legge sul diritto d'autore non è mai l'idea, ma il modo in cui questa viene espressa²².

¹⁹ Corte distrettuale di Amburgo, 27 settembre 2024, n. 310, O 227/23. Osserva V. IAIA, *op. ult. cit.*, 113, come il ragionamento del giudice tedesco si esponga alla critica di eccessivo semplicismo. In questi casi sarebbe stato utile, infatti, operare una distinzione tra «l'attività di analisi computazionale delle immagini ai fini della costruzione del *dataset* e la comunicazione al pubblico dello stesso – unitamente alle fotografie oggetto di addestramento reperibili attraverso i collegamenti ipertestuali in esso contenuti».

²⁰ Per un'analisi delle regole utilizzate dalla giurisprudenza per verificare la sussistenza di un plagio in sede di utilizzo delle opere dell'ingegno si veda C. DI BENEDETTO, *Il plagio-contraffazione di opere dell'ingegno nella giurisprudenza italiana*, in *Il dir. ind.*, 2022, 489 ss.

²¹ L'art. 518 *quaterdecies*, inserito dall'art. 1, comma 1, lett. b), della l. 9 marzo 2022, n. 22, punisce «chiunque, al fine di trarne profitto, contraffà, altera o riproduce un'opera di pittura, scultura o grafica ovvero un oggetto di antichità o di interesse storico o archeologico».

²² Cass. civ., 26 gennaio 2018, n.2039, in *Foro it.* 2018, 3, I, 855; Trib. Milano, 25 luglio 2017, in *AIDA* 2019, 1, 752: il plagio riguardava in questo caso un'opera del pittore Emilio Isgrò, che ha adito l'autorità giudiziaria per inibire alla Sony Music la distribuzione dell'involucro della copertina dell'opera discografica di Roger Waters, dal titolo *Is this the life we really want?*, denunciando la violazione dei diritti d'autore del ricorrente sull'opera *Cancellatura* del 1964. La Sony si è difesa sostenendo che l'opera rappresentata sulla copertina e sul cd aveva un contenuto concettuale completamente diverso dall'opera di Isgrò, volendo denunciare la censura posta in atto dal governo statunitense e, come diretta conseguenza, la cancellazione della cultura. Il Tribunale, sul presupposto che ciò che la disciplina sul diritto d'autore mira a tutelare non è mai l'idea in sé, ma la forma della sua espressione, conferma l'illiceità dell'opera denunciata dall'artista, facendo leva sul raffronto delle due opere, da cui

Ciò è esplicitato dal nostro legislatore con riferimento alle opere create dai programmi per elaboratore, ove si escludono espressamente dall'area della tutela accordata dalla legge autoriale «le idee e i principi che stanno alla base di qualsiasi elemento di un programma, compresi quelli alla base delle sue interfacce»²³. E proprio da simile previsione emerge allora come, ai fini della tutela autoriale, possa rilevare non tanto il semplice utilizzo di immagini per allenare la macchina - che, appunto, hanno lo scopo di aiutare quest'ultima a trovare idee o ispirazioni -, ma le attività che vengono poste in essere nella fase successiva, ovvero quella in cui la *AI*, una volta in possesso del materiale necessario per elaborare nuove idee, le plasma e dà ad esse una forma.

È in questa seconda fase che possono porsi problemi legati alla lesione di diritti autoriali, peraltro non solo nel caso in cui l'autore dell'opera utilizzata abbia negato espressamente la possibilità di utilizzo, ma anche in assenza di diniego, per le ragioni che ora si esporranno.

4. – Non si può negare che, allo stato, gli autori di opere d'arte protette dalla tutela autoriale possano riscontrare difficoltà notevoli in sede di negazione esplicita del consenso all'estrazione delle opere stesse, soprattutto quando un artista sia autore di numerosi lavori. Un dissenso generalizzato e non riferito alla singola opera non sarebbe leggibile dalla macchina, così che tale espressione del dissenso equivarrebbe nei fatti a consenso: la lettera della legge, infatti, è chiara nel prevedere che il titolare dei diritti debba essersi espressamente riservato l'uso delle opere «in modo appropriato»²⁴, pena, nei fatti, l'intellegibilità da parte dell'*AI*.

Emblematica è in tal senso la volontà recentemente manifestata in Francia da SACEM (*Société des auteurs, compositeurs et éditeurs de musique*), la quale con un comunicato ufficiale ha negato il consenso all'estrazione delle opere appartenenti ai suoi 200.000 rappresentanti. Difficilmente tale dichiarazione generalizzata potrà considerarsi sufficiente ai fini di un'azione per violazione del diritto d'autore, dal momento che in questo caso non vi è nessuna manifestazione di dissenso associata alla singola opera, che sia intellegibile dalla macchina²⁵. Sicuramente col passare del tempo si affineranno le tecniche e si elaboreranno strumenti per esprimere facilmente

emerge in modo evidente come il cd abbia ripreso pedissequamente la forma espressiva personale di Emilio Isgro.

²³ Art. 2, comma 8, l. aut. Allo stesso modo, l'art. 9, comma 2, accordi TRIPs, prevede che «La protezione del diritto d'autore copre le espressioni e non le idee, i procedimenti, i metodi di funzionamento o i concetti matematici in quanto tali».

²⁴ Art. 4, comma 3, dir. 2019/790.

²⁵ In tal senso V. IAIA, *La proprietà intellettuale nell'era dell'intelligenza artificiale*, cit., 85.

il diniego in modo circostanziato²⁶; sino a quando questa operazione rimarrà complessa, con il rischio che un eventuale dissenso non sia letto dalla macchina - che potrebbe dunque considerarsi autorizzata a copiare e rielaborare le immagini inserite ai fini del suo addestramento -, l'unico modo per tutelare le opere protette dal diritto d'autore sembra essere quello di considerare questa rielaborazione legittima solo se rispettosa delle norme in materia di diritto d'autore.

In altri termini, l'assenza di un diniego non potrebbe valere anche ai fini del riconoscimento, in capo alla macchina, del diritto di copiare l'opera utilizzata in sede di *text and data mining* e utilizzarla in violazione delle norme autoriali, ad esempio attribuendosene la paternità²⁷.

5. – Allo stesso modo, laddove la macchina rielaborasse l'opera utilizzata, creandone una nuova, per valutare la liceità di quest'ultima i criteri da utilizzare sarebbero quelli usati per i casi in cui ad appropriarsi di un'opera sia un artista in carne ed ossa. Ai fini della liceità dell'opera frutto del lavoro della macchina, occorrerà in particolare verificare che essa rientri nei parametri della cd. *appropriation art*, quella forma d'arte caratterizzata dal fatto di utilizzare deliberatamente creazioni altrui allo scopo di trasmettere un messaggio ulteriore e diverso da quello evocato dall'opera originaria.

Tale arte deve il suo nome all'atto che l'artista compie nel suo processo creativo: egli infatti si appropria, senza autorizzazione, di oggetti, immagini e opere d'arte, per trasformarli e incorporarli nella propria opera²⁸.

²⁶ Per un elenco dei diversi strumenti ad oggi disponibili per l'esercizio dell'*opt-out* si veda P. KELLER, Z. WARSO, *Defining best practice for opting out of ML training*, Open future policy brief n.5, 2023.

²⁷ Si veda sul punto G. GHIDINI, *op. cit.*, 145: «Mettiamo da parte l'ipotesi che i dati in questione siano riprodotti nella loro forma espressiva (distinta dall'idea sottostante, secondo il principio storico ribadito anche dal TRIPs, art. 9.2.) nel 'prodotto finale', l'output, del robot. Ciò costituirebbe indiscutibilmente una contraffazione».

²⁸ W.M. LANDES, *Copyright, Borrowed Images and Appropriation Art: An Economic Approach*, in *U Chicago Law & Economics*, Working paper 113, 2000, 2, fornisce la seguente definizione di *Appropriation Art*: «*Appropriation Art borrows images from popular culture, advertising, the mass media, other artists and elsewhere, and incorporates them into new works of art. Often, the artist's technical skills are less important than his conceptual ability to place image in different settings and, thereby change their meaning. Appropriation art has been commonly described "as getting the hand of art and putting the brain in"*. Tali concetti sono poi stati codificati all'interno del *Copyright Act*, che alla sec. 107 stabilisce che «*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;*

L'ossimoro con il quale questo tipo di arte viene definita - la parola "arte" evoca il concetto di originalità, laddove la parola "appropriativa" evoca un plagio, - mette in luce le difficoltà interpretative che il giurista incontra nel valutarne la liceità, i cui confini, disegnati nel sistema americano attraverso la c.d. *fair use doctrine*²⁹, appaiono meno nitidi nel nostro Paese, ove il *discrimen* tra ciò che è lecito e ciò che è plagio passa attraverso un concetto di parodia dai contorni sfumati.

Né sul punto è stato d'aiuto il legislatore europeo che, nella direttiva 2001/29 sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione, ha previsto la facoltà per gli Stati membri di stabilire eccezioni al diritto di riproduzione delle opere quando «l'utilizzo avvenga a scopo di caricatura, parodia o pastiche», senza peraltro fornire una nozione e definire i confini del concetto di parodia³⁰.

Nel nostro Paese, dalle - invero poche - pronunce che si sono occupate di arte appropriativa, emerge come ad oggi la nostra giurisprudenza non segua una direttrice comune, ispirandosi talvolta al *fair use* americano, talaltra ai principi contenuti nella Costituzione.

Quest'ultima viene in particolare invocata quando alla base della trasformazione sia rinvenibile un intento sarcastico.

Così, in una pronuncia del Tribunale di Venezia, la liceità di un'opera appropriazionista è stata ricondotta alla libertà di manifestazione del pensiero, *ex art. 21 e 33 Cost.*, e all'esimente della parodia e della critica sarcastica³¹. Si trattava in particolare dell'opera di Samson Kambalu, esposta alla biennale di Venezia nel 2015, che riproduceva e assemblava, tra l'altro, anche immagini fotografiche relative all'archivio situazionista di Gianfranco Sanguinetti, contenente disegni, lettere e fotografie in buona parte mai pubblicati prima di allora. Kambalu aveva attinto alla biblioteca dell'università di Yale - che aveva acquistato gli archivi di Sanguinetti - e

(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».*

²⁹ Si è soliti ricondurre l'origine giurisprudenziale della disciplina del *fair use* al caso *Folson v. Marsh* (F. Cas. 342 (1841), riguardante la riproduzione, all'interno di una biografia del presidente G. Washington, di alcuni estratti di sue lettere protette dal *copyright*. Nel decidere sulla legittimità di questa riproduzione, il giudice ritiene di non potersi affidare a un unico criterio quantitativo, ma di dover prendere in considerazione più fattori, quali «*the nature and object of the selections made, the quantity and value of the material used, and the degree in which the use may prejudice the sale, or diminish the profits, or supersede the objects, of the original works*».

³⁰ Art. 5, comma 3, lett. k) dir. 2001/29/CE del Parlamento europeo e del Consiglio, del 22 maggio 2001.

³¹ Trib. Venezia, 7 novembre 2015, in *Riv. dir. ind.*, 2018, 81.

riprodotto il materiale che aveva trovato, talvolta anche mostrando il suo gesto appropriativo (in alcune foto si vede infatti l'artista malawaiano che tiene in mano le fotografie di Sanguinetti).

Il Tribunale in questo caso ha escluso l'illecita appropriazione, rinvenendo nell'attività di Kambalu non solo un'elaborazione complessa e autonoma, ma anche una critica sarcastica dell'opera originale. A tal fine, la pronuncia richiama anche una nota sentenza della Corte di Giustizia che, nel tentativo di armonizzare sul punto la legislazione dei Paesi europei – osservando come la natura facoltativa delle eccezioni contenute nella direttiva³² non consenta agli Stati membri la libertà di stabilire i parametri delle eccezioni in modo non armonizzato tra loro – vanta il pregio di aver fornito un'interpretazione armonica e unitaria del concetto di parodia, come eccezione alla lesione del diritto d'autore: essa, in particolare, avrebbe come caratteristiche essenziali sia «quella di evocare un'opera esistente, pur presentando percettibili differenze rispetto a quest'ultima» sia «quella di costituire un atto umoristico o canzonatorio»³³.

La definizione del concetto di parodia in chiave solo umoristica, peraltro, porta con sé il limite di trascurare il tratto che più la caratterizza, ovvero quello connesso alla reinterpretazione di opere precedenti per finalità non necessariamente umoristiche, ma ad esempio critiche³⁴.

Un caso in cui la nostra giurisprudenza ha ricondotto la trasformazione di un'opera all'arte appropriativa – riconoscendone dunque la liceità – pur in assenza di fini sarcastici o parodistici è quello riguardante le opere di John Baldessari. In seguito all'apertura della mostra *The Giacometti variations*, presso la sede milanese di Fondazione Prada, la Fondazione Giacometti aveva contestato all'artista statunitense di aver riprodotto illecitamente alcune sculture *Grand femme* di Alberto Giacometti.

³² A differenza di altri Paesi europei, quali il Belgio, la Francia, il Lussemburgo, l'Italia ha scelto di non disciplinare espressamente la parodia, la cui applicazione nel nostro ordinamento è dunque il frutto di elaborazioni dottrinali e giurisprudenziali. La pronuncia della Corte di giustizia ha allora il pregio di essere intervenuta là dove la direttiva aveva fallito, fornendo una nozione di parodia che possa essere utilizzata dal diritto di tutta l'Unione. Sul punto si vd. F. GATTILLO, *Appropriazione artistica tra plagio e parodia: analisi in diritto comparato*, in *Arte e diritto*, 2022, 253 ss.

³³ Corte Giust., 3 settembre 2014, C-201/13, in *Riv. dir. ind.*, 2016, 69.

³⁴ Secondo F. BANTERLE, *L'umorismo è il miglior strumento di difesa: la nozione comunitaria di parodia per la Corte di Giustizia*, in *Riv. dir. ind.*, 2016, 89, la definizione di parodia fornita dalla Corte di Giustizia è «potenzialmente limitante», non solo perché si presta a interpretazioni soggettive, ma soprattutto perché non considera «il carattere più autentico della parodia, e cioè quello connesso allo stravolgimento e alla reinterpretazione di opere precedenti, per finalità umoristiche ma anche critiche».

Baldessarri, a onor del vero, aveva prodotto nove sculture molto simili a quelle giacomettiane – richiamate anche nel titolo del suo progetto –, seppur in scala maggiore (oltre 4 metri) e abbigliate con accessori e vestiti ripresi dal mondo del cinema e dell'arte. Nel riconoscere a tali sculture il valore di opere d'arte creative e nuove, e dunque lecite, nonostante l'artista non avesse fini parodistici o satirici nei confronti dell'opera di Giacometti, il Tribunale ha attribuito valore autonomo e autonoma dignità di tutela non solo alle opere parodistiche, burlesche o ironiche, ma più in generale a quelle che rivisitano un'opera altrui e ne mutano il senso³⁵: l'artista statunitense si era infatti servito delle figure esili giacomettiane – sfigurate e trasformate a causa degli orrori della guerra – per denunciare i canoni di magrezza imposti dalla moda.

Le due decisioni citate, seppur indicative dell'inesistenza, ad oggi, di un uniforme modello di composizione delle controversie in tema di arte appropriativa, mettono in luce come, alla base dell'appropriazionismo, vi sia per la giurisprudenza la necessità che l'autore si serva di opere altrui per esprimere un'opinione critica: essa può essere rivolta alla società – in qual caso le opere originali diventano lo strumento per trasmetterlo -, o concretarsi in un messaggio satirico che si rivolge proprio all'opera trasformata.

In questo modo l'eccezione della parodia diventa lo strumento per dare rilievo, in sede di bilanciamento degli interessi in gioco, alla libertà, costituzionalmente garantita, di espressione.

³⁵ Trib. Milano, 13 luglio 2011, in *Giur. comm.*, 2013, 109. È evidente in questo caso l'influenza della dottrina del *faire use* americano e in particolare da quanto deciso nella pronuncia *Blanch v. Koons* 467 F3d 244 (2d Cir. 2006). Il caso riguardava la riproduzione in un dipinto di Koons di una fotografia di Andrea Blanch, realizzata per una pubblicità di sandali di Gucci e raffigurante due piedi incrociati, calzati da tali sandali. L'immagine era stata ripresa da Jeff Koons in un famoso dipinto, "Niagara", di notevoli dimensioni, elemento questo messo in rilievo dalla sentenza. La sentenza, nel valutare come lecita l'opera appropriazionista, aveva rilevato che la fotografia era stata inserita nel dipinto di Koons con notevoli trasformazioni. Tale dipinto, infatti, raffigurava ben quattro coppie di piedi e non solo una, alcune delle quali erano "vestite", altre prive di scarpe, su un fondo completamente diverso da quello della fotografia, e altresì con un'inversione, poiché la fotografia originaria rappresentava le punte dei piedi rivolte verso l'alto, mentre nell'opera pittorica queste si trovano rivolte verso il basso: Koons aveva dunque realizzato un'opera del tutto diversa. Osserva peraltro F. GATTILLO, *op. cit.*, 265, come il Tribunale di Milano in questo caso si sia affidato, «nel proprio iter logico [...] ad alcuni criteri di analisi tipici della dottrina del *fair use*, senza tuttavia curarsi delle fondamentali differenze intercorrenti fra il proprio sistema di diritto e quello statunitense». Sul tema cfr. anche G. SPEDICATO, *Opere dell'arte appropriativa e diritto d'autore*, in *Giur. comm.*, 2013, 118 ss.; A. DONATI, *Quando l'artista si appropria dell'opera altrui*, in *Riv. dir. ind.*, 2018, 89 ss.; G. CASABURI, *Originalità, creatività, elaborazione creativa, citazione e plagio: profili evolutivi*, in *Il foro it.*, 2017, 3779.

La possibilità che l'opera prodotta dalla macchina abbia un fine concettuale e sia portatrice di un messaggio critico, sembra allora un presupposto imprescindibile per applicare le regole sull'arte appropriativa e considerare il prodotto dell'AI lecito anche quando la macchina si sia appropriata – attuandone una trasformazione – di opere altrui: questo è sicuramente allo stato ipotizzabile nel caso di creatività assistita, in cui vi sia un ruolo attivo dell'umano nel guidare l'algoritmo³⁶, ma non si può escludere che un'opera appropriazionista possa in futuro essere il frutto anche di una creatività automatizzata, laddove l'AI riesca a sviluppare abilità tali da trasformare opere immesse nella macchina, proprio al fine di attribuire a queste ultime un significato diverso e nuovo da trasmettere alla società.

Abstract

THE USE OF ARTWORKS IN GENERATIVE AI SYSTEMS AT INPUT AND OUTPUT PHASE

Il presente contributo si propone di analizzare la normativa vigente in materia di utilizzo delle immagini per allenare macchine intelligenti e per creare opere nuove, esaminando le sfide e le prospettive per l'uso delle opere d'arte nei modelli di AI generativa. In particolare, sono approfondite le questioni relative ai limiti del *fair use* e delle eccezioni al diritto d'autore, al concetto di rielaborazione creativa e alle possibili soluzioni regolatorie per bilanciare i diritti degli autori con l'innovazione tecnologica.

The present contribution aims to analyze the current legislation on the use of images to train intelligent machines and to create new works, by examining the challenges and perspectives for the use of works of art in generative AI models. In particular, the issues related to the limits of fair use and exceptions to copyright, the concept of creative reworking and possible regulatory solutions to balance authors' rights with technological innovation are explored.

³⁶ Attualissimo è il dibattito riguardo alla possibilità di riconoscere il diritto d'autore sulle opere create dall'intelligenza artificiale, che attualmente fanno leva soprattutto sulla distinzione tra forme di creatività assistita e forme di creatività automatizzata. In dottrina si veda V. IAIA *op. ult. cit.*, 16 ss.; G. SENA, *Intelligenza Artificiale, opere dell'ingegno e diritti di proprietà industriale e intellettuale*, in *Riv. dir. ind.*, 2020, 325 ss.; C. VENANZONI, *Intelligenze umane e artificiali. Profili giuridici nella circolazione delle opere dell'ingegno e problematiche sull'identità autorale*, in G. ALPA (a cura di), *Diritto e intelligenza artificiale*, Pisa, 2020, 297 ss.; A. BETTONI, *Tutela dell'opera d'arte algoritmica. Spunti comparatistici a margine di una recente pronuncia italiana*, in *Arte e diritto*, 2023, 399 ss.; R. ALFONSI, *Diritto d'autore e computer generated works: una decisione della Corte di Cassazione*, in *Dir. aut.*, 2023, 31 ss.; G. SPEDICATO, *Creatività artificiale, mercato e proprietà intellettuale*, in *Riv. dir. ind.*, 2019, 253 ss.