

Disposizioni e delega al Governo in materia di intelligenza artificiale A.C. 2316

"I dati sono il combustibile fossile dell'AI. Abbiamo raggiunto il picco dei dati e non ce ne saranno più".

Ilya Sutskever, Open AI co-founder - dicembre 2024

"Accogliamo la tua opposizione riguardante le tue informazioni pubbliche, come i commenti e i post pubblici [...] Ciò significa che in futuro non useremo le tue informazioni pubbliche [...] per sviluppare e migliorare i modelli di IA generativa per l'IA su META".

comunicazione META agli utenti - aprile 2025

Premessa

EGAIR - *European Guild for Artificial Intelligence Regulation* è un network che riunisce oltre 100.000 artisti, creativi, editori e associazioni di diversi Paesi europei, nata in Italia su iniziativa dell'Associazione MeFu - "Mestieri del Fumetto" per proporre alle Istituzioni europee una regolamentazione sullo sfruttamento dei dati e delle opere creative da parte delle società di intelligenza artificiale generativa. A tal fine, EGAIR ha contribuito attivamente alla stesura del testo finale del Regolamento europeo "AI Act", oltre ad aver partecipato alle indagini conoscitive sul tema dell'intelligenza artificiale in corso nel Parlamento italiano. EGAIR fa parte del Working Group 1 - Copyright nell'ambito del processo di consultazione per la definizione del General-Purpose AI Code of Practice dell'AI Office presso la Commissione europea.

L'IA generativa e il diritto d'autore

I creativi sono favorevoli allo sviluppo dell'intelligenza artificiale. Ritengono infatti che la loro attività potrebbe beneficiare dall'integrazione degli strumenti di IA generativa, se opportunamente regolamentati affinché entrambe le parti - autori e fornitori di sistemi di IA - possano trarre beneficio dalla reciproca interazione.

Ciò che si contesta non è la tecnologia in sé, ma la pratica messa in atto dalle società di IA generativa di sottrarre agli artisti, tramite **tecniche di "web scraping" selettive**, dati e opere per addestrare i propri modelli a produrre risultati paragonabili ai loro in termini di qualità, ma con costi e tempistiche irrisori.

Questi dati, oltre ad essere **sottratti illecitamente** - all'insaputa dei titolari e quindi senza il loro consenso - sono **utilizzati a scopo di profitto e in diretta e sleale concorrenza nei confronti dei titolari dei dati stessi**. Tutto ciò in palese violazione delle normative europea e nazionale in materia di tutela dei dati e del diritto d'autore, in uno scenario paradossale in cui aziende private trasformano diritti inalienabili di liberi cittadini in asset tecno finanziari.

È ormai noto che i dati sono essenziali e indispensabili per l'attività delle IA generative. Proprio l'Italia, in particolare, è "seduta su una pentola d'oro": il suo inestimabile patrimonio artistico - passato, presente e futuro. Una ricchezza che stiamo regalando ad aziende straniere, le quali non hanno alcuno scrupolo nel rivenderla sotto forma di materiale derivativo/generativo.

Tali sistemi dipendono, infatti, dalla qualità dei dati che vengono immessi nel modello ai fini dell'addestramento e tendono a rafforzare i propri errori qualora vengano addestrati sui contenuti che essi stessi producono. Si instaura perciò un processo degenerativo irreversibile: il "model decay". Ed è per tale ragione che è indispensabile che i dati utilizzati per l'addestramento siano di origine umana e non sintetica. Difatti, nella "corsa all'oro" tra Stati Uniti e Cina per la supremazia nel mercato delle AI, quell'oro siamo proprio noi cittadini, italiani ed europei.

È di qualche giorno fa l'annuncio di META circa l'utilizzo dei dati degli utenti iscritti per l'allenamento dei propri sistemi di IA. All'utente viene data la possibilità di negare il proprio consenso. Completata la procedura, all'utente

viene inviato il seguente messaggio: *“Accogliamo la tua opposizione riguardante le tue informazioni pubbliche [...] e le tue interazioni con le funzioni di IA su META. Ciò significa che **in futuro** non useremo le tue informazioni [...] per sviluppare e migliorare i modelli di IA generativa per l'IA su META”*. Da ciò, sembra che in media 15 anni di dati saranno utilizzati per l'addestramento dei sistemi di IA senza il consenso dei legittimi proprietari. Questa appare un'attività del tutto illegittima e irragionevole aggravata dal fatto che queste società rendono pressoché impossibile eliminare tali dati su richiesta dei titolari, nonostante sia un processo elementare sul piano tecnico.

Ilya Sutskever, il creatore di OpenAI, alcuni mesi fa ha dichiarato che i dati rispondono agli stessi principi di mercato dei combustibili fossili: sono pochi, finiranno e sono preziosissimi. Le Istituzioni devono difendere i propri cittadini da questo tsunami. L'Italia, con la sua storia e il suo patrimonio inestimabile, dovrebbe essere la capofila nella definizione di un quadro regolatorio per un utilizzo etico e “umano centrico” dell'IA.

Tralasciando l'impatto in termini etici del fenomeno, che già ora implica lo scenario grottesco nei fornitori di IA i quali invece di supportare l'essere umano, mirano a sostituirlo nei suoi mercati di riferimento sfruttandone direttamente il lavoro per plagiare e svenderlo su base sistemica con la promessa di una tecnologia intelligente, si pone un problema giuridico, legato al concetto stesso di diritto d'autore e all'efficacia degli strumenti previsti dalla normativa per garantirne il rispetto.

D'altronde, sulla necessità di garantire un'adeguata tutela del diritto d'autore e della proprietà intellettuale insistono le stesse Commissioni Lavoro e Attività produttive della Camera dei Deputati, nei documenti conclusivi sulle Indagini conoscitive rispettivamente promosse sul tema dell'intelligenza artificiale.

La Commissione Lavoro nel documento relativo all'indagine conoscitiva sul rapporto tra intelligenza artificiale e il mondo del lavoro, evidenzia che: *“Nell'era dell'IA, particolare attenzione va riconosciuta alla tutela del diritto d'autore e della proprietà intellettuale, con l'obiettivo primario di salvaguardare le opere dell'ingegno e i dati personali. Tale esigenza si manifesta con particolare urgenza a fronte della capacità dell'IA di generare contenuti e di analizzare grandi quantità di dati, che rischia di compromettere i diritti di autori e titolari di dati. In ordine a tale ambito, l'aggiornamento del quadro normativo deve tenere conto delle specificità dell'IA per individuare meccanismi efficaci a protezione delle opere e dei dati, nonché per prevenire e reprimere eventuali violazioni”*.

La Commissione Attività produttive, il cui focus si è concentrato sulle opportunità e i rischi per il sistema produttivo, evidenzia che: *“l'IA pone un tema di rispetto del diritto alla riservatezza dei dati, nella loro raccolta e nel loro trattamento, nonché di tutela del diritto d'autore e delle proprietà industriali, nei confronti dei creatori di contenuti. Nel corso dell'indagine conoscitiva è stato osservato che le deroghe, consentite ad esempio dall'ordinamento statunitense, all'utilizzo di opere altrimenti tutelato dal diritto d'autore per scopi scientifici potrebbero aver permesso la manipolazione e lo sfruttamento economico di numerose opere, senza il consenso e la partecipazione degli autori, soprattutto nel campo delle arti visive”*. Anche in questo caso, si richiama la *“necessità di una regolamentazione etico-legale efficace”* che garantisca la trasparenza e il controllo dei contenuti, *“prevedendo strumenti specifici a protezione di inventori e artisti”*.

Le proposte di modifica al disegno di legge

EGAIR accoglie con favore la volontà del Governo di implementare nella normativa nazionale il Regolamento europeo AI Act, con un disegno di legge sull'intelligenza artificiale che integra i principi e le regole introdotte a livello comunitario. Apprezziamo l'impostazione del provvedimento, che intende mantenere al centro l'autonomia e il potere decisionale umano e promuovere una visione “umano-centrica” di sviluppo delle tecnologie di intelligenza artificiale. Il fatto che il Governo sia intervenuto anche sul tema del diritto d'autore è per noi molto rilevante. L'Italia, infatti, è la culla dell'arte e del design, patria di artisti stimati e valorizzati in tutto il mondo. Proprio per questo è fondamentale che assuma il ruolo di capofila nella tutela degli artisti e dei loro diritti, anche in questo nuovo ambito dell'era digitale rappresentato dallo sfruttamento delle opere da parte dei sistemi di IA generativa.

Proprio alla luce di tali considerazioni, esprimiamo la nostra perplessità circa le modifiche apportate al Capo IV sul diritto d'autore nel corso dell'esame in prima lettura. Appare una **grave lacuna non aver previsto alcuna sanzione specifica per le "violazioni AI" su scala industriale delle norme sul diritto d'autore**, così come avviene, ad esempio, per il settore della pirateria. Senza l'introduzione di sanzioni adeguate, queste multinazionali straniere con ricavi miliardari proseguiranno indisturbate la loro operazione di sciacallaggio. È importante e urgente tutelare i prodotti della creatività italiana, famosa in tutto il mondo, da tali pratiche di concorrenza sleale, in palese violazione della normativa europea e nazionale di tutela della proprietà intellettuale.

Apprezziamo il richiamo alla Convenzione di Berna per la protezione delle opere letterarie ed artistiche con particolare riferimento all'articolo 6-bis della stessa che specifica esplicitamente che, anche dopo una eventuale cessione di diritti, *"l'autore conserva il diritto di rivendicare la paternità dell'opera e di opporsi ad ogni deformazione, mutilazione od altra modificazione, come anche ad ogni altro atto a danno dell'opera stessa, che rechi pregiudizio al suo onore od alla sua reputazione"*.

Di seguito i punti che riteniamo più rilevanti e sui quali chiediamo al Governo di intervenire:

- le eccezioni per scopo di ricerca;
- il consenso preventivo;
- gli obblighi di trasparenza;
- le clausole vessatorie.

L'utilizzo strumentale delle eccezioni per scopo di ricerca

Le società di IA generativa utilizzano strumentalmente l'eccezione per scopo di ricerca prevista dalla direttiva Copyright (n. 2019/790) e dal relativo recepimento nell'ordinamento nazionale. Nonostante in Europa e in Italia esista già una normativa che disciplina il legittimo uso dello scopo di ricerca, **con riferimento all'addestramento dell'IA negli ultimi abbiamo assistito all'utilizzo strumentale di tale eccezione** al solo fine di sfruttare commercialmente i dati ottenuti. Pertanto, è necessario rafforzare e specificare in dettaglio gli articoli 70-ter e 70-quater della legge sul diritto d'autore relativi allo scopo di ricerca, delimitando con maggior precisione tale legittimo uso e prevedendo sanzioni adeguate anche al fine di limitare i rischi di violazione.

Chiediamo che, pertanto:

- *l'utilizzo dei dati rispetti specifici criteri di trasparenza e informazione: devono essere comunicati l'inizio e la fine dell'attività; lo scopo, gli obiettivi e i dati utilizzati; la necessità sottesa all'utilizzo di materiale tutelato; deve essere giustificata la proporzionalità dell'uso dei dataset rispetto allo scopo nonché indicata la successiva modalità di protezione dei dati;*
- *qualora il soggetto che effettua la ricerca intenda rendere pubblici i dati, dovrà chiedere il consenso preventivo ed esplicito ai titolari del diritto per ogni singola opera utilizzata. Ove tale consenso venga negato, questi ultimi devono essere cancellati entro uno specifico termine. Qualora, al contrario, il consenso venga accordato, deve essere prevista una remunerazione proporzionata e adeguata.*

La necessità di un consenso preventivo

Il concetto giuridico di diritto d'autore su un'opera implica già di per sé una riserva sui diritti dell'opera stessa. Ne deriva, pertanto, che per utilizzare un'opera coperta dal diritto d'autore vada sempre chiesta l'autorizzazione preventiva al titolare del diritto. Perché questo non dovrebbe valere nel caso del loro sfruttamento per l'addestramento dei modelli di IA generativa?

Per fare un paragone con il concetto di proprietà privata, è come dire che un ladro possa entrare in una casa e rubare quanto c'è al suo interno solo perché non c'è un cartello all'esterno che glielo vieti e che indichi tutti gli oggetti presenti al suo interno che non possono essere rubati.

È quindi paradossale che il meccanismo alla base della tutela delle opere coperte dal diritto d'autore sia l'opt-out, che richiede che i diritti su un'opera siano espressamente riservati, quando tale riserva è già implicita nel diritto d'autore. **Appare ormai evidente il totale fallimento dell'opt-out** che ha consentito alle società di IA di appropriarsi di materiale protetto anche nei casi in cui vi era un'espressa riserva da parte dei titolari dei diritti. L'opt-out, infatti, introdotto in un periodo precedente a quello dello sviluppo e della rapida diffusione dell'intelligenza artificiale, **nasce già vecchio rispetto allo stadio di evoluzione tecnologica in cui ci troviamo** e, quindi, al contesto in cui deve operare. Oltre a non essere coerente con il concetto di diritto d'autore a cui viene associato, il meccanismo dell'opt-out è in ogni caso inefficace in termini di tutela, tanto da venire continuamente aggirato.

Coerentemente con il concetto giuridico di diritto d'autore, il **meccanismo per tutelare tali opere**, in tutti gli ambiti e quindi anche in quello del loro sfruttamento da parte delle società di IA generativa, **deve essere l'opt-in**. Inoltre, proprio come accade in tutti gli ambiti di sfruttamento di un'opera artistica coperta dal diritto d'autore, anche in quello dell'intelligenza artificiale i titolari delle opere dovrebbero essere liberi di poterne negoziare l'utilizzo, negando l'autorizzazione allo sfruttamento, oppure concedendola in cambio di una retribuzione o per un periodo di tempo determinato.

In caso di violazione, **è necessario prevedere sanzioni adeguate e proporzionate** ai ricavi miliardari delle società che forniscono sistemi e modelli di IA, peraltro già previste nella normativa italiana per le violazioni del diritto d'autore su scala industriale.

Chiediamo, dunque, che i fornitori di sistemi e modelli di IA generativa chiedano ai titolari dei diritti un consenso per l'utilizzo di materiale tutelato che sia preventivo, libero, esplicito e specifico per ciascuna opera. Ove accordato il consenso, deve essere prevista una remunerazione adeguata e proporzionata. Qualora, al contrario, tale consenso venga negato, in caso di violazione, sono previste sanzioni adeguate.

La necessità di imporre obblighi di trasparenza

Riteniamo necessario sanare la situazione pregressa, estendendo e rafforzando gli obblighi di trasparenza che l'AI Act introduce in capo ai fornitori di modelli di IA generativa. Proponiamo di introdurre un obbligo in capo a sviluppatori e fornitori di sistemi di intelligenza artificiale di rendere pubblici registri esaustivi - invece delle sintesi previste dal Regolamento europeo - contenenti tutte le opere protette dal diritto d'autore utilizzate per l'addestramento dei propri modelli. Tali registri dovrebbero essere facilmente accessibili sul web per consentire ai titolari aventi diritto di poter verificare il materiale utilizzato e, eventualmente, chiederne la cancellazione.

Chiediamo, dunque, che i fornitori di sistemi di IA generativa tengano e aggiornino registri esaustivi e pubblici relativi a tutte le opere protette dal diritto d'autore utilizzate per l'addestramento dei propri modelli.

Le clausole vessatorie previste nei contratti

Negli ultimi tempi, diverse categorie professionali quali **illustratori, attori, doppiatori e persino architetti e ingegneri**, si vedono sottoporre dai rispettivi committenti **contratti che prevedono la cessione generalizzata dei propri dati personali e relativi alle opere ai fini dello sfruttamento per l'addestramento dei modelli di IA**.

A mero titolo di esempio, all'interno dei contratti sottoposti da una piattaforma per la partecipazione ai concorsi di architettura, si legge una clausola riferita alla cessione dei dati per l'utilizzo dei progetti presentati ai fini dello sfruttamento dei propri sistemi di IA¹. Si tratta di clausole che le categorie professionali si trovano di fatto costrette ad accettare, pena la perdita della prestazione principale oggetto del contratto, senza che venga

1. "Il Partecipante accetta che l'Organizzatore abbia il diritto di utilizzare tutte le informazioni, inclusi disegni, concetti, soluzioni o altri risultati creativi, negli AI machine learning processes o di trasferirli a terzi che utilizzeranno tali informazioni negli stessi processi".



EUROPEAN GUILD
FOR ARTIFICIAL
INTELLIGENCE
REGULATION

corrisposto uno specifico corrispettivo economico per questi ulteriori utilizzi. Si tratta di un vero e proprio sfruttamento a danno di numerose categorie professionali, che riteniamo inaccettabile.

Chiediamo, dunque, che siano considerate nulle nei contratti le clausole in cui è prevista la cessione generalizzata di dati sensibili e protetti. Qualora i titolari dei diritti intendano cedere al committente i dati, tale cessione è disciplinata e quale prestazione principale oggetti del contratto e remunerata adeguatamente.

Allegato - Proposta emendativa

Dopo l'articolo 25, inserire il seguente articolo:

"Art. 25-bis

(Tutela del diritto d'autore dall'attività dei sistemi e modelli di intelligenza artificiale generativa)

1. Alla legge 22 aprile 1941, n. 633, sono apportate le seguenti modificazioni:

a) All' articolo 70-ter, dopo il comma 1, aggiungere il seguente:

1-bis. Chiunque utilizzi opere tutelate dal diritto d'autore per l'estrazione di dati per scopi di ricerca scientifica:

- a) indica pubblicamente data di inizio e di fine della ricerca e scopo e obiettivo della ricerca; motiva la necessità sottesa all'utilizzo di dati opere tutelati dal diritto d'autore; giustifica la proporzionalità dell'uso del set di dati rispetto allo scopo; indica le informazioni sulla raccolta e sulla protezione dei dati dai successivi sfruttamenti;
- b) qualora intenda utilizzare a scopo di profitto o cedere a terzi i risultati della ricerca ovvero i dati utilizzati, chiede il consenso preventivo ed esplicito ai legittimi titolari di tali dati. Ove il consenso venga negato, i dati utilizzati sono cancellati entro 30 giorni dalla ricezione del diniego di utilizzo, pena l'applicazione delle sanzioni previste dagli articoli 171 e 171-ter. In caso di accordo di cessione tra le parti, ai titolari dei dati utilizzati è riconosciuta una remunerazione adeguata e proporzionata al valore economico della cessione e dei suoi successivi addestramenti e sfruttamenti economici.

b) Dopo l'articolo 70-sexies sono inseriti i seguenti:

Art. 70-septies

1. Fatto salvo quanto previsto dagli articoli 70-ter e 70-quater, chiunque sviluppi ovvero metta a disposizione sistemi di intelligenza artificiale generativa, destinati a generare, con vari livelli di autonomia, contenuti quali testi complessi, immagini, audio o video, richiede preventivamente il consenso dei titolari delle opere protette dal diritto d'autore e dei relativi dati ai fini del loro utilizzo, anche parziale, per l'addestramento dei propri modelli. Il consenso di cui al periodo precedente può essere espresso dal titolare avente diritto in forma scritta ovvero mediante apposite misure di protezione.
2. Salvo diverso accordo tra le parti, l'utilizzo delle opere di proprietà di terzi protette dal diritto d'autore, in qualsivoglia forma, per l'addestramento dei sistemi di intelligenza artificiale generativa è disciplinato da termini e condizioni contrattuali precedentemente stabiliti con i titolari aventi diritto. Tale utilizzo è soggetto ad una remunerazione adeguata e proporzionata ai ricavi derivanti dallo sfruttamento delle opere oggetto del contratto, anche tramite accordi di licenza con il titolare avente diritto o accordi collettivi tramite le associazioni di categoria maggiormente rappresentative dei titolari aventi diritto.
3. In caso di violazione delle disposizioni di cui al comma 1, si applicano le sanzioni previste dagli articoli 171 e 171-ter.

Art. 70-octies

1. Gli sviluppatori e fornitori di sistemi di intelligenza artificiale generativa redigono registri contenenti dati esaustivi relativi a tutte le opere protette dal diritto d'autore utilizzate per l'addestramento dei propri modelli. Tali registri sono pubblicati sul sito internet dell'Agenzia per l'Italia digitale e liberamente consultabili.



EUROPEAN GUILD
FOR ARTIFICIAL
INTELLIGENCE
REGULATION

2. Con decreto del Presidente del Consiglio dei ministri, da adottarsi entro sessanta giorni dalla data di entrata in vigore della presente legge, sono disciplinate le modalità di comunicazione e pubblicazione dei dati contenuti nei registri di cui al precedente comma, improntate a criteri di semplicità e immediatezza.

c) Dopo l'articolo 156-ter è inserito il seguente:

Art. 156-quater

1. Nei contratti che hanno come prestazione la realizzazione, interpretazione o esecuzione di un'opera protetta di cui all'art.2, sono nulle le clausole in cui è prevista la cessione al committente dei dati biometrici, immagini, fotografie, musiche, testi e disegni, appartenenti o realizzati dall'autore, dall'artista interprete esecutore o dal professionista, ai fini del loro successivo utilizzo o della cessione a terzi, a qualsiasi scopo, per l'addestramento di applicazioni di intelligenza artificiale.

2. Qualora l'autore, l'artista interprete esecutore o il professionista intenda cedere al committente, in tutto o in parte, i dati di cui al comma 1, ai fini dell'addestramento di applicazioni di intelligenza artificiale, ha diritto ad una remunerazione specifica e distinta da quella relativa alla prestazione principale oggetto del contratto. Tale remunerazione è irrinunciabile ed adeguata e proporzionata al valore economico dell'opera protetta oggetto del contratto e dei suoi successivi addestramenti e sfruttamenti economici».