



**FEDERAZIONE ITALIANA
EDITORI GIORNALI**

CONTRIBUTO

della

Federazione Italiana Editori Giornali

per l'Ufficio di presidenza della VII Commissione (Cultura, scienza e istruzione) della Camera dei deputati nell'ambito della discussione della risoluzione 7-00185, recante «iniziative per garantire la tutela del diritto d'autore nell'ambito dell'utilizzo delle tecnologie di intelligenza artificiale».

Roma, 1° marzo 2024

Gli Editori associati alla FIEG ringraziano il Presidente e gli Onorevoli componenti della VII Commissione (Cultura, scienza e istruzione) della Camera dei deputati per l'invito a rendere un contributo nell'ambito della discussione della Risoluzione 7-00185, presentata dall'onorevole Amorese, recante «*Iniziativa per garantire la tutela del diritto d'autore nell'ambito dell'utilizzo delle tecnologie di intelligenza artificiale*».

Con il presente documento, nell'esprimere condivisione e apprezzamento per l'impegno del Governo all'adozione delle iniziative ivi elencate – finalizzate a salvaguardare le opportunità dell'AI (*Artificial Intelligence*), affrontandone al contempo le relative sfide strategiche – la FIEG intende evidenziare alcune questioni ritenute di più immediata rilevanza e attinenza per il settore dell'editoria giornalistica e d'informazione.

Con l'avvento dell'intelligenza artificiale ci troviamo di fronte a una sfida epocale per il mondo dei media, superiore allo stesso impatto della rete od all'introduzione degli *smartphone* e pari, come sostiene in una recente intervista Padre Benanti (il primo a parlare di “algoritica”), all'introduzione della stessa tecnologia di stampa.

Con specifico riferimento al mondo dell'informazione giornalistica e di attualità, occorre evidenziare, da un lato, l'enorme e rivoluzionario potenziale dell'AI come strumento per la raccolta, l'analisi, la produzione e la distribuzione delle notizie. Tramite l'utilizzo di algoritmi avanzati, infatti, i giornalisti possono automatizzare processi ripetitivi, analizzare grandi quantità di dati in tempo reale e generare notizie *data driven*, garantendone tempestività e precisione. Dall'altro, va anche rilevato come la capacità delle AI di creare contenuti possa essere sfruttata per produrre e distribuire notizie false o manipolative ad una velocità e pervasività mai viste prima d'ora, con il serio rischio di erodere la fiducia nelle fonti di informazione e, in ultima analisi, nella stessa professione giornalistica.

Vanno, quindi, evitati gli errori commessi in passato in situazioni analoghe, come quando l'avvento della rete e l'assenza di una regolamentazione della stessa ha portato, oltre all'innegabile sviluppo della diffusione dell'informazione, a conseguenze significative sulla struttura industriale e sulla tenuta del settore dei media tradizionali e a problemi ancora irrisolti, come ad esempio il rispetto del *copyright* in rete. In tal senso, la Risoluzione in esame coglie nel segno laddove intende impegnare il Governo ad una serie di iniziative e di azioni espressamente finalizzate a salvaguardare e sostenere

le opportunità offerte dall'AI nel mondo dei media, affrontando al contempo le relative sfide strategiche che fin da oggi si profilano.

Sfide che, dal punto di vista degli editori associati alla FIEG, paiono snodarsi, seppure con le inevitabili semplificazioni, essenzialmente in due direzioni:

- 1) tutela del diritto d'autore;
- 2) trasparenza, riconoscibilità e tracciabilità dei contenuti prodotti dall'AI generativa.

1) TUTELA DEL DIRITTO D'AUTORE

Il riferimento è alla nota questione dello sfruttamento da parte degli sviluppatori di sistemi di AI (e dei diversi soggetti che partecipano alla relativa catena del valore) di materiali editoriali protetti dal diritto d'autore, consultabili sul *web*, per il *training* degli algoritmi; nonché al problema della riproduzione di tali contenuti nelle risposte offerte dai programmi che interagiscono con gli utenti simulando conversazioni umane, i c.d. *chatbot*, che tendono a riproporre porzioni di articoli, dando luogo a copie "pervasive" dei contenuti protetti.

Sul punto può essere utile richiamare quanto dimostrato da un recente studio della *News Media Alliance* americana, l'associazione che riunisce 2.200 editori (da *Hearst* a *Condé Nast*, dal *New York Times* al *Washington Post*), dal quale, in estrema sintesi, è emerso che:

- nei *Large Language Model* (LLM) si utilizzano ampiamente contenuti informativi di giornali e media digitali, generalmente senza autorizzazione o collegamento alla fonte originaria;
- il peso dei contenuti che arrivano dagli editori è significativamente superiore rispetto a quelli liberamente utilizzabili raccolti sul web (da 5 a 100 volte);
- il contenuto editoriale è sfruttato anche nelle risposte, specie dai *chatbot* che riproducono pedissequamente pezzi e frasi degli articoli, dando luogo a c.d. "copie pervasive" che violano i diritti degli editori;
- la copia dell'AI generativa "per addestramento" non serve a uno scopo diverso dalle opere originali perché i LLM ingeriscono (cioè, copiano) le notizie in modo da imitare la stessa forma espressiva originaria e, quindi, competono anche direttamente con il contenuto protetto che è stato copiato e utilizzato per addestrarli.

Le esperienze statunitensi sulla tutela del *copyright* rispetto agli strumenti di intelligenza artificiale menzionate nelle premesse della Risoluzione in esame e la recente causa del *New York Times* nei confronti di *ChatGpt* confermano l'azionabilità di alcuni strumenti giuridici già consolidati, rispetto

ai quali tuttavia occorre individuare più dettagliatamente quali diritti tutelare e con quali più efficaci modalità.

Il settore editoriale potrebbe giovare della elaborazione di schemi contrattuali di licenza, che fungano da utile strumentario nella regolazione dei rapporti tra gli editori e gli sviluppatori, mediante l'introduzione di specifiche clausole atte ad escludere l'utilizzabilità dei contenuti oggetto di fornitura o di consultazione nell'ambito di sistemi di AI. Tuttavia, per garantire l'effettività degli strumenti di tutela, è necessario che la questione della protezione dei diritti autoriali sia affrontata non solo rispetto alle opere prodotte dall'IA generativa – definendo come e in quale misura possano a loro volta essere protette da diritto d'autore – ma anche, e con maggiore urgenza, rispetto alle opere afferenti ai contenuti utilizzati per addestrare gli algoritmi.

Margini di intervento significativi sembrano attestarsi proprio sul punto dell'effettività della tutela, ostacolata principalmente dall'asimmetria tecnologica e informativa tra le parti. In questo ambito, partendo da quanto previsto nell'*AI Act*, si ritiene assolutamente opportuna l'introduzione a livello nazionale, per via legislativa, di un obbligo di tenuta di un registro dei materiali protetti utilizzati per le finalità da addestramento. A tal fine, merita di essere evidenziata la posizione espressa, durante i lavori sull'*AI Act*, dalle associazioni europee degli editori di giornali (ENPA, che rappresenta gli editori di quotidiani ed EMMA che rappresenta gli editori di periodici e di altri media), cui è associata la FIEG. In una lettera pubblica indirizzata ai legislatori europei, firmata congiuntamente anche dalla EFJ-Federazione europea dei giornalisti, venne sottolineata la necessità di *“(…) avere una conoscenza precisa dei contenuti che sono stati utilizzati dai sistemi di AI per la formazione, l'input o qualsiasi altro scopo. La legge sull'AI dovrebbe, pertanto, basarsi sulla proposta del Parlamento di cui all'art. 28b, paragrafo 4(c), in modo da garantire che i fornitori di sistemi di AI generativi siano tenuti a tenere un registro completo, esauriente e aggiornato dei contenuti utilizzati per l'addestramento, l'input o per qualsiasi altro scopo. La proposta del Parlamento di pubblicare un 'riassunto dettagliato' dei contenuti protetti da copyright utilizzati per la formazione è sì necessaria, ma attualmente formulata in modo poco chiaro e di difficile attuazione. Un semplice riepilogo, inoltre, non riuscirebbe a fornire ai titolari dei diritti una chiara indicazione dell'eventuale sfruttamento dei loro contenuti. Pertanto, è necessario rivederne la formulazione per garantire una soluzione praticabile ed efficace. Proponiamo che i sistemi di AI siano tenuti a 'registrare automaticamente' tutti i contenuti e i dati immessi nel sistema per la formazione, l'input o per qualsiasi altro scopo. Questo processo di registrazione automatica non graverebbe in modo sproporzionato sui fornitori di sistemi di AI e consentirebbe ai creatori di verificare se i loro contenuti sono stati utilizzati. La trasparenza nella legge sull'AI non riguarda la modifica delle norme sul copyright. Tuttavia, la trasparenza è una condizione necessaria affinché i titolari dei diritti possano*

capire se i sistemi di AI rispettano le loro scelte e le loro preferenze in materia di estrazione di testi e dati. In altre parole, senza trasparenza i titolari dei diritti non potranno esercitare i loro diritti in base alle attuali norme sul copyright. Solo con la trasparenza i titolari dei diritti possono valutare se il loro opt-out (articolo 4 della Direttiva “Copyright”) è stato preso in considerazione nella pratica.”

Il richiamo all’art. 4 della Direttiva c.d. *Copyright* - il cui comma 3 afferma la necessità che la riserva dell’utilizzo dei materiali protetti da parte dei titolari dei diritti avvenga “*in modo appropriato, ad esempio attraverso strumenti che consentano la lettura automatizzata in caso di contenuti resi pubblicamente disponibili online*” – è particolarmente utile nel comparto editoriale, laddove sia in grado di rispondere effettivamente allo scopo di escludere l’utilizzabilità dei contenuti in forza dell’eccezione *text and data mining*. Con particolare riferimento a questo profilo, occorre semplificare il più possibile le procedure di espressione del consenso/divieto all’utilizzo dei propri contenuti anche a fini di addestramento, ad es. prevedendo una specifica dicitura – sulla falsariga della clausola di “riproduzione riservata” vigente nella disciplina generale del diritto d’autore (LDA n. 633/1941) – che consenta ai titolari dei diritti di manifestare, in maniera univoca e nei confronti di tutti gli operatori del mercato, i limiti di utilizzo delle loro opere.

La possibilità di una effettiva valorizzazione – anche in ambito AI – dei contenuti protetti presuppone l’effettività degli strumenti di tutela e, quindi, in concreto, un rafforzamento e una semplificazione delle modalità di esercizio, da parte dei titolari, del diritto di vietare o limitare l’utilizzo dei propri contenuti.

2) TRASPARENZA, RICONOSCIBILITÀ E TRACCIABILITÀ DEI CONTENUTI PRODOTTI DALL’AI GENERATIVA

Com’è noto, a FIEG aderiscono editori di quotidiani, periodici ed agenzie di stampa. I contenuti sono forniti su supporto sia cartaceo sia digitale (anche attraverso l’abbonamento a siti web) verso una Clientela Utenti (B2C) e Professionisti (B2B). Le condizioni generali di fornitura differiscono a seconda della clientela di riferimento, della tipologia di contenuti, e dei diritti di uso riconosciuti (dalla mera consultazione di carattere individuale, all’uso redazionale, a quello per pubblicazione o per rassegna stampa). In tutti questi casi (fatta eccezione per la copia cartacea), il Cliente riceve dei contenuti digitali che possono potenzialmente esser fatti oggetto di rielaborazione da sistemi di *machine learning*.

Molte aziende stanno rivedendo le condizioni generali di fornitura inserendo delle specifiche esclusioni relativamente all’utilizzo dei contenuti oggetto di fornitura o di consultazione da elaborazioni da parte di sistemi AI. È, però, del tutto evidente che senza un adeguato sistema di tutela,

anche con specifiche misure di *enforcement* sanzionatorio da parte del legislatore sovranazionale e nazionale, tutte le cautele ed attenzioni poste in essere dalle aziende editoriali nell'utilizzo dell'AI potrebbero vanificarsi nel momento in cui un'attività non controllata di rielaborazione con AI avesse luogo ad opera di terzi subito dopo l'avvenuta diffusione dei contenuti. È un tema estremamente delicato sul quale, riteniamo, ci sia l'esigenza di prestare una particolare attenzione.

Anche la gestione del *copyright* acquista nuove ed ulteriori complessità con la diffusione dell'AI. La quasi assoluta coincidenza tra OTT e sviluppatori di sistemi di AI è evidente a tutti. Questi operatori, che stanno mostrando una palese resistenza ad uniformarsi alle previsioni della normativa UE sui diritti connessi e agli iter negoziali in ambito AGCOM (si veda il caso del Regolamento sull'equo compenso introdotto dalla Direttiva c.d. *Copyright*), sono di fatto già in possesso di contenuti digitali raccolti dai siti *web* delle aziende (e non solo) attraverso *tools* dedicati come *robot.txt* e meccanismi simili di estrazione più o meno occulta di contenuti.

L'uso continuato e non autorizzato danneggerà i mercati esistenti che riconoscono il valore dei contenuti di qualità e, col tempo, gli stessi modelli di *Generative Artificial Intelligence* si deterioreranno. È fondamentale che le protezioni sul *copyright* siano adeguatamente applicate e che elevati standard di qualità e responsabilità siano il fondamento di queste e di altre nuove tecnologie. Se i sistemi di IA si sviluppano in concorrenza con l'industria culturale, non solo rischiano, col tempo, di impoverirla, ma finiranno anche col produrre a loro volta, nel lungo termine, contenuti di più modesta qualità, trovando nei settori creativi sempre minori risorse cui attingere.

Gli editori FIEG concordano sull'opportunità di promuovere lo sviluppo e l'impiego di soluzioni di marcatura dei contenuti prodotti mediante AI, affinché l'editore possano apporle, garantendo così la tracciabilità del contenuto medesimo. Rilevante, in quest'ambito, è l'esperienza già maturata da ANSA, che ha recentemente sperimentato una soluzione denominata "*ANSA check*" la quale consente di certificare l'origine ANSA delle notizie mediante la tecnologia *blockchain*.

Inoltre, si pone il tema del "livello di trasparenza". Anche qui tutte le attenzioni delle aziende editoriali che stanno autoregolamentando – o che presto lo faranno – la gestione dell'AI, anticipando le normative europee e nazionali in via di finalizzazione, potrebbero divenire vane senza regole chiare e vincolanti per i grandi sviluppatori dell'AI che oggi forniscono servizi senza render conto del proprio operato e sfruttando la condizione di oggettiva debolezza contrattuale in cui si trovano gli editori nazionali.