



EUROPEAN GUILD
FOR ARTIFICIAL
INTELLIGENCE
REGULATION

Siamo un Gruppo di artisti, creative, editori e associazioni europei uniti nel portare all'attenzione del pubblico come i nostri dati e le nostre proprietà intellettuali stiano venendo sfruttati senza il nostro consenso e su una scala mai vista prima. Questa situazione senza precedenti ci ha portati a unire le nostre forze per raggiungere le istituzioni europee. Se anche tu credi che i tuoi dati personali e il frutto del tuo lavoro creative non debbano essere impunemente sfruttati per profitto da una manciata di società e multinazionali, unisciti a noi in questa battaglia.

L'estate 2022 ha visto l'emergere di una nuova e incredibile tecnologia: le AI text-to-image (TTI) o image-to image. Ogni volta che viene sviluppata una tecnologia così rivoluzionaria, la nostra società deve monitorare il suo sviluppo, al fine di evitare qualunque tipo di danno o l'infrazione dei diritti umani. Questo non è ancora accaduto con la tecnologia AI. Queste tecnologie sono in grado di generare immagini o altri contenuti a partire da prompt (comandi) testuali o altri media forniti dall'utilizzatore. Per fare ciò, un'AI deve essere addestrata su un dataset di media. La qualità di un'AI generative è definita dalla qualità del suo dataset – per esempio, riguardo alle immagini, più foto e illustrazioni vengono usate nell'addestramento di un'AI, più stili quell'AI potrà replicare e più cose potrà fare. Ne consegue che i prodotti venduti dalle società di AI sono il risultato di operazioni su dataset, i quali contengono ogni tipo di dato, inclusi milioni di immagini coperte da copyright, foto private e altro materiale sensibile. Questi file sono stati raccolti attraverso un'operazione di scraping indiscriminato di tutto il web, senza alcun consenso da parte dei proprietari e delle persone rappresentate e stanno venendo utilizzati dalle società di AI per trarne profitto. Alcune di queste società stanno inoltre utilizzando i nomi degli artisti i cui lavori sono stati estratti dal web per pubblicizzare le loro AI e gli stili che esse possono replicare. Questo sfruttamento del nostro lavoro e dei nostri dati non solo non rispetta i requisiti minimi imposti dai diritti umani che regolano la nostra società: stanno anche danneggiando severamente il mercato creativo, segnandolo potenzialmente per sempre. Vediamo tutto ciò come solo l'inizio di una crisi che colpirà ogni tipologia di lavoro, creative o no. Il mercato delle arti visive è stato il primo a venir colpito unicamente per le sue vulnerabilità, che lo rendono una preda facile. È tempo di cambiare tutto ciò.

Questi sono i punti chiave che vogliamo introdurre:

- 1) Tutti i dati relativi a persone o opere, in ogni forma essi siano, siano essi già digitalizzati come file di testo, audio, video o di immagine o catturati dalla realtà attraverso videocamere, microfoni o qualsiasi mezzo di registrazione, non devono poter essere utilizzati per il training di modelli di AI senza l'esplicito consenso informato dei loro legittimi proprietari.
- 2) L'utilizzo di nomi di persone, nomi d'arte o opere che non sia coperto da una licenza di sfruttamento per il training di AI deve essere proibito dai software che consentono di inserire un "prompt" testuale o vocale per richiedere la generazione di un'immagine, video, testo o suono.
- 3) L'utilizzo di video, immagini, suoni e testi che non sia coperto da una licenza di sfruttamento per il training di AI deve essere proibito dai software che consentono di inserire media per richiedere la generazione di un'immagine, video, testo o suono.
- 4) Deve essere istituito un sistema di indicizzazione e certificazione standardizzato, "human e machine readable" delle attività delle AI e dei dataset di immagini, testi o suoni interamente o parzialmente prodotte con AI. Diciture come "interamente generato con AI", "realizzato utilizzando materiale generato con AI" dovrebbero diventare lo standard.
- 5) Non appare sufficiente il discrimine tra materiale "copyrighted" e "di pubblico dominio". Nei dataset possono infatti trovarsi anche dati personali sensibili, quindi protetti dalle leggi sulla privacy, ma non da copyright. Al tempo stesso, esistono casi di materiale pubblicato in un momento in cui non si poteva prevederne l'utilizzo nel contesto di dataset per il training dei modelli di Intelligenza Artificiale. Ogni singolo dato che viene inserito nei modelli di training deve essere curato e autorizzato dai legittimi proprietari e quindi inserito in maniera volontaria e informata dai singoli autori. Le aziende di AI devono produrre internamente materiali originali per il training o licenziare contenuti esterni secondo termini e condizioni contrattuali precedentemente stabilite con gli autori interessati.

Scopri di più su www.egair.eu