“An AI” for your thoughts? Qualche riflessione sull’uso delle nostre conversazioni per il training dell’intelligenza artificiale generativa

di Michela Leggio - Osservatorio sullo Stato digitale, 22 novembre 2023

La Generative AI richiede un volume sempre più grande di dati per essere addestrata e resa più performante, e sono i suoi stessi fruitori, più o meno consapevolmente, a metterli a disposizione. Ma è possibile trovare un giusto bilanciamento tra diritto alla privacy e interesse a progettare tecnologie sempre più avanzate?

L’intelligenza artificiale (Artificiale Intelligence, AI) fa ormai parte delle nostre vite ed è noto ai più come per il suo funzionamento servano essenzialmente tre elementi: un algoritmo, un hardware con elevata potenza computazionale e una grande quantità di dati.

Meno noto è invece il processo con cui essa si ottiene e cioè l’addestramento dell’algoritmo. Per spiegarlo, si può ricorrere alla metafora con cui Turing chiudeva il suo famoso saggio sull’intelligenza artificiale nel 1950, dove questa era paragonata alla mente di un bambino: così come quest’ultimo impara andando a scuola, si può dire che anche la prima necessiti di “apprendere”, scovando collegamenti tra dati per essere poi in grado di riprodurre autonomamente la logica che vi sta dietro, simulando così il ragionamento umano. Per raggiungere questo risultato è fondamentale, però, avere a disposizione un’elevatissima quantità di dati che siano al contempo rispettosi di un certo standard qualitativo.

Con l’arrivo della AI generativa, termine usato per descrivere le tecnologie digitali “in grado di creare diversi tipi di contenuti come testi, audio, immagini, video”, questa fase è diventata ancora più importante. Ciò è vero, in particolare, per i Large Language Model – LLM’s, anch’essi parte della generative AI, ovvero reti neurali che, grazie a modelli strutturati su innumerevoli parametri, sono addestrate per mezzo di miliardi di dati a svolgere un determinato compito, come suggerire una parola, una frase o un intero discorso: ne è perfetto esempio ChatGPT di OpenAI. Proprio tale tecnologia ha destato l’attenzione dei giganti del web come Google, Microsoft e Meta, i soli (o i pochi) ad avere le conoscenze tecniche, le risorse finanziare necessarie e l’accesso a un patrimonio di dati immenso per poter sviluppare un sistema di intelligenza artificiale generativa.

Se tuttavia un’intelligenza artificiale sempre più performante abbisogna di un training a sua volta più intenso, è gioco forza che i dati in possesso delle Big Tech non siano sufficienti, ragion per cui si è posto il problema di reperirli altrove. Per far fronte a tale mancanza, si è ricorsi innanzitutto al c.d. web scraping, scaricando dati da pagine web liberamente accessibili e attirando l’ira di numerose testate giornalistiche, tra cui il New York Times. Non bastando i contenuti liberamente reperibili online, però, si sono iniziati utilizzare anche dati provenienti dagli stessi utenti che ogni giorno si avvalgono di applicazioni offerte dalle piattaforme.

Soffermandosi su questo profilo, è emerso infatti come Google stia utilizzando i messaggi di posta su Gmail per allenare l’intelligenza artificiale, analizzando le risposte alle proposte di “Help Me Write”, la funzione di Gmail che suggerisce agli utenti come concludere una frase. Allo stesso modo Microsoft sta utilizzando le chat con Bing per allenare i bot a rispondere meglio alle domande e perfino Zoom, qualche mese fa, ha affermato di essere intenzionato a usare i contenuti delle video chat per migliorare i sistemi di intelligenza artificiale a cui sta lavorando. Peraltro, anche i post sui social non sarebbero immuni, specie se provenienti da profili non privati.

Ma quale libertà hanno le Big Tech di utilizzare le interazioni degli utenti per addestrare l’intelligenza artificiale e quali sono i rischi per la privacy?

La questione non appare perfettamente sovrapponibile a quelle emerse in passato, sempre relative alla privacy online. Per quanto, infatti, non vi siano dubbi che le informazioni estratte dalle comunicazioni degli utenti siano qualificabili alla stregua di dato personale ai sensi del GDPR, deve evidenziarsi come ciò che importa, nel training dei LLM's, sia la forma o la stessa struttura con cui l'informazione è trasmessa, piuttosto che le sue componenti individuali e personali. Ciò che interessa agli sviluppatori di AI sono quindi il linguaggio, l'interazione scritta e le connessioni logiche tra parole, non importando il loro collegamento con la persona fisica da cui l'espressione proviene.
Oltre a sedare i timori di una sorveglianza digitale, i casi sopra analizzati consegnano un ulteriore elemento: i dati captati dalle interazioni degli utenti si collocano sempre all'interno di un'applicazione loro offerta "gratuitamente". Si pensi all'account Gmail e al servizio di scrittura guidata, così come la piattaforma Zoom, le chatbot di Bing o ancora i social network. Emerge quindi una logica di corrispettività, similmente a quanto affermato dal Consiglio di Stato con la pronuncia n. 2631/2021, poichè sembra intravedersi uno scambio tra i dati estrapolabili dalle interazioni degli utenti ai fini di training e l'utilizzo delle applicazioni che le Big Tech offrono e offriranno sempre più in futuro. Sembrerebbe quindi che il training dell'intelligenza artificiale e, conseguentemente, lo sviluppo di sistemi più performanti e avanzati vada a beneficio degli stessi consumatori che se ne avvalgono e che partecipano, inconsapevolmente, alla loro stessa creazione.
È bene però sottolineare che, per quanto estranea a logiche di profilazione degli utenti, tale pratica presenta comunque dei rischi, specie sotto il profilo della sicurezza del dato personale, che risulta comunque esposto ad attacchi informatici dai quali potrebbe derivare una sua diffusione incontrollata e un suo riutilizzo da parte di agenti malevoli per altri scopi.
Anche per questa ragione è imprescindibile che la scelta relativa all'utilizzabilità delle proprie interazioni online sia sempre rimessa alla persona da cui esse provengono. La possibilità di scegliere l'opt-out dovrebbe essere sempre presente e comunicata in forma trasparente e chiara, e non, come spesso avviene, presentata all'interno di disclaimer il cui contenuto non è sempre intelligibile da parte dell'utente medio. Questa sembra essere stata la scelta adottata da Google che ha da poco aggiornato la propria policy di privacy garantendo adesso la possibilità di negare il proprio consenso all'utilizzo dei dati personali per lo sviluppo di intelligenza artificiale.
Un'altra via per ovviare al problema potrebbe essere poi l'anonimizzazione o la pseudoanonimizzazione dei dati raccolti, per quanto poi si ponga sempre il problema di controllare l'effettiva anonimità del dato. Un'altra proposta, che potrebbe costituire una soluzione definitiva al problema, proviene invece dallo stesso mondo della tecnologia. Di recente, infatti, si è iniziato a parlare di dato sintetico, ovvero di dati creati dalle stesse macchine e somiglianti ai dati del mondo reale che possono sostituirsi ad essi per diversi scopi, tra cui anche l'addestramento dell'intelligenza artificiale.
Non resta quindi che attendere e vedere se non possa essere proprio la tecnologia ad offrire delle soluzioni a problemi che da essa stessa originano.