

È il momento di parlare (seriamente) di etica dell'intelligenza artificiale

Mancano meno di due settimane alla chiusura della finestra temporale che il **Garante della privacy** italiano ha dato a **OpenAi** per [adeguarsi alle sue prescrizioni sull'uso dei dati personali](#) e *correggere la rotta* su **ChatGPT**: come vi avevamo raccontato, la startup fondata da **Sam Altman** ha tempo fino al 30 aprile prossimo per rispettare le regole europee in tema di informativa sull'uso dei dati personali, diritti degli interessati (utenti e non) e base giuridica del trattamento delle informazioni per allenare l'algoritmo. Al di là di come finirà questa storia, comunque, i problemi e le alzata di sopracciglio legati alle **intelligenza artificiale**, e in particolare agli algoritmi conversazionali come per l'appunto ChatGPT, sono ben lontani dall'essere risolti – ne è testimone l'ormai [famosa lettera](#) in cui **Elon Musk**, i ricercatori di **DeepMind** e altri imprenditori e scienziati hanno chiesto uno stop di sei mesi nello sviluppo dell'intelligenza artificiale, in attesa di avere “*regole comuni e protezioni dalle conseguenze più deleterie*”.

Che se ne condividano o meno i contenuti, comunque, la lettera racconta solo parte della storia. E, a voler pensar male, l'impressione è che almeno alcuni dei suoi autori stiano soltanto cercando di *prendere tempo* per non rischiare di essere tagliati fuori dal succulento mercato dell'intelligenza artificiale ([e difatti...](#)). A sostenerlo, tra gli altri, sono gli esperti di [SIpEIA](#), la **Società Italiana per l'Etica dell'Intelligenza Artificiale**, associazione fondata da esperti di informatica, etica e giurisprudenza delle **Università degli Studi di Torino**, **Sapienza Università di Roma**, **Università della Calabria**, **Università di Bologna** e del **Consiglio Nazionale**

delle Ricerche, che sta per pubblicare un **manifesto** su ChatGPT e gli altri Large Language Model (Llm) – di cui si è tra l'altro discusso il 17 aprile scorso a Torino, nel corso dell'evento *“ChatGPT: promesse e illusioni. Limiti e potenzialità del fenomeno informatico del momento”*. *“Parte della classe degli intellettuali organici e degli imprenditori vuole ‘raccattare le briciole’ di questi strumenti – ci ha raccontato **Guido Boella**, professore al Dipartimento di Informatica dell'Università di Torino e cofondatore di SIpEIA – e usa la [visione lungotermista](#) per puntare il dito verso problemi lontani e non ben definiti, anziché evidenziare quelli molto più immediati e concreti, relativi in particolare alle interazioni tra economia, società e tecnologia”*. Come vi avevamo raccontato, la narrazione lungotermista cara a Elon Musk e colleghi mette in guardia, per esempio, rispetto alla possibilità che *“menti non umane potrebbero superarci di numero, essere più intelligenti di noi, renderci obsoleti e rimpiazzarci”*: ma si tratta, come [hanno più volte ribadito i massimi esperti del settore](#), di una suggestione distopica e assurda dal punto di vista scientifico e tecnologico, specie se si considera che un'intelligenza artificiale del genere non solo non è nemmeno ancora in vista, ma non è neanche in alcun modo chiaro se sarà mai possibile svilupparla.

Una riflessione necessaria

Per una volta, quindi, sarebbe meglio guardare il dito anziché la luna, e preoccuparsi di questioni più immediate. Magari guardando a quello che è successo nel passato: *“Quando la comunità scientifica cominciò a sviluppare le biotecnologie – dice ancora Boella – fu subito chiaro che c'era l'esigenza di definire una nuova etica, perché quella tradizionale non aveva gli strumenti sufficienti per avere a che fare con quelle innovazioni. Oggi siamo in uno scenario molto simile: i cambiamenti portati dalle intelligenze artificiali richiedono un cambiamento dell'etica. Di più: a differenza delle biotecnologie, **l'intelligenza artificiale ha ripercussioni su***

molti più aspetti della nostra quotidianità, e dunque definirne un'etica è molto più urgente. È per questo che due anni fa abbiamo fondato SIpEIA, società scientifica e di advocacy composta da informatici, filosofi, giuristi ed eticisti".

Detto, fatto: proprio su questo è incentrato il *position paper* di SIpEIA. *"Bisogna evitare – si legge nel documento – che il business model che caratterizza il capitalismo della sorveglianza non diventi il modello di monetizzazione anche dei Llm, perché l'impatto sarebbe molto maggiore. Infatti, i Llm creano 'relazioni sintetiche' con gli utenti, coinvolgendoli in una relazione anche emozionale e portando a nuove e più profonde modalità di estrazione dei dati, nonché possibilità di manipolazione degli utenti. Noi umani, infatti, attribuiamo a chi comunica con noi stati mentali ed emozioni, che i Llm by design non hanno, e siamo quindi esposti a manipolazione da parte loro".*

I dubbi e le paure

Ma quali sono, nello specifico, **i rischi** ravvisati dagli esperti, quelli che rendono cogente la definizione di linee guida etiche? Anzitutto la questione dell'**estrazione dei dati** – quella che ha mobilitato il Garante – che, per forza di cose (le Ai vivono di dati), diventerà sempre più intrusiva. La sociologa **Shoshana Zubov**, nel suo *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*, pubblicato nel 2019, ha paragonato l'estrazione dei dati a quella del petrolio: ne serve sempre di più, e bisogna scavare sempre più in profondità. *"Fino a non molto tempo fa – dice Boella – Google raccoglieva i dati delle ricerche degli utenti; poi si è passato a un livello ancora più personale, cioè scandagliare quello che si pubblica sui social, i testi sulle app di messaggistica e così via. Con ChatGPT e simili l'estrazione sarà ancora più personale, perché ancora più personale è il dialogo che l'utente ha con un*

*chatbot. Un'estrazione così profonda può portare a nuovi livelli di pubblicità e addirittura di manipolazione: supponiamo di chiedere a ChatGPT di scrivere un racconto sull'autunno; sapendo, per esempio, che in passato abbiamo fatto ricerche sul Giappone, potrebbe citarci le foglie di acero del Giappone, per indurci a comprare un volo per il Giappone. Cosa succederebbe invece nel caso ci fossero delle elezioni politiche? Lo scenario – e la manipolazione – potrebbero essere **ancora più preoccupanti**".*

Un'altra questione affrontata nel manifesto è quella relativa al **mondo della formazione**: la possibilità che **compiti a casa, elaborati o anche intere tesi siano scritte non dagli studenti ma da un Llm in real time** ha suscitato un'ondata di preoccupazione nel mondo delle scuole e dell'università. E anche in questo caso gli esperti ravvisano **una grande ipocrisia** nella narrativa corrente: "OpenAi" dice Boella "ha messo a disposizione un sistema che riconosce se è un testo è scritto o meno da un essere umano. Peccato **funzioni il 25% delle volte, cioè quasi mai**. Avrebbe potuto fare una cosa molto più semplice, che non richiede neanche l'utilizzo di tecnologie Ai. OpenAi, come tutte le piattaforme web e i motori di ricerca, tiene traccia nei suoi registri (i logfile) di tutte le richieste degli utenti e di tutte le risposte fornite dal chatbot per migliorare il prodotto e per raccogliere informazioni su di noi (da monetizzare). Certo per identificare il plagio da parte di studenti più motivati nel copiare, che fanno una parafrasi del testo ottenuto da ChatGPT, occorrerebbe solo qualche semplice strumento di Ai che vada a dare una misura della distanza del testo prodotto dal bot. **Ma siamo sicuri che nella maggior parte dei casi basterebbe un semplice 'ctrl+F' sul file di log di ChatGPT?**".

E ancora: un altro tema eticamente rilevante è quello relativo alle possibili interazioni di ChatGPT e simili con i **minori**. "Lascereste i vostri figli a parlare con uno sconosciuto?" è la domanda retorica degli esperti di SIpEIA "per di più sapendo che è [uno sconosciuto che può soffrire di](#)

[allucinazioni?](#)”. Il *position paper* non offre soluzioni, ma sottolinea la necessità di una maggiore sensibilizzazione e soprattutto di un dibattito più sensato e centrato sul presente. *“È davvero il momento di agire: ma al centro della nostra preoccupazione non dovrebbero esserci immaginarie ‘potenti menti digitali’. Al contrario, **dovremmo concentrarci su pratiche di sfruttamento molto reali e molto concrete** da parte delle aziende che stanno sviluppando questi strumenti, e che stanno rapidamente centralizzando il potere e aumentando le disuguaglianze sociali”*.

La sfida del confronto

C'è dell'altro. All'evento ha partecipato anche **Elena Esposito**, professoressa di sociologia dei processi culturali e comunicativi alle Università di Bielefeld, in Germania, e Bologna. Esposito è autrice di [Comunicazione artificiale – come gli algoritmi producono intelligenza sociale](#), un libro appena edito da **Bocconi University Press** in cui invita ad affrontare il “*problema*” delle intelligenze artificiali da un altro punto di vista: anziché chiederci se e come le macchine siano diventate più intelligenti di noi (e abbiamo visto che non è così), bisognerebbe capire come intervengono nella comunicazione. Ossia, in altre parole, parlare di **comunicazione artificiale** più che di **intelligenza artificiale**: *“Il fatto che possiamo comunicare con le macchine – spiega Esposito – non implica che esse abbiano una loro intelligenza che deve essere spiegata ma che, anzitutto, la comunicazione sta cambiando. **L'oggetto di questa ricerca non è l'intelligenza, che è e rimane un mistero, ma la comunicazione, che possiamo osservare e di cui sappiamo già molto.** Occorre un concetto di comunicazione che sia in grado di tener conto anche delle possibilità che il partner comunicativo non sia un essere umano ma un algoritmo. Il risultato, che può essere osservato già oggi, è una condizione in cui disponiamo di informazioni di cui spesso nessuno può ricostruire né comprendere la genesi, ma che ciononostante non*

sono arbitrarie. Le informazioni generate automaticamente dagli algoritmi non sono affatto casuali e sono del tutto controllate, ma non dai processi della mente umana. **Come possiamo controllare questo controllo, che per noi può essere anche incomprensibile?** Questa è, a mio parere, la vera sfida che ci pongono oggi le tecniche di machine learning e l'uso di big data". La tesi dell'esperta è quindi che se le macchine contribuiscono all'intelligenza sociale non è perché hanno imparato a pensare come noi, ma perché hanno imparato a partecipare alla comunicazione: ed è con questo scenario che dobbiamo imparare a confrontarci. Senza paura, se possibile: "Platone diceva che la scrittura avrebbe fatto perdere l'uso della memoria agli esseri umani – conclude Esposito – e in parte è successo. Con le intelligenze artificiali, e con la comunicazione artificiale, potremmo perdere qualcos'altro. Ma va bene così. Perderemo qualcosa, ma acquisiremo dell'altro".