

Poteri e responsabilità sulle intelligenze artificiali

Alberto Puliafito, giornalista

6.11.2024

“Da un grande potere derivano grandi responsabilità” è una frase che pronunciano Peter e Ben Parker (rispettivamente, Spiderman e suo zio) nel film *Spider-Man* del 2002 diretto da Sam Raimi. La storia della frase e delle sue modifiche successive, però, ha qualche anno in più, com'è facile immaginare. Nell'ambito delle ia, le personalità della Silicon valley non usano in modo responsabile né la loro visibilità né il loro potere: i profili social di chi ha in mano le aziende che oggi sviluppano nuove tecnologie sono una vetrina, uno strumento di marketing quasi gratis. Elon Musk va perfino oltre spingendosi a fare palesi campagne elettorali (poco sappiamo di quelle meno palesi). Soldi e visibilità insieme garantiscono a queste élite il mantenimento e l'amplificazione dei loro privilegi.

Scendendo di qualche gradino, anche le persone che godono di grande visibilità nell'ecosistema dei media hanno, in qualche modo, potere e dovrebbero agire con più responsabilità, perché i loro messaggi – spesso ipersemplicati – si diffondono più velocemente di altri. È un'epoca in cui è impossibile essere generalisti e avere un'opinione su tutto. Ma, al tempo stesso, la necessità di mantenere una visibilità costante, di cavalcare le tendenze del momento e di parlare delle cose di cui parlano tutti gli altri favorisce la creazione di contenuti che rischiano di essere inaccurati. In generale, questo accade anche nella conversazione pubblica sulle ia: appaiono molto spesso contenuti “forti”, “socializzabili”, che contribuiscono a costruire l'immaginario collettivo. È importante che questo immaginario non sia distorto.

Leggi anche
Musk non è al di sopra della legge

Facciamo un paio di esempi. Luca Bizzarri, attore e conduttore televisivo, ha realizzato una puntata del suo podcast *Non hanno un amico* intitolata *Intelligenza artificiale o ballista seriale?*. Nell'episodio Bizzarri racconta di aver usato ChatGpt per chiedergli informazioni: non

sappiamo se la versione gratuita o quella a pagamento, non sappiamo con quali prompt. Cosa succede poi? Che ChatGpt inventa. Poi ritratta. Poi “mente” ancora quando Bizzarri torna a chiedergli se si ricorda cosa avesse detto. Infine, l’attore commenta: “L’intelligenza artificiale a cui affideremo la medicina, i trasporti, la produzione industriale, la nostra vita è un raccontaballe”.

La frase, come tutta la puntata, è molto efficace perché Bizzarri fa bene il suo lavoro. Tuttavia ci sono molte imprecisioni. Prima di tutto, c’è un’umanizzazione eccessiva di ChatGPT: per quanto funzionale al racconto, non è così che vanno le cose. ChatGPT non è umano. Poi sappiamo che un large language model non è un oracolo onnisciente né un’enciclopedia e non andrebbe consultato come tale; abbiamo imparato che queste macchine possono avere “allucinazioni” – il termine è infelice e troppo umanizzante, ma purtroppo dobbiamo tenercelo – e che le versioni a pagamento hanno prestazioni superiori a quelle gratuite. Inoltre non è affatto vero che a quel tipo di modelli “raccontaballe” vengono affidate decisioni in settori cruciali come la diagnostica medica, per esempio.

Stabilmente tra i venti podcast italiani più ascoltati, con grande visibilità, *Non hanno un amico* di fatto ha diffuso un’ipersemplificazione sulle ia e ChatGpt che rischia di inquinare tutto il discorso sulle intelligenze artificiali: quella puntata è uscita su varie piattaforme, è stata ascoltata, commentata, diffusa, linkata o condivisa con altre persone per “spiegare le intelligenze artificiali”.

Anche chi ha incarichi istituzionali si lascia andare spesso a dichiarazioni che lasciano più di una perplessità. Una larga fetta del mondo politico internazionale ha ceduto, pubblicamente, alla narrazione catastrofista sulle ia, spesso evitando o relegando ai margini i discorsi scomodi: sulle armi, sulla sorveglianza, sull’impatto ambientale. Non fa eccezione l’Italia.

Leggi anche
Privacy

Paolo Benanti, francescano, docente alla Pontificia università gregoriana, è stato nominato da Giorgia Meloni presidente della commissione sull’intelligenza artificiale per l’informazione. Il 17 ottobre, nell’audizione al senato per la cosiddetta commissione Segre su “Fenomeni di intolleranza, razzismo, antisemitismo e istigazione a odio e violenza”, ha parlato di vari problemi legati alle ia. Ho seguito spesso le sue audizioni e i suoi interventi: con una comunicazione molto efficace, anche fatta di aneddoti personali e esempi per arrivare a tutti, introduce e ripete una serie di concetti che fanno parte delle sue convinzioni sul tema.

Qualche esempio: dice che per guidare le automobili viene richiesta una patente e che questo modello potrebbe servire anche per usare le ia, per poi precisare “non so se e quanto sia

efficace e applicabile. Però potrebbe essere un modello a cui ispirarsi”. Attribuisce agli algoritmi e alle notifiche un’efficacia manipolatoria: “il grande problema dei sistemi di cui stiamo parlando”, dice, “è che [...] senz’altro sono manipolativi e questo è il paradosso: noi potremmo avere fenomeni d’odio senza neanche reale convincimento”. Mette in dubbio la scientificità di vari studi sulle ia dicendo che ne fioriscono di “analoghi a quelli che sono fioriti nel campo del tabacco e delle armi, cioè studi che [...] si classificano come scientifici ma hanno alle spalle lobby a vari livelli”. Sostiene che “le intelligenze artificiali [...] danno risposte in una maniera unica che potremmo definire oracolare. Quello che la macchina dice è un contenuto al quale credere come valido, in forza di una fiducia che a priori uno mette sulla macchina”.

Se le semplificazioni di Bizzarri erano facili da confutare, l’impianto del discorso di Benanti è molto più solido e coerente, contiene elementi reali e concreti e non ci sono prese di posizione tipo “ChatGpt mente”. Però le cose sono più complesse.

Sì, la profilazione serve per provare a venderti più prodotti e i messaggi pubblicitari sono in qualche modo manipolatori, esattamente come la propaganda politica, da sempre, serve a convincerti di qualcosa. Sì, ci sono studi che presentano una commistione di interessi pubblici e privati e alcuni provengono da aziende che hanno tutto l’interesse a magnificare i risultati delle ia, ma il mondo scientifico contiene al suo interno anche gli anticorpi per contrastare gli errori e la propaganda e non propone verità assolute. L’anticorpo principale è il metodo scientifico stesso, per cui non si propongono verità a cui credere ma si segue un processo per produrre una conoscenza affidabile e verificabile ma non necessariamente definitiva. Sì, ci sono rischi legati alle ia. Ma l’idea di una patente è un’idea prima di tutto politica. Sì, con le ia si producono anche contenuti falsi, ma sono le persone a produrli e anche con la scrittura e la parola si producono contenuti falsi. E poi, chi dice che diamo fiducia a priori alle macchine?

Mettendo tutti questi concetti in un unico calderone, il messaggio che passa continua, purtroppo, a essere fuorviante. Il rischio è di aprire la strada – anche se animati da ottime intenzioni – a soluzioni con effetti collaterali anche molto pericolosi. La patente per le ia, casomai venisse in mente a qualcuno di applicarla, non sarà mai per gli Elon Musk. L’ia che racconta bugie e che dunque non va usata da tutti perché produce anche contenuti falsi potrebbe portare a limitazioni dell’accesso all’informazione o favorire forme di controllo sulle informazioni, su chi può produrle o diffonderle.

Ecco perché nei podcast, nelle commissioni, sulle newsletter e, in generale, nel discorso pubblico abbiamo tutti grandi responsabilità.

Questo testo è tratto dalla newsletter Artificiale.

Iscriviti a
Artificiale

Cosa succede nel mondo dell'intelligenza artificiale. Ogni venerdì, a
cura di Alberto Puliafito.

Vedi tutte le
newsletter

Iscriviti

Iscriviti a
Artificiale

Cosa succede nel mondo dell'intelligenza artificiale. Ogni venerdì, a cura di Alberto Puliafito.

Iscriviti

Vedi tutte le newsletter

Internazionale pubblica ogni settimana una pagina di lettere. Ci piacerebbe sapere cosa pensi
di questo articolo. Scrivici a: posta@internazionale.it