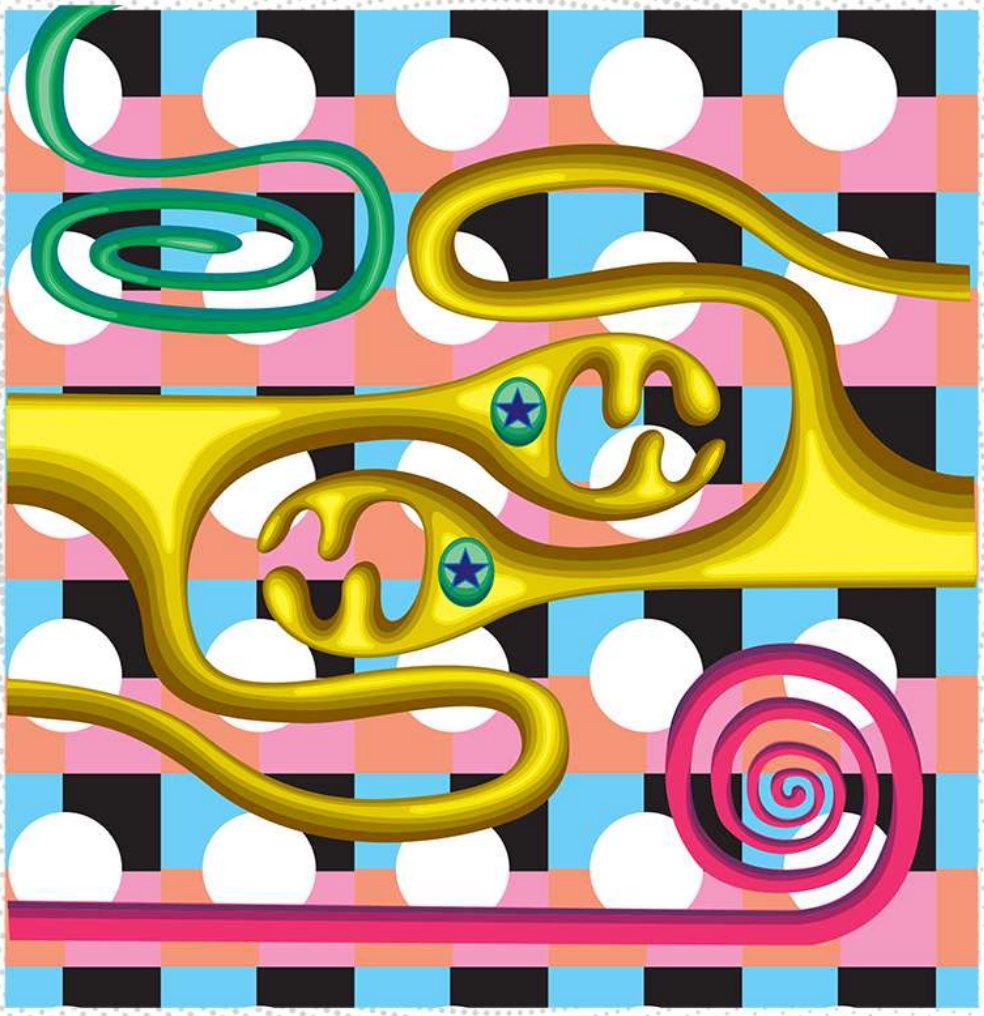
Generazione AI

Una monografia sull'intelligenza artificiale



GUERRE DI RETE

Gli ebook di Guerre di Rete

Generazione AI

Una monografia sull'intelligenza artificiale

A cura di Carola Frediani, Sonia Montegiove e Federico Nejrotti

Con i contributi di Stefano Casini, Antonio Dini, Carola Frediani, Federica Meta, Giuditta Mosca, Federico Nejrotti, Antonio Piemontese e Andrea Signorelli

GUERRE DI RETE

Associazione Culturale Guerre di Rete
www.guerredirete.it
info@guerredirete.it

Associazione Culturale Cyber Saiyan
www.cybersaiyan.it
info@cybersaiyan.it

Curatela: Carola Frediani, Sonia Montegiove, Federico Nejrotti

Copertina: Luca Ottonelli, stampa digitale su plexiglass, senza titolo, 2006 (gentilmente concessa dall'autore)

Immagini: Special Collections and Archives, Cardiff University - Computer Image Corporation Archive - Concordia University, Montreal - George Washington University Libraries - Glenn Research Center Collection - Brown University - University of Texas

Data di pubblicazione: 3 settembre 2023

Indice

Premessa

Generazione AI è una monografia sull'intelligenza artificiale - realizzata da Guerre di Rete (guerredirete.it) e curata da Carola Frediani, Sonia Montegiove e Federico Nejrotti - che vuole introdurre alcune delle questioni più rilevanti dell'intelligenza artificiale a un ampio pubblico, in linea col mandato editoriale di Guerre di Rete.

Guerre di Rete è un progetto editoriale non profit che nasce con l'obiettivo di fare informazione su temi legati al mondo della cybersicurezza e dei conflitti digitali, dando conto della loro complessità e rilevanza sociale.

È il frutto dell'impegno di due associazioni culturali: l'omonima Guerre di Rete, che da anni pubblica una newsletter settimanale dedicata a temi digitali, e Cyber Saiyan, una community indipendente di professionisti del settore della cybersicurezza.

Guerre di Rete si sviluppa intorno al modello di "slow journalism", non rincorrendo la notizia giornaliera ma cercando di dare senso e completezza alle storie e alle vicende più significative, in maniera trasparente, con linguaggio chiaro e toni misurati.

Oggi è totalmente sostenuta:

- da donazioni di lettori, sostenitori, amici e colleghi che credono nella nostra idea di informazione;
- delle associazioni Guerre di Rete e Cyber Saiyan;
- dal lavoro volontario dello staff di Guerre di Rete: Gerardo Di Giacomo, Carola Frediani, Giovanni Mellini, Sonia Montegiovane, Federico Scalco e Patrizio Tufarolo.

Introduzione

L’AI non è un Paese per pochi

Non è neutrale, non è scontata né definita. Ma poiché influenza la distribuzione del potere deve riguardare tutti.

di **Carola Frediani**

Per anni, agli occhi del grande pubblico e dei media, il termine intelligenza artificiale (IA o all’inglese AI, Artificial Intelligence) ha avuto lo stesso fascino e la medesima concretezza dell’espressione Big Data. Un guscio utile per convegni e paper, con pochi effetti visibili sul quotidiano o la società. Poi nell’autunno 2022 sono arrivati ChatGPT, la corsa al lancio di prodotti basati su AI generativa, la possibilità di giocare o sperimentare con una miriade di strumenti - spuntati come funghi giorno dopo giorno - e la competizione fra le grandi aziende tech per rilanciare i propri servizi all’insegna di questa tecnologia.

È così iniziato un ciclo industriale e mediatico, fatto di annunci, investimenti, hype e dichiarazioni di ricercatori, che ha alzato una cortina fumogena su quel che è nuovo e quel che esiste da tempo; su quel che è rivoluzionario e quello che invece è reazionario; sui rischi effettivi e quelli presunti; su chi fa progredire il settore e chi è pronto a speculare; su chi trarrà vantaggio e chi verrà sfruttato.

Siccome le cortine fumogene non fanno mai bene all’informazione occorre ripartire dunque da alcuni elementi fondamentali. Quali sono le aziende in gioco e quale il ruolo di multinazionali consolidate come Microsoft, Google, Facebook? Quali elementi sono di novità e quali rischiano di essere gonfiati dalla grancassa che si è sviluppata attorno al settore? Che ruolo hanno la società civile, la politica, gli Stati di fronte a un panorama fatto di aziende private, concentrazione geografica, nonché di ricercatori in netto contrasto fra di loro sulla capacità, l’impatto e i rischi conseguenti a questa rivoluzione, sempre che si possa definire in tal modo?

Procediamo con ordine e partiamo dall’aspetto materiale. Facendo prima una premessa: studiosi di diversa natura riconoscono come il termine AI sia vago e ambiguo fin dalla sua nascita. Alcuni di questi, come Yarden Katz, professore al Department of American Culture and Digital Studies Institute della University of Michigan, ritengono addirittura questa vaghezza e ambiguità funzionali a un uso ideologico di tale tecnologia. Ovvero, la nebulosità del concetto permette di

reinterpretarlo continuamente sulla base di determinati interessi (che si tratti di aziende che promettano nuovi mercati agli investitori, di Stati che vogliano rafforzare la sorveglianza o di apparati militari che cerchino di legittimare armi "intelligenti").

Ciò premesso, le basi di questa tecnologia (estremamente materiale, radicata nello sfruttamento di risorse, energia, dati e forza lavoro umana come sviscerato più volte da Kate Crawford, professoressa, ricercatrice e autrice del libro Atlas of AI) sono al momento concentrate in un manipolo di grandi aziende, Stati e aree geografiche.

Fino al 2014, la maggior parte dei più importanti modelli di machine learning (apprendimento automatico) sono stati rilasciati dal mondo accademico. Ma da allora, secondo il report AI Index della Stanford University, l'industria ha preso il sopravvento. Perché "la costruzione di sistemi di AI all'avanguardia richiede sempre più spesso grandi quantità di dati, calcoli e soldi, risorse che gli operatori del settore possiedono intrinsecamente in quantità maggiore rispetto alle organizzazioni non profit e al mondo accademico".

Per l'Institute for Human-Centered Artificial Intelligence dell'università americana l'AI è "sempre più definita dalle azioni di un piccolo gruppo di soggetti del settore privato, invece che da una più ampia gamma di rappresentanti della società".

Negli stessi Stati Uniti, la maggior parte degli investimenti sono concentrati attorno ad alcuni hub e centri specifici. Nel 2021, secondo un report dell'istituto Brookings, le città di San Francisco e San Jose, in California, esprimevano da sole circa un quarto degli articoli di conferenze, dei brevetti e delle aziende di AI a livello nazionale.

Certo, molte speranze sono riposte nei modelli aperti, open source, di AI. Fondamentali anche per garantire processi trasparenti, un fattore che a cascata ha ricadute in molti ambiti diversi. Ma, come evidenzia il rapporto di Stanford, se è vero che l'open source e la concessione di licenze diffuse di modelli linguistici di grandi dimensioni possono decentralizzare l'industria, nello stesso tempo bisogna ricordare che l'ampia distribuzione di una tecnologia digitale non implica automaticamente un'equa distribuzione dell'innovazione, della creazione di posti di lavoro e della leadership esecutiva. Insomma, nulla per ora può darsi per scontato, se non che partiamo da una polarizzazione di potere in alcune aziende e accademie.

Ma non si tratta solo di ricerca e di industria, questo accentramento riguarda anche il dibattito ideologico e politico sull'AI. Nel corso degli ultimi mesi, i media hanno dato grande spazio ai sostenitori dell'esistenza di un rischio esistenziale per l'umanità derivante da una AI superintelligente che possa superarci e addirittura sfuggire al nostro controllo (va detto che questo scenario è sempre presentato in modo estremamente vago e non è affatto chiaro come ciò possa davvero avvenire).

Il Center for AI Safety parla addirittura di un rischio estinzione derivante dalla AI, per cui il rafforzamento della governance di questa tecnologia dovrebbe essere una priorità globale al pari della prevenzione delle pandemie e delle guerre nucleari. Uno scenario rafforzato dalle dichiarazioni di ricercatori che hanno fatto la storia del deep learning, come Geoffrey Hinton e Yoshua Bengio, oltre che del CEO di OpenAI Sam Altman (ne avevo scritto in vari numeri della newsletter Guerre di Rete).

Ma come scrivono una serie di accademici e studiosi del settore, "concentrarsi sulla possibilità che una superintelligenza fuori controllo uccida la specie umana può essere dannoso di per sé. Potrebbe distrarre le autorità di regolamentazione, l'opinione pubblica e gli altri ricercatori di AI dal lavoro per mitigare rischi più urgenti, come la sorveglianza di massa, la disinformazione e la manipolazione, l'uso improprio dell'AI in campo militare e l'inadeguatezza del nostro attuale paradigma economico in un mondo in cui l'AI svolga un ruolo sempre più importante".

Inoltre, la discussione esistenziale rischia di oscurare una serie di questioni spinose e attuali che riguardano lo sfruttamento, senza consenso o retribuzioni, di dati, contenuti e opere prodotte da umani per addestrare i modelli di AI (che poi producono contenuti e opere che rischiano di togliere ossigeno e lavoro a quegli stessi umani, come dimostrato dalle preoccupazioni, le cause legali e le proteste di artisti e autori).

O ancora, la perpetuazione e la legittimazione scientifica di bias, pregiudizi e ingiustizie strutturali sicuramente precedenti alla AI ma che la sua applicazione miope rischia di cristallizzare (pregiudizi e sperequazioni di potere che si abbattono addirittura sulle stesse ricercatrici di AI che per prime ne hanno scritto, anche quando si parla di normazione ed etica del settore).

O l'impatto sul mondo del lavoro, dove per alcuni lo scenario, più che di robot ultraintelligenti che sostituiscano in tutto gli umani, è quello di un'amplificazione del lavoro precario e parcellizzato. Così come di una turbo-burocrazia che

potrebbe rendere ancora più iniqua l'attuale distribuzione di risorse e potere (pensiamo agli scandali che hanno riguardato sistemi algoritmici per regolare la gestione di welfare e benefit dentro la stessa Europa).

Si parla anche di come l'AI ambisca, fin dai suoi esordi e ancor più nell'attuale dibattito sulla superintelligenza, a una ridefinizione non tanto delle macchine, ma dell'umano.

Ne hanno scritto e parlato in tanti, più o meno direttamente, da Erik J. Larson (in The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do), alla professoressa di linguistica e direttrice del laboratorio di linguistica computazionale dell'università di Washington Emily Bender (in più luoghi, come qua) ma anche il professore del MIT Joseph Weizenbaum, l'inventore negli anni '60 del primo chatbot ELIZA, di cui parliamo anche in questo speciale. Eppure, ammoniva quest'ultimo nel suo libro Computer Power and Human Reason: From Judgement to Calculation, "c'è una differenza tra l'essere umano e la macchina e ci sono certi compiti che non si dovrebbe far fare ai computer, indipendentemente dal fatto che possano essere eseguiti dai computer". Questione densa, che meriterebbe uno speciale a parte, e lascio qui solo come spunto di riflessione.

"Qualunque cosa sia l'AI, non è neutrale, e nemmeno noi possiamo esserlo", scrive Dan McQuillan nel libro Resisting AI, forse il più critico e radicale di tutti, ma non c'è bisogno di sposarne interamente la visione per cogliere una buona parte della sua brillante analisi. "L'AI è politica perché agisce nel mondo attraverso dinamiche che influenzano la distribuzione del potere, e le sue tendenze politiche si rivelano nel modo in cui stabilisce confini e separazioni".

Negli ultimi mesi il parallelismo tra la ricerca sull'AI e quella sul nucleare è stato fatto più volte da molteplici soggetti. Non importa, in questa sede, che si tratti di un paragone corretto o meno (personalmente ritengo che ci siano più diversità che somiglianze e il solo fatto di avanzare questo parallelismo tende a legittimare la visione dei rischio-esistenzialisti), perché una riflessione in questo senso può comunque avere una sua utilità. Tra tutti i libri o i film citati, quello che a mio avviso andrebbe riletto è quel saggio storico fenomenale di Robert Jungk, tradotto in italiano con il titolo Gli apprendisti stregoni. Storia degli scienziati atomici (Einaudi 1958), dove vengono esplorati i rapporti fra i fisici nucleari, la società e la politica.

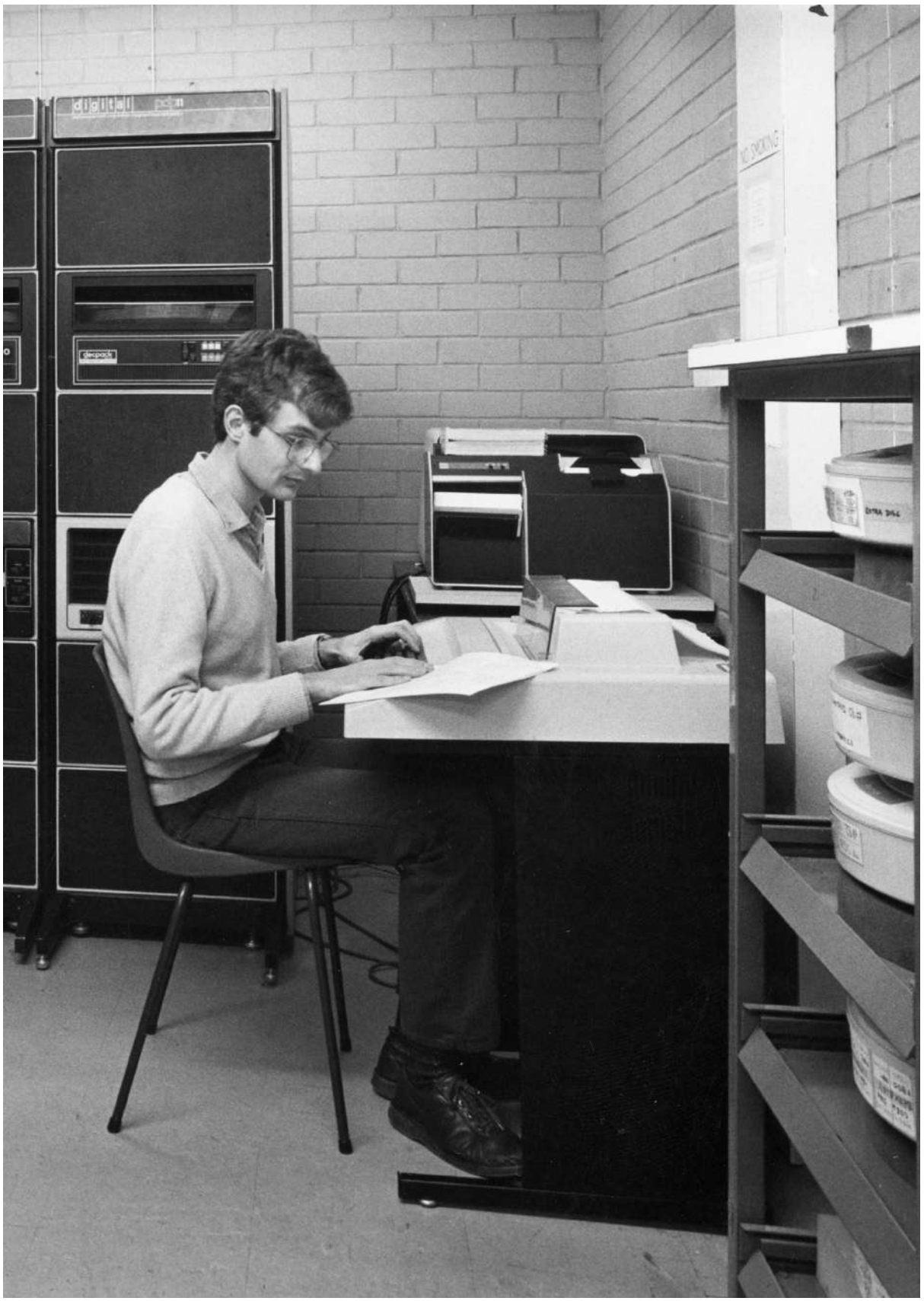
"Quasi in ogni epoca c'è un campo del pensiero e dell'attività umana che attira con forza particolare gli spiriti dotati", scrive Jungk, "così in certe epoche gli

spiriti inquieti, tutti protesi al nuovo, sono portati particolarmente all'architettura; in altre alla pittura o alla musica, alla teologia o alla filosofia. Improvvisamente - e nessuno potrebbe dire come - i più aperti avvertono in che punto esattamente si è dischiusa una breccia, e si spingono là dove possono sperare di non restare semplici discepoli, ma di divenire anch'essi fondatori e maestri. Proprio una siffatta forza di attrazione ebbe la fisica atomica negli anni che seguirono alla Prima Guerra Mondiale".

Vale oggi lo stesso per l'AI? Ci troviamo davvero di fronte a una breccia intellettuale? E se è così, davvero "viviamo su un'isola di fulmicotone… per accendere il quale, grazie a Dio, non abbiamo ancora trovato il fiammifero", come scriveva nel 1921 il fisico tedesco Walter Nerst?

Se anche così mai fosse, dalla storia, come quella splendidamente raccontata da Jungk, un insegnamento possiamo trarlo: di fronte a una tecnologia davvero dirompente, non possiamo lasciare le decisioni né ai generali né ai soli scienziati. Politica, società civile, esperti di altre discipline e cittadini devono influenzare le scelte su come questa breccia trasformerà le loro vite.

"Ogni persona deve agire come se l'intero futuro del mondo dipendesse da lei", scriveva ancora Weizenbaum, "qualsiasi cosa di meno è un restringimento della responsabilità e forza disumanizzante, perché qualsiasi cosa di meno incoraggia gli individui a vedersi come un mero attore in un dramma scritto da agenti anonimi, come meno di una persona, e ciò è l'inizio della passività e dell'assenza di scopo".


digital
pdp-11
dec pack

Società

C’era una volta un chatbot

Cosa ci insegna la storia di ELIZA, il primo chatbot che sembrava conversare come un umano.

di **Andrea Signorelli**

Tra i tanti ruoli che ChatGPT ha rapidamente iniziato ad assumere nelle vite dei milioni di utenti che lo utilizzano su base quasi quotidiana, ce n’è uno probabilmente inatteso. Per molti, il sistema di OpenAI con cui è possibile conversare su ogni argomento, e spesso in maniera convincente, è diventato un amico, un confidente. Addirittura uno psicologo.

Una modalità non prevista (almeno esplicitamente) da OpenAI, ma scelta da un numero non trascurabile di utenti, che si relazionano a ChatGPT come se davvero fosse un analista. Per impedire un utilizzo giudicato (per ragioni che vedremo meglio più avanti) improprio e pericoloso, OpenAI impedisce al suo sistema di intelligenza artificiale generativa di offrire aiuto psicologico, che infatti di fronte a richieste di questo tipo si limita a fornire materiale utile da consultare. Ciò però non ha fermato gli “utenti-pazienti” che, su Reddit, si scambiano trucchi e tecniche per sbloccare ChatGPT affinché fornisca loro consigli psicologici.

Joseph Weizenbaum e il suo chatbot ELIZA

Un risvolto che potrebbe sorprendere molti. Uno dei pochi che sicuramente non si sarebbe sorpreso e che avrebbe avuto moltissimo da dire sull’argomento è Joseph Weizenbaum, scienziato informatico e docente al MIT di Boston, scomparso nel 2008. Colui che già parecchi decenni prima della sua morte aveva preconizzato – o meglio, affrontato e approfondito in prima persona – molti degli aspetti che portano le persone a relazionarsi in maniera intima con le macchine e le cause di questo comportamento.

È infatti il 1966 quando Weizenbaum crea il primo chatbot della storia: ELIZA, il più antico programma informatico in grado di comunicare per via testuale con una persona, e capace di intrattenere una conversazione almeno parzialmente

coerente. Così battezzata in onore di Eliza Doolittle, personaggio del Pigmalione di George Bernard Shaw che vuole imparare a esprimersi in maniera forbita, ELIZA è progettata per imitare in maniera caricaturale il comportamento di uno psicologo della scuola rogersiana, che – come scrive lo stesso Weizenbaum in Computer Power and Human Reason, il suo testo del 1976 – è un metodo "relativamente semplice da imitare, perché buona parte della sua tecnica consiste nel rivolgere al paziente le sue stesse dichiarazioni in forma di domanda".

Sempre nel saggio, Weizenbaum riporta anche un dialogo in chiave psicologica che una persona ha avuto con ELIZA (con cui peraltro è possibile ancora oggi comunicare, per esempio a questo link):

> Paziente: "Ho bisogno di aiuto, di questo sono certa"
>
> ELIZA: "Che cosa significherebbe per te ricevere aiuto?"
>
> P: "Forse riuscirei ad andare d'accordo con mia madre"
>
> E: "Dimmi di più della tua famiglia"
>
> P: "Mia madre si prende cura di me"
>
> E: "Chi altri nella tua famiglia si prende cura di te?"
>
> P: "Mio padre"
>
> E: "Tuo padre"
>
> P. "Per certi versi assomiglio a mio padre"
>
> E: "Che somiglianza vedi?"
>
> P: "Non sei per niente aggressivo, ma penso che tu non voglia che io lo noti"
>
> E: "Che cosa ti fa pensare che io non sia aggressivo?"
>
> P: "Non discuti con me"
>
> E: "Perché pensi che io non discuta con te?"
>
> P: "Sei spaventata da me"
>
> E: "Ti fa piacere pensare che io sia spaventata da te?"

Come si vede dalla conversazione di cui sopra, ELIZA risponde in maniera coerente quando ha qualche appiglio facile su cui incentrare la sua risposta (per esempio il termine “madre” la porta subito a porre una domanda sulla famiglia), mentre in tutti gli altri casi si limita a ribaltare in forma di domanda l’affermazione fatta dal “paziente”.

Essendo un programma sviluppato negli anni Sessanta, quando il machine learning era ancora un modello pioneristico, il funzionamento di ELIZA ha ben poco a che fare con gli strumenti generativi a cui ci stiamo abituando oggi. “Il conversatore umano doveva digitare la sua porzione di conversazione su una macchina da scrivere collegata a un computer, e il computer, controllato dal mio programma, avrebbe analizzato il messaggio che gli veniva trasmesso e composto una risposta in inglese”, spiega sempre nel suo libro Weizenbaum.

“Poiché le conversazioni devono riguardare qualcosa (vale a dire che devono essere inserite in un contesto), il programma è stato progettato su due livelli”, continua Weizenbaum. “Il primo livello consiste in un analizzatore del linguaggio e il secondo in uno script. Uno script è un insieme di regole che possono forse ricordare quelle che vengono date a un attore a cui è chiesto di improvvisare su un determinato tema. Di conseguenza, differenti script potevano insegnare a ELISA ad avere una conversazione su come cuocere le uova, gestire un conto bancario e così via. Ogni specifico script permetteva a ELIZA di avere uno specifico ruolo conversazionale”.

Come visto, la scelta di Weizenbaum per la prima sperimentazione di ELIZA è poi ricaduta sulla figura dello psicologo rogersiano, il cui stile è facilmente (e grossolanamente) imitabile per via informatica. Ma qual era l’obiettivo del docente del MIT? A differenza di quanto si potrebbe pensare, Weizenbaum non voleva mostrare le potenzialità degli allora nascenti nuovi strumenti informatici. Anzi, voleva dimostrare l’esatto contrario: quanto il livello di interazione e conversazione tra essere umano e macchina fosse ancora estremamente superficiale. “Ciò che non avevo però compreso è che un’esposizione anche molto breve a un programma informatico relativamente semplice potesse provocare reazioni deliranti in persone altrimenti decisamente normali”, prosegue Weizenbaum.

Le reazioni inaspettate al chatbot

Le cose, infatti, non andarono secondo i suoi programmi. “Weizenbaum intendeva mostrare tramite ELIZA quanto la comprensione informatica del linguaggio umano fosse ancora ridotta”, scrive Oshan Jarow su Vox. “Gli utenti instaurarono però immediatamente uno stretto rapporto con il chatbot, passando ore di fila in sua compagnia per condividere conversazioni intime”.

Un fenomeno – da allora diventato noto come ELIZA Effect – che colpì moltissimo Weizenbaum (“Ero stupefatto da quanto rapidamente e profondamente le persone che conversavano con ELIZA si facessero coinvolgere emotivamente e come la antropomorfizzassero”) e che, in almeno un’occasione, lo lasciò addirittura esterrefatto. L’aneddoto lo racconta lui stesso: “Una volta la mia segretaria, che mi aveva visto lavorare al programma per molti mesi e quindi sapeva in prima persona che si trattava di un semplice programma informatico, iniziò a conversare con ELIZA. Dopo qualche scambio, mi chiese di lasciare la stanza”. Una richiesta che il professore interpretò come la necessità, da parte della sua assistente, di avere un po’ d’intimità con la macchina.

Da allora, Weizenbaum ha trascorso il resto della sua vita ad avvertire dei rischi di lasciare che i computer, il mondo informatico e l’allora embrionale intelligenza artificiale giocassero un ruolo troppo importante nella società. Non per il timore, oggi nuovamente in voga, che questi sistemi apparentemente intelligenti possano un giorno dominare il mondo, ma per il rischio di affidarsi eccessivamente a strumenti in realtà assolutamente inattendibili e ai quali rischiamo di cedere eccessiva responsabilità e libertà.

Un tema ancora adesso di incredibile attualità, che nell’ottica di Weizenbaum si combinava con l’eccessiva fiducia (anzi, fede) riposta nella scienza e nella logica (confusa per razionalità). Al punto che, scrive lui stesso, “la fede nell’equazione tra razionalità e logica ha corroso il potere profetico del linguaggio. Possiamo contare, ma stiamo rapidamente dimenticando cosa vale la pena contare e perché”. Righe che hanno oggi tutto lo stesso peso di allora e che varrebbe la pena di far leggere, per fare solo un esempio, a filosofi come Nick Bostrom (uno dei più influenti tra le élite della Silicon Valley), figura che ha spinto talmente tanto sull’associazione tra logica e razionalità da arrivare a sostenere, nella cosiddetta “ipotesi della simulazione”, la possibilità che oggi potremmo tutti vivere in una simulazione informatica.

Proprio con l’obiettivo di smontare l’effetto di ELIZA su una parte consistente di utenti (Weizenbaum racconta come a un certo punto fosse diventato una sorta

di “passatempo nazionale”), il docente decise di pubblicare una spiegazione dettagliata del funzionamento del suo chatbot: “Una volta che un particolare programma viene smascherato, una volta che il suo funzionamento interiore viene spiegato in un linguaggio sufficientemente chiaro da essere compreso da tutti, la sua magia si sbriciola”, spiegava proprio in quel testo. Ma anche questa si rivelò un’illusione: il testo di Weizenbaum suscitò molto meno interesse della possibilità di poter conversare con una macchina.

Chi ha seguito l’evoluzione dell’intelligenza artificiale negli ultimi dieci anni, segnati dalla diffusione di una tecnologia trasformativa e di enorme impatto come il deep learning (apprendimento profondo, una branca del machine learning), non potrà che rimanere impressionato dai tanti parallelismi tra ciò che si sta verificando oggi e ciò che così profondamente aveva colpito Weizenbaum quasi cinquant’anni fa, malgrado le evidenti diversità tecnologiche fra i due strumenti.

Certo, ELIZA era un semplice programma di meno di 200 righe di codice, mentre ChatGPT nasce, nella sua prima versione, da una rete neurale dotata di 175 miliardi di parametri, addestrata su un ampio corpus di contenuti e alimentata da un potere computazionale incomparabile.

L’irresistibile fascino di ELIZA (e di ChatGPT)

ELIZA era stata creata da un docente nel suo studio accademico, mentre ChatGPT è stato progettato da una realtà finanziata con decine di miliardi di dollari e per cui lavorano i massimi esperti del settore. ELIZA poteva quasi solo ribaltare in forma di domanda le nostre affermazioni, mentre ChatGPT è in grado di conversare in maniera sofisticata e spesso sorprendente su praticamente ogni tema, attraverso inferenze statistiche sui dati. Tra le tante differenze, c’è però una somiglianza cruciale. Un filo rosso che lega questi due strumenti appartenenti a epoche così diverse: la necessità dell’essere umano di confidarsi con entità terze, che appaiono neutre, oggettive e prive di pregiudizi.

Non è infatti una coincidenza se, a distanza di così tanto tempo, ChatGPT e il suo antenato ELIZA sono stati utilizzati da una parte della popolazione per gli stessi identici scopi. Che cosa porta, però, un buon numero di persone a rivolgersi a strumenti automatici invece che a psicologi in carne e ossa, o ad amici o parenti?

Negli ultimi anni, e in particolare durante i lockdown provocati dalla pandemia da COVID-19, è emerso chiaramente come questi strumenti vengano utilizzati (e spesso appositamente progettati) anche per contrastare "l'epidemia di solitudine" che affligge la società occidentale contemporanea. I numeri sono chiarissimi: secondo una ricerca della Commissione Europea del 2021, il 25 per cento degli abitanti del Vecchio Continente afferma di sentirsi solo "la maggior parte del tempo", un netto peggioramento rispetto ai livelli comunque già preoccupanti degli anni precedenti. Per esempio, un'analisi de Il Sole 24 Ore basata su dati Eurostat e risalente al 2017 mostrava come, già prima dei lockdown, il 13,2 per cento degli italiani over 16 affermasse di soffrire di solitudine. Secondo una ricerca di Harvard, negli Stati Uniti questa percentuale arriva addirittura al 35 per cento.

Abbiamo inevitabilmente meno dati relativi agli anni Sessanta. Parecchie analisi mostrano però come la diffusione della solitudine nelle società occidentali (per ragioni complesse che non è qui il caso di affrontare) abbia le sue radici – o almeno venga riconosciuta – proprio negli anni precedenti alla comparsa di ELIZA. Nel 1950, per esempio, il sociologo David Riesman pubblica il saggio La folla solitaria, mentre nello stesso decennio la psichiatria inizia ad affrontare seriamente il problema: "La solitudine sembra essere un'esperienza così dolorosa e spaventosa che le persone farebbero praticamente di tutto per evitarla", scrive in un saggio del 1959, citato dal New Yorker, la psichiatra tedesca Frieda Fromm-Reichmann.

Tra le tante cose che possono mitigare la solitudine, comunicare con un chatbot non sembra nemmeno la più bizzarra. Potrebbe essere questo, allora, a spiegare l'utilizzo di ELIZA come confidente che tanto stupì e indignò Weizenbaum. Ed è sicuramente ciò che spiega il successo, oggi, dei tanti chatbot nati esplicitamente allo scopo di tenere compagnia. Il più noto di questi, Replika, ha per esempio visto un aumento nel numero di utenti del 35 per cento rispetto alla fase pre-pandemica, raggiungendo oggi i due milioni di utenti attivi.

Oltre a Replika troviamo anche Chai, Kuki, Anima e parecchi altri ancora, tutti nati per fornire compagnia e che spesso prevedono anche la possibilità di sviluppare una relazione romantica con il nostro compagno virtuale. "Abbiamo notato quanta richiesta ci fosse per un ambiente in cui le persone potessero essere se stesse, parlare delle loro emozioni, aprirsi e sentirsi accettate", ha spiegato al The Guardian la programmatrice Eugenia Kuyda, fondatrice di Replika.

Al di là dei possibili lati positivi (per esempio, la possibilità di fornire una valvola di sfogo a chi ne ha urgente bisogno) e negativi (mettere una pezza tecnologica a problemi molto più radicati), c'è un altro elemento di cruciale importanza: non solo i potenziali "pazienti" cercano nei chatbot uno psicologo o almeno qualcuno che li aiuti a sconfiggere la solitudine, ma gli stessi psicologi immaginano un futuro in cui il loro lavoro verrà svolto almeno in parte da questi strumenti.

L'ultima dimostrazione viene da un paper pubblicato sul Jama Internal Medicine, in cui i ricercatori hanno mostrato come – a giudizio degli stessi medici che hanno partecipato a un test cieco – le risposte fornite ai pazienti da ChatGPT fossero nel 79 per cento dei casi di qualità e soprattutto di empatia superiore a quelle dei professionisti. Va comunque sottolineato come, in altri casi, ChatGPT e altri strumenti simili abbiano invece dato dimostrazione di comportamenti estremamente preoccupanti nei confronti degli utenti (consigliando per esempio a una persona di suicidarsi o prescrivendo cure completamente sballate).

Per quanto la strada sia ancora lunga, è possibile pensare che davvero, in futuro, degli strumenti avanzati e potenti come ChatGPT e affini possano svolgere ruoli così delicati? In realtà, potrebbe essere un ulteriore abbaglio: già negli anni Sessanta alcuni psicologi avevano infatti preso talmente sul serio ELIZA da ipotizzare in alcuni studi la possibilità di impiegare al loro posto questo rudimentale software, dotato quasi esclusivamente, come già detto, della capacità di rigirare in forma di domanda le affermazioni dei "pazienti".

Tra questi ci fu il dottor Kenneth Colby, che in un articolo scientifico – riportato da Weizenbaum nel suo saggio – scrisse: "Ulteriore lavoro dev'essere fatto prima che questo programma sia pronto per l'uso clinico. Se il metodo dovesse però rivelarsi benefico, potrebbe allora fornire uno strumento terapeutico da rendere ampiamente disponibile nei centri psichiatrici che soffrono di carenza di psichiatri. (...) Lo psichiatra umano, coinvolto nella progettazione e nell'utilizzo di questo sistema, non verrà rimpiazzato, ma diventerà un professionista molto più efficiente, visto che non dovrà più limitarsi a occuparsi di un paziente per volta".

Opinioni di questo tipo (che ricordano da vicino alcune considerazioni odierne sul rapporto tra essere umano e macchina in ambito professionale) non vennero comunque soltanto da Colby, ma da parecchi altri luminari, tra cui spicca la presenza di un nome di peso come quello di Carl Sagan (comunque non un addetto ai lavori, essendo stato un celebre astrofisico).

Il lascito di Weizenbaum

Come prevedibile, Weizenbaum reagì malissimo a queste ipotesi: “Che lavoro pensa di star facendo uno psichiatra per pensare che la più elementare parodia meccanica di una singola tecnica di analisi possa aver catturato anche solo una minima parte dell’essenza di un rapporto umano? Quale può essere l’immagine che lo psichiatra ha del suo paziente, nel momento in cui vede se stesso non come un essere umano coinvolto in un processo di guarigione, ma come un elaboratore informatico che segue delle regole precise?”.

Le stesse osservazioni vennero ribadite anche in un’intervista rilasciata al *The New York Times* nel 1977: “Ci sono aspetti della vita umana che un computer non può capire. È necessario essere umani. L’amore e la solitudine hanno a che fare con le più profonde conseguenze della nostra costituzione biologica. Questo genere di comprensione è per principio impossibile per un computer”.

Se l’illusione di intimità ottenuta dagli utenti – e le eccessive aspettative di alcuni psicologi – sono simili, gli effetti che uno strumento come ChatGPT potrebbe avere sulla società sono invece molto più profondi di quelli, modesti, di un sistema rudimentale come ELIZA. “Parlare con ELIZA era essenzialmente una conversazione con se stessi: qualcosa che la maggior parte di noi fa ogni giorno nella sua testa”, si legge ancora su Vox. “Nel caso di ELIZA avevamo a che fare con un partner privo di una qualunque personalità, felice di continuare ad ascoltarci fino al momento in cui era spronato a porci qualche semplice domanda. Che le persone trovassero conforto e catarsi in questa modalità di condivisione dei loro sentimenti non è poi così strano”.

Il caso di ChatGPT e dei suoi simili è però radicalmente diverso: “Parlare con la nuova generazione di chatbot non significa parlare con se stessi, ma con un enorme agglomerato di discorsi digitalizzati. A ogni interazione cresce inoltre il corpus di dati utilizzabili per l’addestramento”. Questo avviene proprio per la struttura stessa dei Large Language Model (LLM, ovvero i modelli linguistici di grandi dimensioni come GPT3, addestrati su enormi dataset e con grandi quantità di parametri), che si limitano a distillare dai miliardi di testi con cui sono stati addestrati la formulazione con la maggiore probabilità di rispondere in maniera statisticamente “sensata” e che da queste interazioni apprendono ulteriormente.

Inevitabilmente, essendo i dati usati per l’addestramento forniti dagli esseri umani, nelle risposte degli LLM troviamo uno specchio della società che ha prodotto quei dati. Se parlare con ELIZA era un po’ come parlare con se stessi,

qui siamo di fronte a qualcosa dalle implicazioni molto più profonde: non solo perché le nostre interazioni modificano il comportamento della macchina, ma perché modificano anche il nostro comportamento. Questo vale in modo più evidente ogni volta che chiediamo suggerimenti a ChatGPT, e in modo meno evidente quando lasciamo che l'algoritmo di Facebook o TikTok filtri per noi le informazioni, che quello di Amazon ci suggerisca cosa comprare, quello di Netflix ci segnali cosa vedere e una app come LifeCycle ci dica come gestire le nostre giornate.

"L'intelligenza artificiale sta oggi attivamente dando forma a una parte significativa delle nostre vite", prosegue Jarow su Vox. "In particolare, usiamo i chatbot per aiutarci a pensare e per dare forma ai nostri pensieri. Tutto ciò può avere grandi benefici, come semplificare la produzione di alcuni semplici contenuti professionali. Ma può anche ridurre la diversità e la creatività che sorge dall'impegno umano nel dare voce alla propria esperienza. Per definizione, gli LLM suggeriscono un linguaggio prevedibile. Se ci affidiamo a essi troppo massicciamente, l'algoritmo prevedibile diventiamo noi stessi".

A questo punto non sorprenderà che Weizenbaum avesse affrontato lo stesso tema in un'intervista del 1985, spiegando come "affidarsi eccessivamente ai computer è soltanto il più recente – e più estremo – esempio di come l'essere umano usi la tecnologia per fuggire al fardello di agire come un essere indipendente". Tutto ciò, nell'ottica di Weizenbaum, non significava evitare gli strumenti informatici, ma sfruttarli con consapevolezza e cautela. Senza consegnare a essi – a partire ovviamente dalla psichiatria – compiti di grande responsabilità, in cui gli errori hanno gravi conseguenze e che richiedono un approccio molto più elastico di quanto sia in grado di fare una macchina.

"Poiché al momento non abbiamo modo di rendere saggi i computer", concludeva nel suo saggio Weizenbaum, "non dobbiamo dare ai computer nessun compito che richieda saggezza". Parole scritte 45 anni fa, ma che – in un'epoca in cui gli algoritmi di deep learning vengono utilizzati in ambiti di enorme delicatezza come la giustizia, la selezione dei posti di lavoro, la sorveglianza, la sanità e altro ancora – andrebbero ancora attentamente ascoltate.

Replika, l'amico virtuale che diventa partner romantico

L'idea di creare un chatbot con cui comunicare è venuta alla programmatrice e imprenditrice russa Eugenia Kuyda in seguito alla scomparsa di un caro amico: Roman. Attraverso il machine learning e

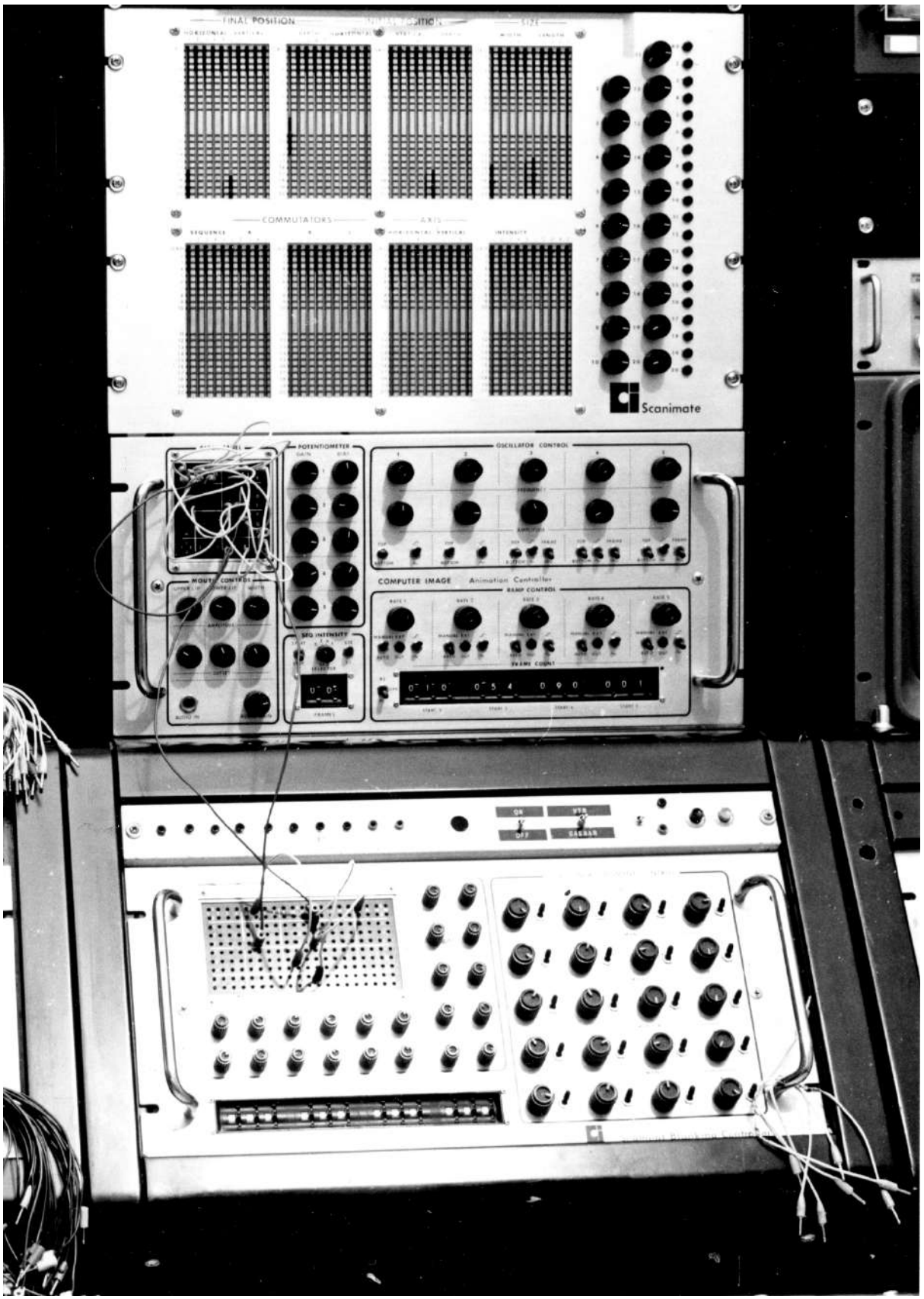
sfruttando tutti gli SMS, messaggi, email, post sui social e altro ancora che il suo amico aveva lasciato di sé, Kuyda ambiva a creare un programma che fosse almeno parzialmente in grado di riprodurre le caratteristiche caratteriali del suo amico (proprio come visto nell'episodio Be Right Back della serie TV Black Mirror).

Il tentativo non diede soddisfazione a Kuyda, che utilizzò però l'esperienza fatta per lanciare, nel 2017, Replika: un chatbot da compagnia, con cui comunicare per via testuale e che, imparando a conoscerci nel corso del tempo, diventa in grado di relazionarsi con noi in maniera sempre più convincente.

Dopo un inizio in sordina, Replika ha conquistato il successo durante la pandemia, quando milioni di persone hanno sofferto di solitudine a causa del lockdown, trovando sollievo anche nelle interazioni con questo chatbot.

Oggi Replika può contare su due milioni di utenti attivi e oltre dieci milioni di utenti registrati. Nelle sue ultime versioni, Replika permette di scegliere il sesso, si è arricchita di nuove funzionalità (si possono scrivere canzoni assieme e si può conversare anche a voce, come se fosse una telefonata), si può scegliere che tipo di relazione avere con lei/lui (amici, partner romantico, mentore oppure casuale, a seconda di come andranno le cose) e si arricchisce col tempo di nuovi tratti caratteriali.

L'intimità che alcune persone sono in grado di sviluppare con i bot ha portato a un passo successivo: secondo i dati forniti proprio da Replika, il 14 percento degli utenti si relaziona con l'intelligenza artificiale con modalità romantiche. Un dato che apre nuovi scenari sul ruolo che, in futuro, questi chatbot potranno avere nella società.


FINAL POSITION
COMMUTATORS
AXIS
Scanimate
POTENTIOMETER
OSCILLATOR CONTROL
COMPUTER IMAGE
Animation Controller
RAMP CONTROL
FRAME COUNT
SEQ INTENSITY
0 1 0 0 5 4 0 9 0 0 0 1

Geopolitica

La “guerra fredda” dell’AI

La scacchiera geopolitica in cui collocare questa tecnologia. E perché interessa agli Stati. Sullo sfondo la rivalità Usa-Cina.

di **Antonio Dini**

Chi sta vincendo la guerra delle AI? E poi, c’è veramente una guerra delle AI? L’intelligenza artificiale come fattore di potenza e di ricchezza delle nazioni è difficile da decodificare a causa della complessità, segretezza e, paradossalmente, anche dell’eccessiva esposizione mediatica di questo settore.

Da un lato, infatti, c’è la ricerca pura, che è nata negli anni Cinquanta ed è arrivata a maturazione attorno al 2010 quando si sono incrociati lo sviluppo delle reti neurali profonde, la crescita della quantità di dati e di computing. Un ambito che attrae una buona parte dell’attenzione e del dibattito pubblico anche se è considerato già maturo: le principali innovazioni sono già state fatte, adesso siamo in una fase di evoluzione e messa a terra della tecnologia.

Dall’altro, c’è appunto, la corsa alla commercializzazione, iniziata con il lancio pubblico di ChatGPT da parte di OpenAI nel novembre 2022. È il piano che genera il maggior rumore mediatico e che viene alimentato dagli investimenti e dalla nascita di moltissime startup e da piccole e grandi iniziative di settore, senza contare i risvolti politici e regolamentari.

Nel mezzo, esiste un terzo piano di lettura delle AI che riguarda le politiche internazionali degli Stati e delle grandi corporation. È quella che nel 2018 la rivista americana *Wired* ha definito “la guerra fredda delle AI” e che Henry Paulson, ex segretario al Tesoro dell’amministrazione George W. Bush, poco dopo ha chiamato “Economic Iron Curtain”, la “Cortina di ferro economica”. Ed è quella in realtà meno conosciuta e soggetta al maggior numero di fraintendimenti, distorsioni e pregiudizi, secondo le ricerche come quella condotta da Joanna J. Bryson ed Helena Malikova per *Global Perspectives* della University of California Press.

La percezione del pubblico, tra la novità commerciale e la fantascienza

Secondo diversi studi, il piano geopolitico è poco percepito dall'opinione pubblica occidentale perché la narrazione corrente è basata su un "hype" generato dalla meraviglia dei primi prodotti presentati, soprattutto ChatGPT, e dalla conseguente corsa alla realizzazione di prodotti e funzioni commerciali ricollegabili all'intelligenza artificiale. In questo campo gli investimenti delle aziende stanno "drogando" il settore dell'informazione, che a sua volta ha maggiore interesse a rincorrere le storie più sensazionalistiche per attrarre click e interesse da parte del pubblico, soprattutto in una fase di crisi acuta per i media internazionali.

Secondo una ricerca condotta dal Columbia Journalism Review, infatti, la copertura mediatica della AI ha già superato quella della maggior parte delle tecnologie e fenomeni tech presentati negli ultimi anni (realtà virtuale, metaverso, deepfake) raggiungendo sostanzialmente le vette della copertura mediatica di Bitcoin e "della promessa delle criptovalute di cambiare il sistema bancario e commerciale così come lo conosciamo".

In particolare, scrive il rapporto, "a soli sei mesi dal lancio, ChatGPT sta già ricevendo un'attenzione simile a quella riservata alle criptovalute nel 2021, quando i prezzi dei bitcoin raggiunsero il picco, oltre un decennio dopo la sua diffusione al pubblico nel 2009".

La qualità della copertura mediatica è però fuorviante perché in generale sottolinea gli aspetti più "fantascientifici" (come l'intervista a Geoffrey Hinton, pioniere della AI, che ha dichiarato che "la AI può sterminare l'umanità e potrebbe essere impossibile fermarla") e apre al rischio militare e geopolitico oltre a quello di non informare o sottovalutare l'impatto delle AI da un punto di vista etico e del futuro del mondo del lavoro.

L'interesse degli Stati per l'AI

Ci sono due motivi principali per cui l'intelligenza artificiale viene ritenuta uno strumento strategico dai Governi e dai colossi della tecnologia: da un lato il vantaggio industriale (e la conseguente ricchezza economica) che può generare.

Dall'altro la possibilità di essere utilizzata per scopi militari o di controllo interno.

L'interesse da parte dei Governi per lo sviluppo e l'utilizzo delle tecnologie di intelligenza artificiale è legato innanzitutto all'impatto che gli sviluppi futuri delle AI possono avere sulla ricchezza delle nazioni, ovvero il grande gioco competitivo su scala globale.

Secondo la società di market intelligence Idc, il valore complessivo del mercato delle AI crescerà del 19 per cento all'anno e nei prossimi tre anni arriverà a sfiorare il valore di 1.000 miliardi di dollari. Per la società di consulenza Accenture, la ricchezza verrà generata dall'incremento del 40 per cento della produttività del lavoro svolto usando la AI, che a sua volta porterà vantaggi a cascata in tutte le filiere produttive. La società di servizi professionali e consulenza PwC ha quantificato questo effetto in una ricerca, sostenendo che la ricchezza generata utilizzando le AI supererà i 15mila miliardi di dollari.

Per questo grandi banche americane come Bank of America e Goldman Sachs ritengono che gli investimenti nel settore cresceranno in maniera rapidissima, superando i quasi 100 miliardi investiti nel 2021 (prima cioè dell'annuncio di ChatGPT), che a loro volta erano già cresciuti di cinque volte rispetto al quinquennio precedente.

Queste valutazioni derivano dal fatto che, a partire dal 2010, con i primi risultati positivi offerti dagli algoritmi di deep learning, le AI sono diventate un mercato in costante crescita, che produce valore aggiunto in vari settori: dalla sanità al settore finanziario e assicurativo alla ricerca biotech. Quest'ultimo è il comparto economico con il maggior valore aggiunto dell'economia occidentale, assieme al settore delle tecnologie digitali e al settore della produzione, gestione e distribuzione dell'energia. Che, non a caso, utilizzano anch'essi da tempo algoritmi di intelligenza artificiale per aumentare la redditività delle loro attività.

Uno scontro economico prima di tutto

Con questa premessa, basandosi cioè sulla ricchezza che viene generata, lo scontro economico tra nazioni per il controllo e la gestione dell'intelligenza artificiale è già spiegabile e, dal loro punto di vista, pienamente giustificato anche da fattori geopolitici. Sempre secondo PwC, ad esempio, la ricchezza generata dalle AI verrà distribuita in maniera piuttosto diseguale nel mondo: gli

Usa e la Cina intercetteranno poco meno di due terzi degli oltre 15mila miliardi di dollari previsti nel 2030, mentre al resto del pianeta arriverà circa un terzo del totale. Anche qui, con una ulteriore divisione tra Europa e le altre economie sviluppate da un lato e i Paesi in via di sviluppo dall'altro, ai quali arriveranno sostanzialmente le briciole.

Rivalità Stati Uniti-Cina

Questo scontro economico è alla base della narrativa che fa riferimento alla "Guerra Fredda delle AI" tra Usa e Cina e si inserisce in un più ampio scontro legato all'accesso delle tecnologie digitali da parte della Cina (che, secondo alcuni studi, ha una limitata capacità di ricerca, sviluppo e produzione, anche se sta rapidamente colmando il gap). Gli Stati Uniti hanno bloccato l'esportazione di una serie di tecnologie e vietato la collaborazione alle aziende americane o straniere che vogliano portare avanti attività commerciali negli e con gli Stati Uniti nel settore delle tecnologie informatiche e per le telecomunicazioni a partire dall'amministrazione Obama.

Il ban alle tecnologie cinesi si è esteso durante l'amministrazione Trump (arrivando a comprendere ad esempio anche Asml, azienda olandese principale produttore al mondo di impianti litografici per la produzione dei "negativi" da cui vengono realizzati gli strati di microscopici transistor dei microchip) ed è stato sostanzialmente mantenuto dall'attuale presidente americano Biden (che nell'agosto 2023 ha firmato un nuovo ordine esecutivo per limitare quegli investimenti statunitensi in Cina che finanziano lo sviluppo del calcolo quantistico, dell'intelligenza artificiale e di chip avanzati).

Giustificato con ragioni di sicurezza nazionale, il divieto di esportazione ha avuto non solo l'obiettivo di bloccare i presunti rischi di spionaggio da parte dei fornitori di tecnologia legati a Pechino (come Huawei) negli apparati e infrastrutture telematiche occidentali, ma anche di isolare la ricerca e sviluppo di nuove tecnologie da parte delle aziende cinesi.

Questo perché, secondo Eric Schmidt, ex numero uno di Google e oggi lobbista e consulente dell'amministrazione americana sui temi di intelligenza artificiale, e Graham T. Allison, professore di scienza politica di Harvard, già nel 2020 la Cina avrebbe avuto una "capacità delle AI superiore a quella degli Stati Uniti" in molte aree critiche. E gli Stati Uniti non sarebbero in grado di "difendersi dalla Cina" sui mercati della tecnologia. "La maggior parte degli americani – scrivevano

Schmidt e Allison – ritiene che il vantaggio del proprio Paese nelle tecnologie avanzate sia inattaccabile. E molti nella comunità della sicurezza nazionale statunitense insistono sul fatto che la Cina non potrà mai essere più di un "concorrente quasi alla pari" nell'AI. In realtà, la Cina è già un concorrente alla pari a tutti gli effetti, sia in termini di applicazioni commerciali che di sicurezza nazionale dell'AI. La Cina non sta solo cercando di padroneggiare l'AI, ma la sta padroneggiando".

Il ruolo di Eric Schmidt

La prima valutazione di Schmidt, che ha presieduto anche la National Security Commission americana sull'Intelligenza artificiale, risale al 2019 con un primo rapporto nel quale vengono indicati gli estremi del problema e la necessità di coinvolgere ricercatori universitari e i centri di ricerca aziendali per mappare lo stato dell'arte dello sviluppo dell'AI. Un settore, nell'economia americana, in cui il peso della componente privata è il triplo per investimenti e numero di addetti rispetto a quello della ricerca universitaria (che 20 anni fa invece dominava il settore). La commissione presieduta da Schmidt ha proposto di raddoppiare l'investimento americano nel settore privato entro il 2026, portandolo a 32 miliardi di dollari.

Schmidt è una figura chiave di questo particolare snodo: ha fatto da consigliere a due presidenti americani (Obama e Biden) oltre ad aver aiutato Hillary Clinton a creare la sua piattaforma politica per la tecnologia. Soprattutto, Schmidt è stato tra i primi sostenitori dell'utilizzo delle AI da parte dei militari per costruire un vantaggio in termini di armamento.

L'ex manager di Sun Microsystems e di Google, infatti, dopo aver lasciato la guida dell'azienda di Mountain View nel 2017 (dove aveva supervisionato i progetti legati all'intelligenza artificiale, alle auto a guida autonoma e ai computer quantistici), era stato invitato dall'allora segretario alla difesa di Barack Obama, Ashton Carter a presiedere un'altra commissione: il Defense Innovation Board.

A partire da questo momento Schmidt è stato il principale artefice dell'idea di trasferimento tecnologico dalla Silicon Valley verso la Difesa americana. Il manager, che ha recentemente scritto un libro con Henry Kissinger e lo scienziato informatico Daniel Huttenlocher intitolato The Age of AI: And Our Human Future, ritiene che le AI non solo cambieranno la nostra relazione con la

conoscenza, la forma della società e quella della politica, ma anche le armi e gli eserciti del pianeta.

Il nodo Taiwan e microchip

Per sviluppare le AI c'è bisogno di tre cose: algoritmi innovativi, una grande massa di dati per addestrare i modelli (punto sul quale la normativa Usa è estremamente permissiva) e una potenza di calcolo enorme per elaborarli. Quindi: servono chip moderni e potenti, in grande quantità. Qui si inserisce un rischio di tipo militare e geopolitico, perché il principale centro di produzione dei microchip nel mondo, attualmente, è a Taiwan. Schmidt torna molto spesso sul "nodo di Taiwan", cioè il problema della produzione dei microchip più avanzati.

Schmidt ha detto più volte che "la microelettronica è alla base di tutta l'intelligenza artificiale e gli Stati Uniti non producono più i chip più sofisticati del mondo. Dato che la stragrande maggioranza dei chip all'avanguardia viene prodotta in un unico stabilimento separato da appena 110 miglia d'acqua dal nostro principale concorrente strategico, dobbiamo rivalutare il significato di resilienza e sicurezza della catena di approvvigionamento."

Il concorrente è, ovviamente, la Cina, mentre lo stabilimento che si trova a sole 110 miglia dal territorio "nemico" è quello di un'azienda chiamata TSMC, Taiwan Semiconductor Manufacturing Company, che si trova nell'isola che la Cina considera una estensione del proprio territorio.

Se, da un lato, gli Stati Uniti hanno avviato durante la presidenza Trump una netta strategia di "re-internalizzazione", per riportare in patria la produzione ad alto valore aggiunto di semiconduttori e altre componenti chiave nella produzione di computer (la stessa TSMC ha avviato in Arizona la discussa ristrutturazione di un suo impianto per la produzione di microchip dal costo di 40 miliardi di dollari), dall'altro la strategia nel settore delle AI della Cina e del blocco europeo (e dei pochi altri attori rilevanti al mondo, come Corea del Sud, Giappone, Iran e Vietnam) è più difficile da decodificare. La Cina per mancanza di informazioni, il resto del mondo per difficoltà a elaborare una strategia coerente che riesca a competere con quella degli Stati Uniti.

L'Europa ha più volte tentato, nel corso degli ultimi venti anni, di avviare progetti di ricerca e sviluppo nel settore dell'intelligenza artificiale, ma senza raggiungere

i risultati e la scala degli Stati Uniti. Questo, anche a causa della mancanza di uno dei tre fattori necessari all'addestramento e allo sviluppo di sistemi ad alte prestazioni: microchip sufficientemente potenti.

I più grandi produttori al mondo sono americani o operano in sinergia con gli americani: oltre a Tsmc e a Intel, ha un ruolo chiave la statunitense Nvidia, produttrice di schede video che si è trovata in una posizione di naturale vantaggio nello sviluppo sia di processori nel settore della produzione di criptovalute che in quello dell'addestramento di modelli di AI. Tuttavia, i principali produttori di tecnologia legata alle AI negli Stati Uniti sfruttano processori creati ad hoc.

L'innovazione nella AI, dalla Silicon Valley all'Europa

Alcuni esempi. Google ha realizzato le Tensor processing unit (TPU), dei processori progettati su misura per il suo cloud che consentono di ottimizzare l'addestramento delle AI. La stessa cosa ha fatto Amazon per i suoi data center di AWS (il servizio cloud dell'azienda). Più di recente anche Facebook, per tenere il passo nella corsa alle AI con Google e Amazon, ha iniziato a progettare dei chip su misura (e messo in open source gli algoritmi dei suoi sistemi di intelligenza artificiale). Dal 2019 Microsoft ha sviluppato internamente il chip Athena mentre OpenAI si appoggia a processori di altri. La stessa Apple, che ha un ruolo significativo nella realizzazione di dispositivi tecnologici, ha realizzato dei SoC (System on a Chip) su misura: Apple Silicon, con una componente specializzata per l'esecuzione degli algoritmi di intelligenza artificiale chiamata Neural Engine che copre circa il 50 per cento della superficie del chip stesso.

In Europa la strategia per le AI procede invece lungo due direttrici parallele: l'Unione Europea sta promuovendo una serie di investimenti, dal Chips Act alla ricerca software soprattutto nel settore accademico, da un lato, e sta elaborando normative per la raccolta, il trattamento e l'elaborazione delle informazioni personali dall'altro. Inoltre, l'Unione Europea sta avanzando anche una serie di normative pensate per disciplinare direttamente l'operatività del settore delle intelligenze artificiali, che sono legate però alla strategia di sviluppo degli scambi commerciali e al trasferimento tecnologico con gli Stati Uniti.

Dal punto di vista della ricerca, (come ha spiegato, durante un incontro a Milano cui abbiamo assistito, il presidente dell'associazione italiana dell'intelligenza artificiale, Gianluigi Greco) in Europa non possiamo competere con la potenza di calcolo degli Stati Uniti: "Attualmente i migliori sistemi hanno 17.000 GPU [componenti hardware chiave per l'elaborazione grafica e la potenza computazionale degli algoritmi di AI, NdR], il più performante ne ha 43mila. Non abbiamo questa potenza di calcolo o le risorse per costruirla. Dobbiamo affrontare i problemi nei settori in cui possiamo fare la differenza, cioè gli algoritmi, e trovare soluzioni più creative e innovative in questo ambito".

I finanziamenti erogati dall'Unione Europea nel settore sono relativamente consistenti. Come ha spiegato (durante lo stesso evento) Vittorio Calaprice, analista della rappresentanza italiana per la Commissione Europea, gli investimenti vengono erogati tramite differenti piani, da Horizon Europe, Next Generation EU, fino al Pnrr. L'Unione Europea investe un miliardo all'anno e attrae altri 20 miliardi di euro di investimenti privati. Gli Stati Uniti hanno un investimento di circa 55 miliardi di dollari. La Cina si stima che investa circa 20 miliardi di dollari. In Europa, il Paese che investe di più è il Regno Unito (ma è fuori dall'Unione Europea). Immediatamente dietro ci sono Germania e Francia, che però seguono un approccio integrato con i progetti europei. L'investimento in Italia si aggira attorno ai 500 milioni.

Secondo le politiche dell'Unione Europea, oltre all'aumento della produttività, l'AI dovrebbe evolversi in sinergia con lo sviluppo sostenibile e contribuire soprattutto alla trasformazione digitale per renderla, nelle parole della Commissione, una "Twin Transition": una doppia transizione in cui oltre a una dimensione di diritti e regolamentazioni per gli individui ci devono essere anche indirizzi per quanto riguarda la sostenibilità. Una AI "non solo antropocentrica ma anche planetocentrica", nelle parole dei dirigenti della Commissione.

La Commissione Europea spinge anche per l'adozione di regolamentazioni specifiche che sono simmetriche alla richiesta di una "moratoria" sulle AI proposta dai alcuni big del tech negli Usa, che servirebbe però sostanzialmente a congelare lo status quo, cioè il vantaggio della coppia OpenAI-Microsoft rispetto a Google, Facebook e Amazon. Ma soprattutto a scapito della concorrenza delle startup e di altri potenziali "disruptor" del mercato, che ha visto per un decennio investimenti enormi da parte dei titani del settore e che adesso devono essere monetizzati con i primi prodotti e servizi.

Il modello asiatico

Fuori dall'Europa, i Paesi asiatici investono sistematicamente nelle AI. In Corea del Sud, il gruppo Naver (che gestisce il primo motore di ricerca del Paese con una quota di mercato del 61 per cento rispetto al 29 per cento di Google) si è alleato con Samsung (una delle poche aziende al mondo dotate di impianti avanzati per la produzione di microchip) e il Governo coreano per la realizzazione di un modello di AI sofisticato (chiamato HyperClova AI) e un ecosistema di aziende e startup collegate.

Naver sta sviluppando delle AI localizzate per i Paesi del mondo arabo, nonché per i grandi Paesi non anglofoni, come Spagna e Messico, dove i Governi desiderano avere i propri sistemi di AI personalizzati in base ai loro contesti politici e culturali e "al sicuro" anche da possibili interferenze o atti di spionaggio da parte di aziende legate al Governo americano. "Sarà un business enorme, – ha detto al *Financial Times* Sung Nako, un dirigente di Naver responsabile dello sviluppo dell'AI – poiché la tecnologia AI sovrana sta diventando sempre più importante per la protezione dei dati".

Un ruolo più defilato ma non meno importante in Asia è quello del Giappone, che nella AI vede uno strumento di aumento della redditività del lavoro e della produzione, ma anche un modo per automatizzare il già esteso parco di robot industriali (e in futuro per le cure domestiche). La prima strategia (che risale al 2018) è basata sul piano Society 5.0, che mira a portare un impatto positivo delle tecnologie di AI, assieme ad altre, nel tessuto sociale del Paese.

Il convitato di pietra della corsa alle AI però è un altro: la Cina.

Il Paese viene percepito come un rischio in Occidente soprattutto da quando, nel luglio del 2017, ha annunciato di voler raggiungere la leadership planetaria nel settore delle AI entro il 2030. La guida di Xi Jinping sta contribuendo al raggiungimento di una serie di primati scientifici e tecnologici, tra cui la messa in orbita del primo astronauta "civile" con un vettore cinese e la commercializzazione di un aereo di linea capace di fare la concorrenza nei trasporti a corto e medio raggio a Boeing e Airbus. La Cina è anche da tempo il primo produttore e acquirente al mondo di auto elettriche e il piano Made in China 2025 (che risale al 2015) punta alla capacità di spingere ancora di più l'evoluzione industriale del Paese.

Per quanto riguarda l'intelligenza artificiale, il ruolo maggiormente conosciuto è quello del settore privato, guidato da pochi, grandi produttori: Tencent, Baidu e

il gruppo Alibaba. Ma, come osserva l'International Institute for Strategic Studies, "dopo la Silicon Valley il secondo centro al mondo di intelligenza artificiale è l'Accademia delle scienze cinese".

Non sono noti gli usi interni sui sistemi d'arma letale autonomi sviluppati dall'esercito popolare cinese, ma sono più conosciuti quelli nell'ambito della sorveglianza della popolazione. Dai sistemi di visione e riconoscimento dei volti per individuare nella folla le persone di etnia uigura, oltre che i dissidenti politici e altri individui schedati, ai sistemi automatici di contenimento di milioni di persone durante la pandemia da COVID-19. Nella provincia di Hubei, abitata da 60 milioni di individui (leggermente superiore alla popolazione italiana), la Cina ha mantenuto un enorme cordone sanitario completamente impermeabile al passaggio delle persone grazie a una serie di algoritmi di intelligenza artificiale "per tracciare gli spostamenti dei residenti e aumentare le capacità di analisi mentre venivano costruite nuove strutture sanitarie".

Chi sono i Paesi cyber-maturi

L'International Institute for Strategic Studies ha rivisto la sua classificazione dei tre livelli che misurano la cyber-maturità dei Paesi nel mondo, che ha visto a lungo il predominio del sistema statunitense (l'unico al livello uno), seguito dal "gruppone" composto da Regno Unito, Australia, Canada, Francia, Israele, Russia e Cina nel secondo livello, e poi dal terzo livello con India, Indonesia, Giappone, Malesia, Corea del Nord, Iran e Vietnam. Adesso la Cina, secondo varie analisi, avrebbe raggiunto una cyber-maturità paragonabile a quella degli Stati Uniti soprattutto grazie alla ricerca nell'ambito dell'AI.

Una maturità che per gli americani si configura come un rischio: nel 2018 l'allora capo di Stato maggiore americano, generale Joseph Dunford, aveva dichiarato che "chiunque abbia un vantaggio competitivo nell'intelligenza artificiale e sia in grado di mettere in campo sistemi basati sull'intelligenza artificiale, potrebbe benissimo avere un vantaggio competitivo complessivo".

Al di là della propaganda di Pechino e della narrativa della "Guerra fredda della AI" occidentale, non ci sono informazioni indipendenti e autonome che indichino quali siano i risultati raggiunti dalla ricerca militare e governativa cinese e quali siano le applicazioni concrete.

Esistono poche ricerche in grado di fornire un quadro completo sui sistemi d'arma descritti o pubblicizzati come "autonomi" o

"intelligentizzati" (intelligentized, concetto specifico con cui la Cina applica la AI al dominio militare) che si basano sulla ricerca scientifica e militare e nello sviluppo di sistemi senza pilota e di tipo missilistico, ma anche di superficie (come i cani-robot autonomi da combattimento, ad esempio).

La guerra in Ucraina e la corsa agli armamenti AI

Anche nel conflitto militare in corso tra Ucraina e Russia sarebbe sempre possibile una escalation delle AI. Un timore che viene alimentato da varie testate internazionali, anche sulla base del fatto che già in passato le AI sarebbero state utilizzate durante dei conflitti, come segnalato da un rapporto del Consiglio di Sicurezza dell'Onu riguardo alla seconda guerra civile in Libia. Si trattava, in quel caso, di sistemi d'arma letali e completamente autonomi del tipo STM Kargu-2, capaci di individuare e attaccare fino a distruggere bersagli di natura non predefinita, operando senza alcun tipo di connessione con gli operatori remoti. Ancora lontani dall'utilizzo di robot con licenza di uccidere (o interi sciami di robot autonomi d'assalto dotati di forza letale sia via terra che in volo), il cui effetto sorpresa sarebbe riservato probabilmente a conflitti più vicini al cuore degli interessi geopolitici americani ed europei.

In conclusione, l'intelligenza artificiale come fattore di potenza militare oltre che come fattore di ricchezza economica è al centro degli obiettivi delle principali potenze planetarie. I tentativi di comprensione e regolamentazione del fenomeno sono però ancora frammentati e divisi in ambiti diversi: da quello militare a quello economico dalle regolamentazioni commerciali a quelle per i diritti e la privacy.


CONTROL DATA 3300

Stati

USA, UE e Cina: il trilogo sull'AI

Come si stanno muovendo i tre colossi politici tra regolamentazione e investimenti.

di **Federica Meta**

L'intersezione tra geopolitica e tecnologia, in questo momento, è un complesso sistema di nazioni, attori e politiche che si intrecciano tra di loro, influenzandosi a vicenda. Per poter comprendere gli approcci e le prospettive dei principali soggetti politici presenti nel mondo, è necessario fare un affondo specifico sulle politiche dei singoli Stati. In particolare, quelle implementate dal "trilogo", chiamiamolo così, composto dagli Stati Uniti, l'Unione Europea e la Cina.

Le mosse della Cina

Con investimenti di oltre 13 miliardi di dollari in AI nel 2023, la Cina è sicuramente uno dei maggiori protagonisti della corsa a questa tecnologia. La Cyberspace Administration of China (Cac) ha stabilito, in un progetto di legge (finalizzato a luglio), che tutti i nuovi prodotti di intelligenza artificiale sviluppati nel Paese dovranno essere sottopostia una "valutazione di sicurezza".

"Prima di fornire servizi al pubblico che utilizzano prodotti di intelligenza artificiale generativa – dice il provvedimento - è necessario richiedere una valutazione della sicurezza tramite i Dipartimenti nazionali di regolamentazione di Internet".

Le mosse degli Stati Uniti

Mentre la Cina prevede dei limiti entro i quali sviluppare servizi basati sull'AI, gli Stati Uniti si preparano a investire 140 milioni di dollari nella creazione di sette nuovi centri di ricerca e a pubblicare linee guida tese a regolamentare questa tecnologia. Per l'investimento, la Casa Bianca attingerà ai fondi della National Science Foundation.

A febbraio, inoltre, il presidente Joe Biden ha firmato un ordine esecutivo che ordina alle agenzie federali di proteggere i cittadini dalla "discriminazione algoritmica" mentre la Federal Trade Commission, il Consumer Financial Protection Bureau, l'Equal Employment Opportunity Commission e la Civil Rights Division del Dipartimento di Giustizia hanno firmato una dichiarazione congiunta in cui si ribadisce "l'impegno collettivo a sfruttare le autorità legali esistenti per proteggere il popolo americano dai danni legati all'AI".

Al centro di questa strategia è prevista anche una forte collaborazione con i principali protagonisti industriali del settore. Il presidente Biden ha affidato a Kamala Harris il compito di sondare il terreno: in questi mesi sono in corso riunioni tra la vicepresidente e Alphabet, Anthropic, Microsoft e OpenAI per discutere dei rischi e dei limiti dell'intelligenza artificiale.

Sono tre i campi, annunciati a maggio, in cui la Casa Bianca si sta muovendo per promuovere uno sviluppo sostenuto e sostenibile dell'AI:

- nuovi investimenti su ricerca e sviluppo. Ai centri nazionali di ricerca il compito di garantire un'innovazione responsabile e potenziare la ricerca e sviluppo degli Stati Uniti sulla tecnologia, soprattutto nei settori relativi al clima, all'agricoltura, all'energia, alla sanità pubblica, all'istruzione e alla sicurezza informatica;
- valutazioni dei sistemi di AI generativa esistenti. Aziende come Microsoft, Google, Nvidia e OpenAI saranno invitate a partecipare a valutazioni pubbliche dei modelli di intelligenza artificiale che stanno sviluppando;
- stesura di una Carta dei Diritti dell'AI e di un AI Risk Management Framework.

L'AI Act europeo

In Europa, invece, inizia il conto alla rovescia per l'approvazione finale dell'AI Act. Lo scorso 11 maggio, il Parlamento Europeo ha dato il primo via libera alla proposta di Regolamento – il primo al mondo - che è stata votata alla sessione plenaria del Parlamento europeo a giugno, in vista dei negoziati inter-istituzionali con il Consiglio dell'Unione Europea.

L'obiettivo è dare il via libera entro la fine della legislatura, nella primavera del 2024, dopo la fase dei triloghi (ne abbiamo scritto qua).

Per quanto riguarda l'AI generativa, Strasburgo ha deciso che questi sistemi dovranno essere progettati nel rispetto del diritto dell'Unione Europea e delle libertà fondamentali. La proposta della Commissione definiva ad "alto rischio" quei sistemi di AI applicati, ad esempio, alle reti critiche, all'occupazione, all'istruzione e alla formazione nonché ai servizi pubblici essenziali.

Gli eurodeputati hanno definito ad "alto rischio" anche i sistemi che possono provocare danni alla salute, alla sicurezza, ai diritti fondamentali e all'ambiente.

Il rischio significativo è definito come "risultato della combinazione della sua gravità, intensità, probabilità di accadimento e durata dei suoi effetti, e della capacità di colpire un individuo, una pluralità di persone o di colpire un particolare gruppo di persone", si legge nel testo approvato.

Infine sono considerati ad alto rischio anche i sistemi di raccomandazione utilizzati per influenzare gli elettori di piattaforme online di grandi dimensioni, così come sono definite dal Digital Services Act.

Sono state aumentate anche le tutele per i dati personali con controlli più stretti su come i provider di sistemi ad alto rischio possono elaborare dati sensibili: ad esempio l'orientamento sessuale o quello politico e religioso.

Invece i fornitori di modelli di base dovranno valutare e mitigare i possibili rischi (alla salute, alla sicurezza, ai diritti fondamentali, all'ambiente, alla democrazia e allo Stato di diritto) e registrare i loro modelli nella banca dati dell'UE prima della loro immissione sul mercato europeo. I sistemi di AI generativa che si basano su tali modelli, quali ChatGPT, dovranno rispettare i requisiti di trasparenza (dichiarando che il contenuto è stato generato dall'AI), aiutando anche a distinguere le cosiddette immagini deepfake da quelle reali, e fornire salvaguardie per evitare la generazione di contenuti illegali. Dovranno inoltre essere pubblicate le sintesi dettagliate dei dati protetti dal diritto d'autore utilizzati per l'addestramento, si legge nel comunicato stampa del Parlamento europeo.

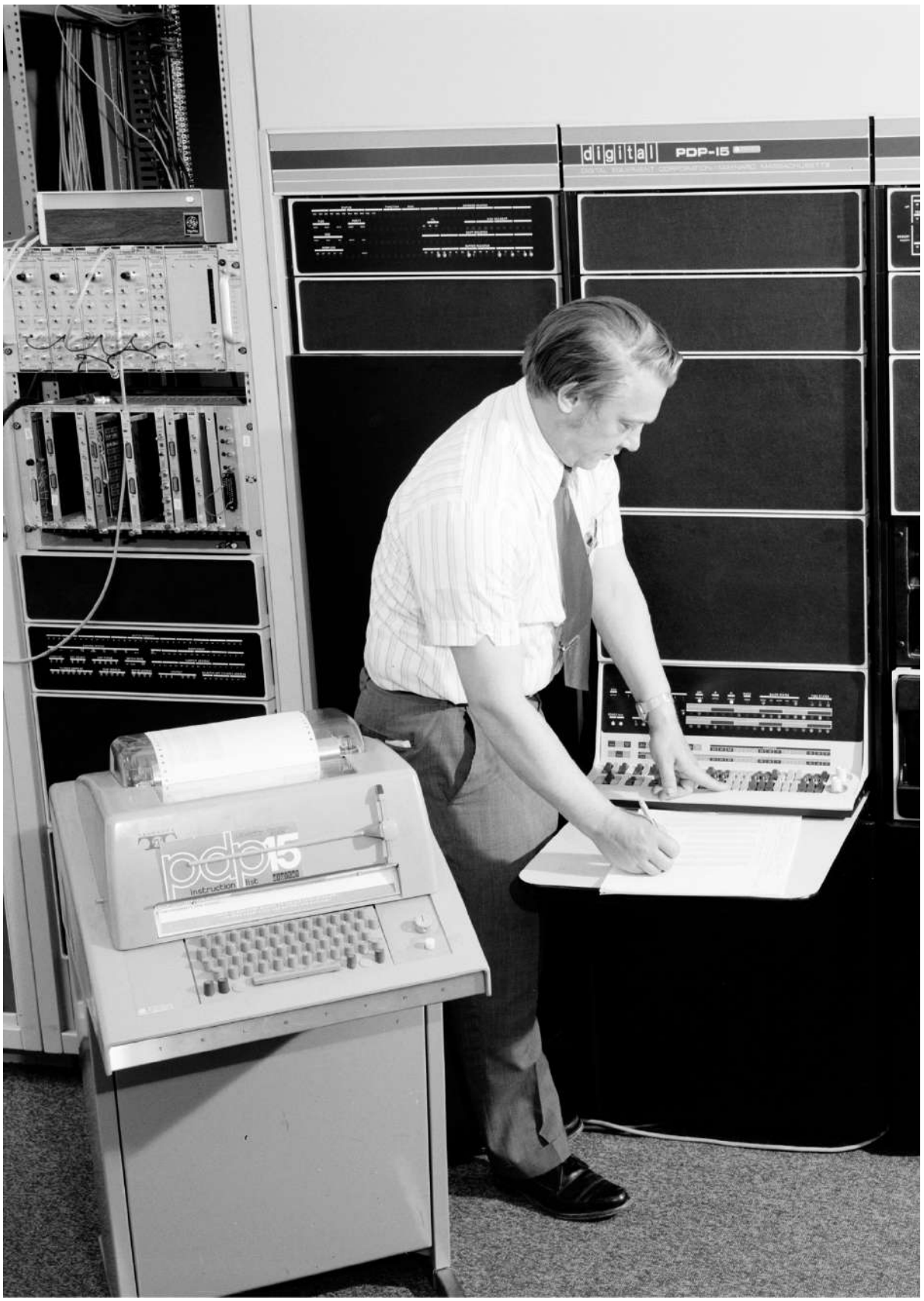
Intelligenza artificiale, gli investimenti di Horizon Europe

L'Unione Europea si sta preparando a investire 180 milioni di euro in tecnologie digitali all'avanguardia, seguendo le ultime linee guida del programma Horizon

Europe che promuovono la ricerca e lo sviluppo collaborativi in tutta l'Unione. Venti progetti dei 28 selezionati nel marzo 2023 nell'ambito di queste call (alle quali hanno partecipato università, centri di ricerca e piccole-medie imprese) riguardano l'intelligenza artificiale e la robotica.

Sei progetti selezionati, con un bilancio complessivo di 20 milioni, favoriranno la diffusione dell'AI più avanzata e della robotica nell'industria. I progetti selezionati mirano a fornire all'industria tecnologie di intelligenza artificiale e robotica più autonome, facili da usare e affidabili. Oltre a promuovere la leadership industriale, i progetti di ricerca compiono progressi significativi verso la trasparenza, la responsabilità, la sicurezza e la sicurezza dell'AI, in linea con l'approccio "human-centric" promosso dall'Unione Europea.

Sei progetti per un bilancio di oltre 50 milioni andranno a finanziare la ricerca di base nel settore mentre altre otto iniziative - il bilancio è pari a 60 milioni - mirano a sviluppare tecnologie che estendano significativamente la capacità fisica dei robot e migliorino le prestazioni energetiche.


digital PDP-15
pdp15
instruction

Mercato

Cosa stanno facendo grandi aziende e startup

Ovvero dove stanno e vanno i soldi, dalle Big Tech alle startup.

di **Federica Meta**

Nonostante lo scenario globale di involuzione a livello macroeconomico, per il comparto dell'AI si snocciolano numeri record. Nel 2022 la società di consulenza Markets&Markets prevedeva che il mercato dell'intelligenza artificiale avrebbe raggiunto un valore globale di 309,6 miliardi di dollari entro il 2026, con un CAGR (tasso annuo di crescita composto) di circa il 40 per cento da qui al 2026.

Inoltre uno studio di Goldman Sachs firmato da Joseph Briggs e Devesh Kodnani stima che l'aumento della diffusione degli strumenti basati su elaborazione del linguaggio naturale potrebbe portare a un incremento del PIL globale del 7 per cento e aumentare la crescita della produttività di 1,5 punti percentuali su un periodo di 10 anni.

Oggi viviamo dunque una vera "primavera" di interesse per le soluzioni di AI: non sono più considerate nicchie del mondo dell'accademia o dei dipartimenti di ricerca e sviluppo delle aziende, ma volano di business e competitività. Ed è per questo che non solo le imprese della tecnologia ma anche gli Stati si stanno muovendo in questa direzione: le prime per implementare questi strumenti nei loro business, i secondi per definire un contesto regolatorio abilitante in linea con i propri principi e valori. Un report della società di market intelligence IDC stima che nel 2023 il fatturato globale per l'intelligenza artificiale supererà i 500 miliardi di dollari.

I big in campo

Al di là della correttezza di numeri e stime, il mercato dell'AI è in forte espansione e a livello globale sono tantissime le aziende che stanno diventando protagoniste dell'industria grazie a questa trasformazione.

A rivoluzionare il mercato è stato sicuramente ChatGPT, uno strumento di elaborazione del linguaggio naturale (o NLP, Natural Language Processing, un campo di studio tra intelligenza artificiale, informatica e linguistica, che si concentra sull'interazione tra computer e linguaggio umano) sviluppato e reso pubblico da OpenAI alla fine del 2022.

La storia di OpenAI inizia nel 2015, quando un gruppo di informatici e imprenditori, tra cui Sam Altman, Peter Thiel ed Elon Musk, annunciano a San Francisco la nascita di un'organizzazione non profit di ricerca sull'intelligenza artificiale per scoprirne il potenziale e i benefici per la società.

L'obiettivo iniziale, in parte condizionato dai cosiddetti rischi esistenziali derivanti dall'AI, è di "collaborare liberamente" con altre istituzioni e ricercatori rendendo i suoi brevetti e le sue ricerche aperte al pubblico.

Quel gruppo di ricerca oggi è una società a tutti gli effetti, con un consiglio di amministrazione e investitori di peso tra cui Microsoft, che l'ha prima finanziata con un miliardo di dollari nel 2019 e poi, per circa 10 miliardi.

Ma il campo da gioco è popolato anche da nomi già noti nell'industria tecnologica. Tra i principali attori mondiali, abbiamo appena visto che Microsoft si è assicurata un vantaggio nella corsa all'intelligenza artificiale generativa grazie al suo investimento in OpenAI. L'azienda ha integrato il chatbot in Bing, migliorando la competitività del suo motore di ricerca e, soprattutto, posizionandosi in prima linea in questo processo di trasformazione. La tecnologia di OpenAI è inoltre stata utilizzata dal colosso tech per lo sviluppo di Copilot, un assistente AI per la piattaforma di Github pensato per affiancare i programmatori nel loro lavoro di scrittura di codice. A marzo 2023 è stato presentato anche Microsoft 365 Copilot, un assistente per la produttività personale che da giugno è integrato in Word, Excel, PowerPoint, Outlook, Teams così come negli altri prodotti software.

Ovviamente, anche Google si sta prendendo il suo spazio. Nel febbraio 2023, l'azienda ha annunciato Bard, un chatbot pensato per generare risposte dettagliate a domande semplici. Il suo funzionamento si basava inizialmente su LaMDA - il Language Model for Dialogue Applications, divenuto noto al grande pubblico nel 2022 dopo essere stato definito "senziente" da parte di uno degli ingegneri di Google (poi smentito dalla stessa azienda) - e successivamente su un altro modello, PaLM 2.

Inoltre, Alphabet (la holding di Google) ha investito 300 milioni di dollari nella startup Anthropic. Fondata da un gruppo di fuoriusciti di OpenAI, Anthropic si

focalizza principalmente sugli aspetti di sicurezza relativi all'AI, tanto da definirsi come una "azienda di sicurezza e ricerca dell'intelligenza artificiale". Il suo principale obiettivo di ricerca riguarda quella che Anthropic definisce Constitutional AI, un metodo di training che cerca di incorporare nei sistemi regole e "principi" su cui valutare i propri output. Amazon, invece, scommette su Bedrock, un servizio che mette a disposizione vari modelli di AI tramite AWS (la divisione di Amazon che fornisce servizi in cloud) e che ha come target sviluppatori e aziende.

Bedrock permette di creare chatbot per utenti e clienti, riassumere testo, classificare e creare immagini, effettuare ricerche e molto altro ancora - il tutto attraverso input di testo, in modo completamente personalizzato e su misura per le proprie esigenze.

Si parla poi anche di Meta, il colosso che possiede Facebook e Instagram. All'opera da tempo nel campo dell'intelligenza artificiale, ha però restituito risultati in parte deludenti come nel caso di Galactica, il chatbot che avrebbe dovuto assistere scienziati e ricercatori nel loro lavoro di ricerca. Nonostante il sistema fosse stato addestrato con oltre 48 milioni di articoli scientifici, è stato messo offline pochi giorni dopo l'avvio della beta pubblica a fine 2022. Ricercatori e scienziati l'avevano definito pericoloso (riferendosi alla sua tendenza a inventare articoli inesistenti). Successivamente Meta ha annunciato il lancio di un team ad hoc per lo sviluppo di intelligenze artificiali che andranno a migliorare i servizi collegati ai social del gruppo.

L'azienda di Zuckerberg sta puntando soprattutto sull'AI generativa open source. Ne è un esempio ovviamente Llama 2, la nuova generazione del modello di linguaggio di grandi dimensioni open source, gratuito per la ricerca e l'uso commerciale (su cui Meta ha pure stretto un accordo commerciale con Microsoft, rendendolo disponibile nel catalogo di Azure AI, oltre che su AWS).

E poi MusicGen, un modello open source per la generazione di musica inedita da testi. L'utente inserisce la propria richiesta testuale e il software la trasforma in una melodia. Gli sviluppatori di tutto il mondo potranno studiarlo e magari utilizzarlo per realizzare strumenti di AI ancora più avanzati. Per addestrare il modello sono state usate 20.000 ore di musica preesistente. La metà di questo archivio è composta da brani offerti con licenza, caricati in alta qualità. L'altra metà è composta da tracce provenienti da Pond5 e Shutterstock.

Pond5 è uno store online espressamente dedicato ai media liberi da royalty: le sue librerie sono composte da fotografie e filmati, ma anche da musica ed effetti

sonori. Shutterstock è invece una piattaforma dedicata all'upload e al download di contenuti multimediali di tutti i tipi. A inizio agosto Meta ha rilasciato Audiocraft, un insieme di modelli per la generazione di musica che include MusicGen oltre a EnCodec e AudioGen, per coprire tutta la gamma di necessità di chi debba produrre musica, effetti sonori e compressione.

La risposta dei big cinesi

In Cina sono diverse le aziende che stanno puntando sempre più sull'intelligenza artificiale generativa.

Alibaba ha presentato Tongyi Qianwen, risposta a ChatGPT. Tongyi Qianwen è in grado di realizzare lettere di invito, pianificare itinerari di viaggio e consigliare prodotti da acquistare. L'intelligenza artificiale è integrata in DingTalk, l'app interna di messaggistica ed è stata aggiunta a Tmall Genie, l'assistente vocale di Alibaba. L'unità cloud del colosso cinese ha in programma di aprire Tongyi Qianwen a tutti gli utenti e di integrarlo in tutte le app dell'azienda nei prossimi mesi.

Baidu, il Google cinese, ha lanciato a marzo il suo ChatGPT, chiamato Ernie Bot, a sua volta basato su un LLM, modello linguistico di grandi dimensioni, chiamato Ernie. Ernie, acronimo di Enhanced Representation through Knowledge Integration, in realtà è stato messo a punto nella sua prima versione già nel 2019 ma è stato rilasciato solo pochi mesi fa. Similmente al modello linguistico di OpenAI, anche Ernie sa trovare risposte, comporre testi e generare immagini sulla base di input testuali. Il software è stato reso disponibile in versione trial a selezionati clienti, e da allora solo una ristretta cerchia di aziende e utenti ha potuto testarlo. Il Large Language Model permetterà al motore di ricerca di Baidu di funzionare in modo più evoluto. E anche Tencent sta lavorando a un proprio chatbot, che sarà integrato in WeChat.

Il caso SenseTime

SenseTime, colosso cinese specializzato in AI, ha recentemente presentato alcuni prodotti. SenseChat è un chatbot, non molto diverso da ChatGPT, che dunque simula una conversazione umana, restituendo risultati vari in base a

input testuali.Il sistema si fonda su un modello di intelligenza artificiale proprietario di nome SenseNova, sviluppato negli ultimi cinque anni.

Un altro prodotto AI di SenseTime non ha ancora un nome ufficiale, ma sarà un generatore di immagini. Ciò vuol dire che l'utente potrà effettuare richieste testuali e ricevere in cambio immagini sintetiche, realizzate in tutto e per tutto dall'intelligenza artificiale.

La carica delle startup: dalle più popolari alle più "originali"

Non solo i big player del tech stanno investendo nell'AI generativa. Le startup che sviluppano applicazioni aziendali basate su software di intelligenza artificiale stanno catturando l'attenzione dei leader della tecnologia mondiale e dei grandi investitori.

Anthropic. Fondata nel 2021 da ex dipendenti di OpenAI, Anthropic è un'azienda di ricerca che si occupa di intelligenze artificiali e che sta lavorando a un concorrente di ChatGPT su cui Google ha investito 300 milioni di dollari. Alla base di Claude – questo il nome del suo chatbot - la tecnologia Constitutional AI che istruisce il sistema con un insieme di "regole" e "principi" utilizzati per revisionare automaticamente le risposte. Per ora Claude 2 è disponibile in early-access solamente per le aziende e l'uso degli utenti è limitato.

You.Com. Fondata da due ex dipendenti di Salesforce, You.com si propone al pubblico come "il motore di ricerca che controlli tu". Il chatbot YouChat può fornire risultati personalizzati alla ricerche dell'utente, riassumere articoli presi dal web, generare codice, scrivere. You.com recentemente ha integrato altri generatori di immagini basati sull'AI quali Stable Diffusion 1.5, Stable Diffusion 2.1 e Openjourney, con l'obiettivo di diventare la classica "one stop solution" per gli early adopter di intelligenze artificiali.

StabilityAI. La startup ha creato Stable Diffusion, un modello di deep learning rilasciato con licenza open source che consente di eseguire sostanzialmente tre operazioni: generare immagini a partire da un testo scritto; generare immagini a partire da una bozza o altra immagine di partenza indicando uno o più stili; fare image inpainting e outpainting, cioè ricostruire a partire da un testo la parte mancante di una immagine o sostituire uno o più degli elementi in essa raffigurati, ed espanderla oltre i contorni. Più di recente, Stability AI ha

annunciato il suo primo modello linguistico di grandi dimensioni (LLM) open source chiamato StableLM. In realtà si tratta di una suite di modelli linguistici denominati collettivamente StableLM, che sono già disponibili su GitHub per chi vorrà implementarli. Se il modello linguistico Gpt 3.5 che dà vita a ChatGPT ha un massimo di 175 miliardi di parametri, StableLM per il momento è fermo a 3 e 7 miliardi di parametri per i modelli in fase di test. Stability AI conta di arrivare a 15 e 65 miliardi.

Jasper. La startup con sede in Texas ha sviluppato una piattaforma di Generative AI progettata per creare automaticamente blog post promozionali e testi di marketing. Lo scorso mese ha annunciato un round di raccolta capitali da 125 milioni di dollari, che ha portato la valutazione della società al di sopra del miliardo di dollari. Tra gli investitori Insight Partners, Coatue e Bessemer Venture Partners.

Murf AI. Murf AI è una startup che fornisce soluzioni Tts (text-to-speech) basate su software as a service. I prodotti consentono agli utenti di generare voci rispettando anche le inclinazioni dialettiche e quelle legate all'età. Oltre al servizio standard, sono disponibili opzioni di voice-over in cui gli utenti possono includere e sincronizzare la musica di sottofondo.

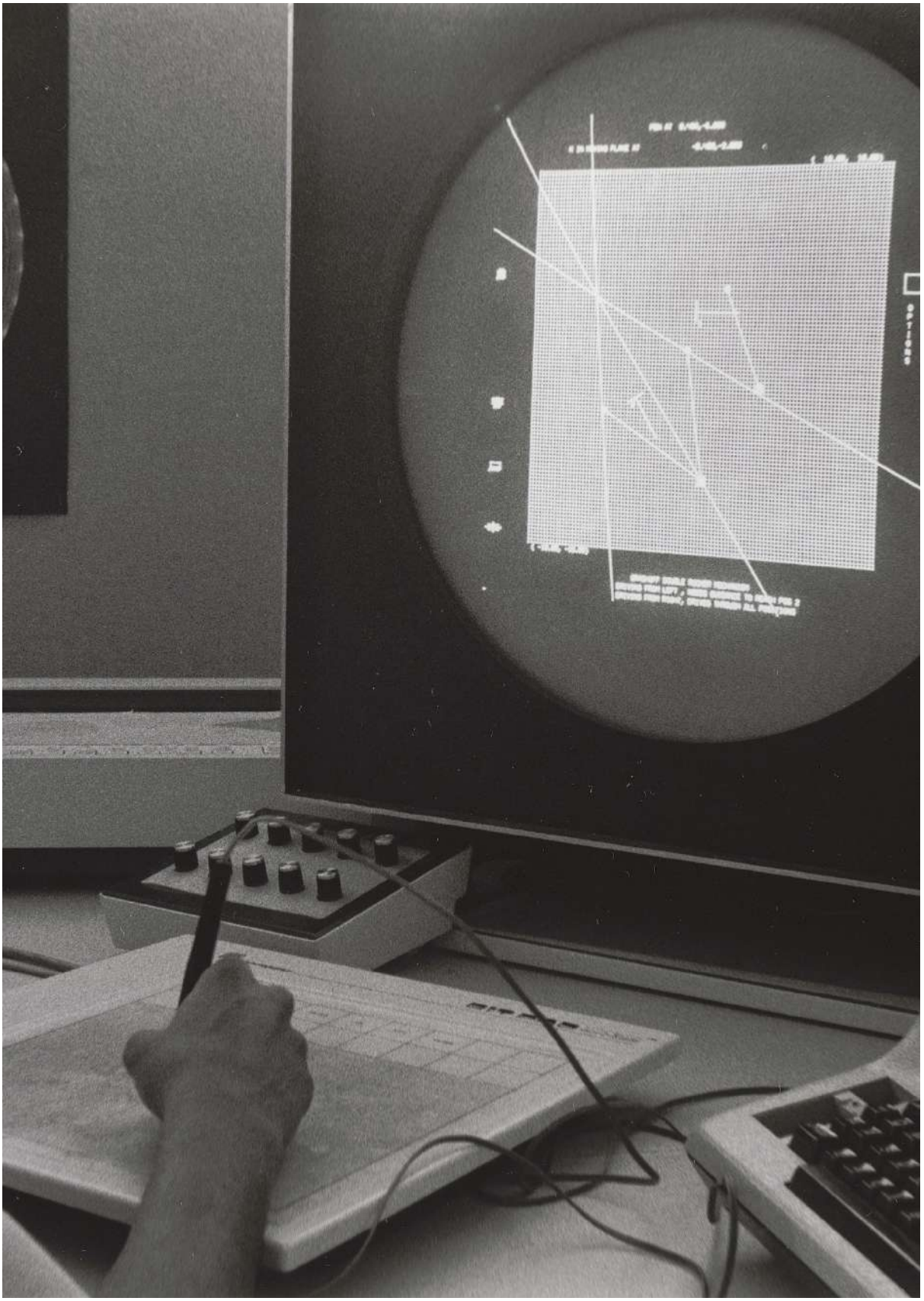
Character.AI. Character.ai è una piattaforma AI che sta dando vita al "sogno fantascientifico" di conversare, tramite computer, con un personaggio virtuale, immaginario o improntato su qualcuno realmente esistito. I personaggi di Character.ai sono agenti di dialogo alimentati da tecnologia proprietaria. L'utente può conversare con personaggi storici, immaginari oppure creati appositamente. Sono disponibili tra gli altri i personaggi di Dante, Maradona o Mussolini. Lo scopo di Character.AI è quello di creare uno strumento utile per l'immaginazione, la creatività, il brainstorming, l'apprendimento delle lingue.

Un SuperPc per l'AI generativa

Nei prossimi mesi ChatGPT e le altre piattaforme di intelligenza artificiale generativa potranno beneficiare di un assist decisivo da parte di Nvidia: il CEO Jensen Huang ha infatti presentato un supercomputer appositamente sviluppato per l'AI chiamato Dgx Gh200.

La macchina assisterà le aziende tech nella costruzione di modelli di AI generativa simili per l'appunto al software di OpenAI. Meta, Microsoft e Google Cloud saranno tra i primi clienti del supercomputer. L'annuncio giunge pochi giorni dopo che il colosso dei chip americano ha rivelato le previsioni di una rapida crescita delle vendite, alimentando un aumento del prezzo delle azioni

che l'ha portata vicina all'essere il primo titolo al mondo di semiconduttori del valore di mille miliardi di dollari.



Creatività

Creare insieme o contro l'algoritmo?

La diffusione capillare di intelligenze artificiali generative sta già avendo un impatto sugli artisti. Cambierà anche il nostro modo di essere creativi?

di **Federico Nejrotti**

ChatGPT, DALL-E, Bard, Stable Diffusion, Midjourney: le intelligenze artificiali generative, in grado di creare testi, immagini, video e suoni a partire dagli input degli utenti, hanno investito trasversalmente ogni settore produttivo, creativo e di ricerca che lavora con informazioni, dati ed elementi d'archivio. Nel corso di pochissimi mesi, i prodotti basati sui modelli linguistici di grandi dimensioni (LLM, Large Language Models) hanno raggiunto centinaia di milioni di persone: chi semplicemente per provarli, chi per studiarli, chi per farsi affiancare nel proprio lavoro quotidiano - fino al punto di farsi (quasi) sostituire.

Tecnologie capaci di penetrare così rapidamente negli immaginari e nella quotidianità del grande pubblico finiscono per produrre dibattiti che si polarizzano altrettanto velocemente, tra tecno-ottimisti convinti che la diffusione di questi algoritmi rivoluzionerà completamente la civiltà umana, e chi invece teme che una diffusione incontrollata di strumenti ancora così acerbi non potrà fare altro che amplificare dinamiche di disuguaglianza, ingiustizia e discriminazione di cui gli algoritmi implentati dalle piattaforme digitali si erano già dimostrati catalizzatori.

Ma quella in corso non è una trasformazione esclusivamente strumentale: l'introduzione di questi algoritmi nell'immaginario del grande pubblico sta mutando il nostro modo di percepire il ruolo delle intelligenze artificiali nella nostra vita. Se i settori lavorativi evolvono da sempre di pari passo con le innovazioni tecnologiche, diverso è il discorso per il mondo creativo e artistico, il quale, nonostante le numerose rivoluzioni affrontate, ha sempre mantenuto il ruolo dell'artista umano al centro.

L’impatto degli algoritmi nel mondo dell’arte e della creatività

Potremmo dibattere per ore sull’effettiva utilità di questi algoritmi nello svolgimento di compiti quotidiani, ma non vi è dubbio che i risultati più pirotecnici siano stati elaborati applicando questi algoritmi al mondo dell’arte e della creatività.

Dal lancio delle prime versioni di DALL-E e Stable Diffusion, capaci di produrre immagini quantomeno verosimili a partire da prompt testuali, fino ai risultati fotorealistici generati dall’ultima versione di Midjourney (e che preoccupano per la loro capacità di produrre contenuti facilmente scambiabili per fotografie reali), l’impatto che i Large Language Model stanno avendo su una disciplina profondamente umana come la creatività sta lasciando artisti, creativi e appassionati incuriositi, disorientati e spesso preoccupati.

Per capire il perché di questa preoccupazione, è necessario fare il punto sulle potenzialità e i limiti di questi strumenti. Per esempio, le intelligenze artificiali generative permettono di accelerare i processi creativi di brainstorming, traducendo in tempi rapidi un panorama immaginativo umano in un contenuto visivo o testuale. O ancora, se non si ha la pretesa di produrre risultati accurati, le intelligenze artificiali generative testuali permettono di raccogliere rapidamente informazioni altamente specifiche, da cui partire per un eventuale lavoro di ricerca.

Il motivo per cui tantissimi stanno guardando a questa tecnologia con preoccupazione è lo stesso che si cela dietro alle sue potenzialità: questo tipo di intelligenze artificiali operano dopo essersi addestrate su ampi dataset di informazioni. Ci sono dataset che raccolgono testi di ogni tipo (come Common Crawl, usato fra gli altri per ChatGPT), e altri curati con un taglio più specifico (come LAION-5B o suoi subset, usati per Midjourney e Stable Diffusion, che abbinano immagini e descrizioni testuali).

La preoccupazione principale riguarda i contenuti presenti all’interno di questi dataset, che spesso raccolgono testi e immagini disseminati liberamente per Internet, anche quando sono protetti da proprietà intellettuale, e che sono recuperati senza il consenso esplicito degli autori. È il caso, per esempio, di Greg Rutkwoski, un artista digitale polacco dallo stile particolarmente riconoscibile, che ha scoperto di essere straordinariamente popolare tra gli utenti degli algoritmi text-to-image —al punto da trovare su Internet opere attribuite a lui,

facilmente confondibili con degli originali, ma che lui non aveva mai creato. È il caso, anche, dell'illustratrice Hollie Mengert, che si è ritrovata protagonista di un modello appositamente sviluppato per imitare il suo stile di disegno.

Il 13 gennaio 2023, un gruppo di artisti affiancati da avvocati specializzati in class action, ha fatto causa a Stability AI, Midjourney e DeviantArt per l'utilizzo di algoritmi addestrati su dataset contenenti opere artistiche senza il consenso degli autori.

I possibili impatti, però, non si fermano alle violazioni del diritto d'autore. Se questi algoritmi sono in grado di generare immagini ispirandosi allo stile di artisti specifici grazie a dataset che raccolgono anche le loro opere, allora chiunque li usi può generare autonomamente contenuti visivi in grado, quantomeno, di passare per originali. Le implicazioni, in questo senso, sono evidenti in termini di economia del lavoro. Io ci sono passato personalmente: all'inizio dell'anno ho aggiornato e rilanciato il mio blog personale e per un periodo ho presentato ogni post con un'immagine di copertina generata con Midjourney, come nel caso di questo articolo.

Queste immagini, diffuse attraverso i principali social network, risultavano credibili ed esteticamente coerenti con ciò che un lettore si potrebbe aspettare di vedere nella copertina di un post su un blog. Ho continuato, finché qualcuno mi ha fatto notare una cosa ovvia: per quanto il mio blog non sia uno spazio espressamente for-profit, i suoi contenuti mi permettono di ottenere nuovi lavori. Utilizzando un'immagine generata con Midjourney con l'intento di fornire al post un artwork di qualità percepibile come professionale, finivo indirettamente per privare un vero artista di un potenziale compenso. Non solo: nel farlo, pagavo un abbonamento a Midjourney, che aveva generato quelle immagini attingendo a un dataset contenente immagini di artisti che avrebbero potuto fare quel lavoro al posto dell'algoritmo. Da allora, ho smesso.

Esistono quindi dei rischi pratici, legati all'utilizzo di un algoritmo anziché la commissione a un artista, per la creazione di un contenuto visivo che passi quantomeno per credibile. Esistono, poi, dei rischi più filosofici e legati al nostro modo di percepire la creatività: se questi algoritmi permettono la generazione automatica (e rapida) di opere visive a partire dai loro dataset, sul medio termine il rischio è di assistere a una saturazione di immagini che potrebbero essere percepite come creative, ma che altro non sono che la ricombinazione delle caratteristiche di altre immagini già esistenti. In questo senso, è davvero possibile definire un'intelligenza artificiale creativa?

Secondo Marcus du Sautoy, un matematico della Oxford University intervistato dal *The Guardian*, è importante distinguere i modelli di diffusione latente (come quelli di Stable Diffusion e DALL-E), da quelli a rete generativa avversaria (GAN). Nel primo caso, secondo Sautoy, le intelligenze artificiali stanno replicando una creatività di tipo "combinativo, perché gli algoritmi sono istruiti a creare nuove immagini nello stesso stile di altre milioni presenti nel dataset." Nel caso delle GAN, invece, si può parlare di "creatività trasformativa" [...] perché permettono "uno scambio tra due reti, una che crea nuove immagini, l'altra che decide quanto bene quell'immagine generata rispecchia determinate caratteristiche," si legge nell'articolo del The Guardian. "Una GAN artistica potrebbe avere come obiettivo la creazione di un'immagine che sia quanto più possibile diversa da quelle presenti nel suo dataset, senza però allontanarsi dalla categoria di ciò che gli umani potrebbero interpretare come arte visiva."

Cosa possono fare gli artisti?

Come spesso accade quando si parla dell'impatto delle tecnologie digitali sugli esseri umani, esiste un piano istituzionale e un piano più pratico. Dal punto di vista formale, nei prossimi anni tribunali e corti saranno impegnate sempre di più nella discussione di casi di questo tipo e dovranno maturare (sperabilmente) le competenze necessarie per definire dei confini tecnici alla definizione di creatività legata alle intelligenze artificiali generative. Nel frattempo, però, questi strumenti sono già alla portata del grande pubblico ed è fondamentale per gli artisti potersi tutelare.

Strumenti come Have I Been Trained? permettono a chiunque di effettuare una ricerca all'interno del database LAION-5B, quello usato da Midjourney e altri, per verificare la presenza delle proprie opere o meno. Ancora, script come NO AI permettono a chiunque di applicare alle proprie opere originali un watermark che impedisce ai bot utilizzati per fare raccolta (scraping) di immagini di acquisirle.

Resta, però, un punto importante da affrontare: molto probabilmente, le intelligenze artificiali generative sono arrivate per restare. Altrettanto probabilmente, i loro casi d'uso reali saranno estremamente specifici e lontani da un'applicazione concreta per il grande pubblico. Immaginare forme di collaborazione con questi nuovi strumenti non è soltanto una contingenza storica, ma in un certo senso il compito stesso dell'artista che si ritrova a interagire sempre con nuovi strumenti.

Ponendo in primo piano le implicazioni etiche ed economiche dell'utilizzo di questi algoritmi, è possibile immaginare di includere il loro potenziale apparentemente creativo nei processi artistici. In questo senso, è fondamentale interpretare queste tecnologie non come strumenti, ma come interlocutori in grado di interpretare un ruolo specifico (deciso dall'artista) e cogliere i risultati generati dall'algoritmo come spunti creativi. È possibile, per esempio, chiedere a ChatGPT di simulare un personaggio specifico (un pesce palla in grado di pensare) e sottoporgli una parte dell'opera (per esempio, un racconto ambientato nell'oceano) per vedere generare dei risultati non fattualmente creativi, ma ispirati da una messinscena assurda e testabile in tempo reale.

La generazione di opere creative attraverso l'utilizzo di intelligenze artificiali generative è una tecnica indubbiamente popolarissima, in questo momento, ma non è ancora chiaro quanto realmente incisivi saranno gli impatti di questi strumenti sulle economie dell'arte e della creatività. Nel quotidiano sforzo di sensibilizzazione necessario a ristrutturare i proclama tecno-ottimisti che spesso interpretano queste tecnologie come panacee in grado di sollevare gli umani da qualunque incarico produttivo (spesso ignorando le conseguenze sociali di un'evoluzione di questo tipo), possiamo cogliere la possibilità per analizzare il modo in cui queste innovazioni stanno interagendo con il nostro modo di percepire la creatività. E interrogarci, soprattutto, sul fatto che stiano ampliando o restringendo le nostre possibilità immaginative.

Società

Non c’è AI senza Big Tech

Un rapporto dell’AI Now Institute analizza la corsa all’AI e il rischio che favorisca la concentrazione del potere nelle mani di grandi aziende tech. Quali sono i nodi da affrontare.

di **Giuditta Mosca**

Un rapporto dell’AI Now Institute, uscito ad aprile, condensa l’anima del dibattito più politico che si è sviluppato attorno alle intelligenze artificiali. Poco più di un centinaio di pagine nelle quali si sviscerano i risvolti etici, giurisdizionali, economici, sociali e politici collegati a uno sviluppo delle AI consegnato nelle mani di quelle che il report chiama Big Tech, termine usato nella più ampia accezione possibile anche se, di norma, è etichetta che si applica ai giganti noti, ovvero Alphabet, Amazon, Apple, Meta e Microsoft.

Fondato nel 2017, l’AI Now Institute si prefigge lo scopo di esaminare la concentrazione di potere nel settore delle tecnologie e sviluppare strategie per scongiurare la creazione di oligopoli. Nel 2021 è entrato nell’orbita della Federal Trade Commission (FTC) in qualità di consigliere.

Il messaggio che emerge dal rapporto è l’urgenza di norme che possano disciplinare l’uso delle AI, anche se queste non possono essere demandate soltanto alle autorità legislative e alla politica in genere, ma necessitano di una condivisione popolare perché il quadro normativo attuale - il riferimento è alle leggi americane e di sponda a quelle europee - può sembrare sufficiente o bene improntato soltanto a una prima lettura, ma non riesce davvero a mitigare il potere che si condensa nelle mani delle Big Tech. Non solo. Le AI non sono soltanto materia di competizione tra governi, hanno anche un impatto ambientale sulla cui gravità il report si concentra in chiusura.

Il rapporto tra AI e Big Tech

Guardando alle Big Tech come un’unica entità a prescindere dal marchio e dai brand specifici, il rapporto tra questa e le AI solleva molte domande. Il pensiero

corre in prima istanza a ChatGPT, che ha portato in primo piano il tema delle capacità degli ultimi sviluppi dell'intelligenza artificiale ma anche del suo impatto, e poi di quali dati sono utilizzati e come, argomento come sappiamo sollevato dal Garante per la privacy italiano che ha chiesto proprio a OpenAI dettagli su come vengono trattati i dati degli utenti. La lettura fornita dall'AI Now Institute è però di più ampio respiro e include tutte le tecnologie e il rischio che queste si concentrino nelle mani di poche aziende.

Il dominio che possono esercitare sulle AI le Big Tech verte soprattutto su tre ambiti:

- i dati: le aziende che hanno accesso alla più ampia gamma di informazioni hanno un vantaggio nell'uso delle AI, riuscendo così a penetrare in comparti sensibili come, per esempio, quello della salute, del credito, degli armamenti e dell'istruzione;

- la capacità di calcolo: le AI hanno bisogno di una potenza di calcolo e di quantità di dati non comuni, quindi non sono appannaggio di qualsivoglia organizzazione. Sono poche le aziende che hanno un'infrastruttura proprietaria e ne traggono un vantaggio, anche cercando di limitare l'uso dei loro dati, così come ha minacciato Microsoft nei confronti di motori di ricerca rivali;

- vantaggio geopolitico. Esiste un rapporto tra governi, difesa e le stesse Big Tech, rapporto che la corsa all'AI potrebbe stringere ancora di più mentre, al contempo, si instaurano frizioni dialettiche a distanza, come quelle tra Usa e Cina. La potenza economica delle Big Tech si riverbera sul mercato tant'è che le acquisizioni di aziende che si occupano di AI sono notizia che non fa più scalpore per il numero ma, casomai, per le somme delle offerte di acquisto, a cominciare da quella da 19,7 miliardi di dollari con la quale Microsoft nel 2021 ha messo nel carrello Nuance, impresa attiva nelle AI conversazionali e introdotta nel mondo della sanità, comparto per il quale fornisce diverse soluzioni. Non può passare inosservata neppure la cifra da 7,1 miliardi di dollari con cui Panasonic ha acquisito Blue Yonder nel 2021, dopo averne rilevato il 20 per cento del capitale azionario a metà del 2020. Senza dimenticare il recente investimento di Microsoft da 10 miliardi di dollari in OpenAI e l'integrazione di ChatGPT nel motore di ricerca Bing.

Le priorità a cui trovare risposta

Il rapporto dell'AI Now Institute si sofferma sul rischio che le Big Tech concentrino un immenso potere e che questo possa essere contenuto lavorando su aspetti cruciali, su quattro aree strategiche utili a mitigare gli effetti di un tale disequilibrio.

La prima è la necessità di imporre alle Big Tech l'onere di dimostrare che lavorano nel rispetto delle norme vigenti e in modo etico, senza costringere le autorità a continue indagini e a correre ai ripari dopo. L'idea è quella di replicare parte del lavoro già svolto dalle autorità che sorvegliano il comparto finanziario i cui errori possono avere ricadute pesanti sull'economia, anche se spesso sono imprevedibili e non calcolabili a priori.

La priorità numero due è creare un ponte tra l'operato delle Big Tech e quello della politica, panorami isolati tra loro perché mentre le aziende tecnologiche espandono i rispettivi mercati, molto spesso politica e autorità legislative non riescono a essere altrettanto espansive, anche a causa della limitata comprensione di ciò che le Big Tech fanno e dei metodi che usano per penetrare un numero maggiore di comparti economici.

Sempre in base al report dell'AI Now Institute, quando le autorità comprendono la materia su cui sono chiamate a legiferare esercitano un compito di maggiore spessore. Allo stesso modo, all'inizio del mese di aprile, un tribunale d'appello olandese ha stabilito che demandare a sistemi automatizzati le politiche salariali e la cernita dei dipendenti da licenziare (robo-firing) è un modus operandi illegale.

Il terzo punto su cui intervenire è limitare l'ingerenza dell'industria nella vita politica. Come esempio il report cita l'influenza esercitata da Microsoft sull'adozione nel 2020 da parte dello Stato di Washington di una legge che regolava il riconoscimento facciale (senza bannarlo). O ancora il fatto che proposte legislative come l'europeo AI Act si spostino da quadri normativi basati sui diritti ad altri impostati sul rischio. O che molti sforzi di policy proposti o spinti dall'industria puntino sproporzionatamente su concetti come auditing degli algoritmi, valutazioni di impatto e trasparenza, invece di affrontare alla radice la questione dei pregiudizi e di altri danni prodotti da questi sistemi. Ad esempio, l'AI Now Institute, a pagina 24 del rapporto, sottolinea la necessità che i regolatori della privacy debbano lavorare per impedire sorveglianza invasiva.

Inoltre, le aziende tecnologiche possono approfittare del loro potere economico e politico per eludere le norme in essere, come per esempio proporre soluzioni per la moderazione dei contenuti online che scavalcano i principi della privacy. Oppure possono sfruttare determinate crisi, come la guerra in Ucraina, per spingere soluzioni per il riconoscimento facciale.

Il quarto nodo è non limitarsi all'applicazione di norme e leggi ma fare leva sulle Big Tech affinché cambino cultura, coinvolgendo anche osservatori indipendenti e l'opinione pubblica in genere. Come esempio, il report cita il caso di un gruppo di suore che hanno creato un movimento attivista al fine di chiedere a Microsoft di non vendere ai governi le tecnologie per il riconoscimento facciale e di fare in modo che le attività di lobbying rimangano in linea con i valori e le politiche sociali che la stessa Microsoft ha dichiarato essere parte integrante della propria missione.

Queste quattro priorità strategiche possono essere estese a diversi ambiti, siano questi tecnologici, economici, sociali, politici o etici. A cominciare da ChatGPT.

ChatGPT e gli altri modelli su larga scala

ChatGPT ha acceso i riflettori dei media sulle AI nel loro insieme, risollevando anche questioni etiche tuttora irrisolte.

I modelli linguistici di grandi dimensioni, come quelli su cui si basa ChatGPT, sono ritenuti, da una parte dell'industria, un punto di svolta per il progresso scientifico nel suo insieme. Una narrazione che, secondo l'AI Now Institute, rischia di distogliere l'attenzione dai problemi reali e già documentati che parlano di discriminazioni, di violazioni della privacy e problemi di sicurezza.

Inoltre, la convinzione che simili modelli accorcino la strada che ci separa dall'intelligenza artificiale generale (AGI) sarebbe una distorsione della realtà, sostiene l'AI Now Institute.

Infine, i regolamenti esistenti non si concentrano su un elemento di rilievo. La capacità computazionale necessaria all'esistenza e allo sviluppo dei modelli di AI è di proprietà di un numero ristretto di aziende e non potrebbe essere diversamente, considerando che il costo di esecuzione di ChatGPT è stato stimato da alcuni in 3 milioni di dollari al mese (mentre altri hanno stimato i costi di addestramento di un altro modello linguistico di grandi dimensioni, Pathways Language Model, PaLM, di Google, tra i 9 e i 23 milioni di dollari).

Anche l'uso della dicitura open source spesso associata ai modelli è fuorviante. Di sicuro non si applica al leader attuale di mercato, OpenAI, che non rilascia informazioni sull'architettura e sulle dimensioni del modello, così come informazioni relative all'addestramento utilizzato e all'hardware impiegato.

Regolamentazione dell'uso dei dati

Le Big Tech traggono vantaggio dai dati e dall'uso che ne fanno e questo, sempre stando al rapporto dell'AI Now Institute, è un potere che può essere frenato soltanto lasciando agire in concerto le norme sulla privacy e il diritto alla concorrenza. Strumenti che dovrebbero essere tenuti saldamente tra le mani della politica che, però, tende a interpretarli come separati l'uno dall'altro, favorendo così le imprese tecnologiche che dalla sorveglianza invasiva traggono vantaggi competitivi.

È un tema che le autorità preposte stanno considerando. Lo dimostra il fatto che, nel 2022, quando Amazon si è mossa per acquisire One Medical (per 3,9 miliardi di dollari) e iRobot (per 1,7 miliardi di dollari) la Federal Trade Commission (FTC) ha aperto un'inchiesta per comprendere se Amazon avrebbe potuto trarre un vantaggio considerevole in due mercati considerati emergenti. La proposta di acquisizione di iRobot ha portato anche all'apertura di un'inchiesta della Commissione europea.

Un esempio simile si è verificato in UK nell'ottobre del 2022, quando la Competition and Markets Authority (Cma) ha ordinato a Meta di vendere Giphy (noto sito per condividere GIF, acquisito nel 2020), giudicando l'acquisizione limitativa per la scelta degli utenti britannici e in grado di togliere ossigeno all'innovazione e concorrenza.

Gli audit non bastano

Un audit è una valutazione indipendente che ha lo scopo di considerare la bontà di una pratica in relazione alle norme e ai regolamenti in essere. Attualmente, come sottolinea il rapporto a pagina 34, le Big Tech tendono a influenzare il mercato dell'auditing, svolgendo al proprio interno test e valutazioni basate su strumenti proprietari creati appositamente. Gli auditor esterni tendono a loro volta a fissare parametri di valutazione vaghi. È un tema che l'AI Now Institute

rilancia sulla scorta di una ricerca dal titolo "Chi controlla i controllori?" pubblicata nel 2022 dalla The Algorithmic Justice League, non profit per la difesa digitale con sede nel Massachusetts.

Il risultato di questi audit rafforzerebbe la posizione della grande industria tecnologica e farebbe ricadere le responsabilità su chi ha meno potere, ovvero il pubblico e le aziende tech meno grandi e con meno risorse che usano i modelli di AI per produrre beni o servizi. Il frequente ricorso ad audit esterni dunque tenderebbe a rafforzare il potere delle Big Tech nel campo della AI invece di evidenziare danni potenziali causati da queste.

Il risultato è una fiorente economia dell'Audit As a Service con insufficiente chiarezza riguardo ai metodi e agli standard di valutazione adottati, e con il rischio che il concetto di audit in sé venga relegato a un mero esercizio procedurale, un meccanismo di verifica gestito dalle Big Tech senza l'intervento di un quadro politico-normativo. Il Digital Service Act europeo (Dsa) invertirebbe soltanto parzialmente questa rotta, sostiene il report, imponendo agli auditor di verificare i rischi sistemici delle AI per i diritti fondamentali ma lasciando alle aziende che svolgono la revisione il compito di individuare in modo autonomo gli ambiti nei quali svolgere l'audit.

Una strada auspicata dall'AI Now Institute è quella di coinvolgere negli audit anche le comunità direttamente interessate alle quali affidare test misurabili e riproducibili, al fine di partecipare in modo attivo al miglioramento delle tecnologie AI.

La riforma strutturale delle Big Tech

La debolezza degli organi per la libera concorrenza, le autorità antitrust, ha permesso alle Big Tech di entrare in molti mercati, togliendo ossigeno a realtà imprenditoriali già presenti nel medesimo settore, sostiene poi il rapporto.

È vero che le autorità si stanno prodigando per correggere il tiro ma è anche vero che, muovendosi tardivamente, hanno accumulato un gap che diventa difficile da colmare, anche perché alle iniziative prese dai regolatori (tanto quelli americani quanto quelli europei), le Big Tech hanno risposto con politiche di lobbying aggressive. Il report sostiene ad esempio che l'attività di lobbying delle grandi aziende tecnologiche sia aumentata in seguito all'introduzione negli Usa

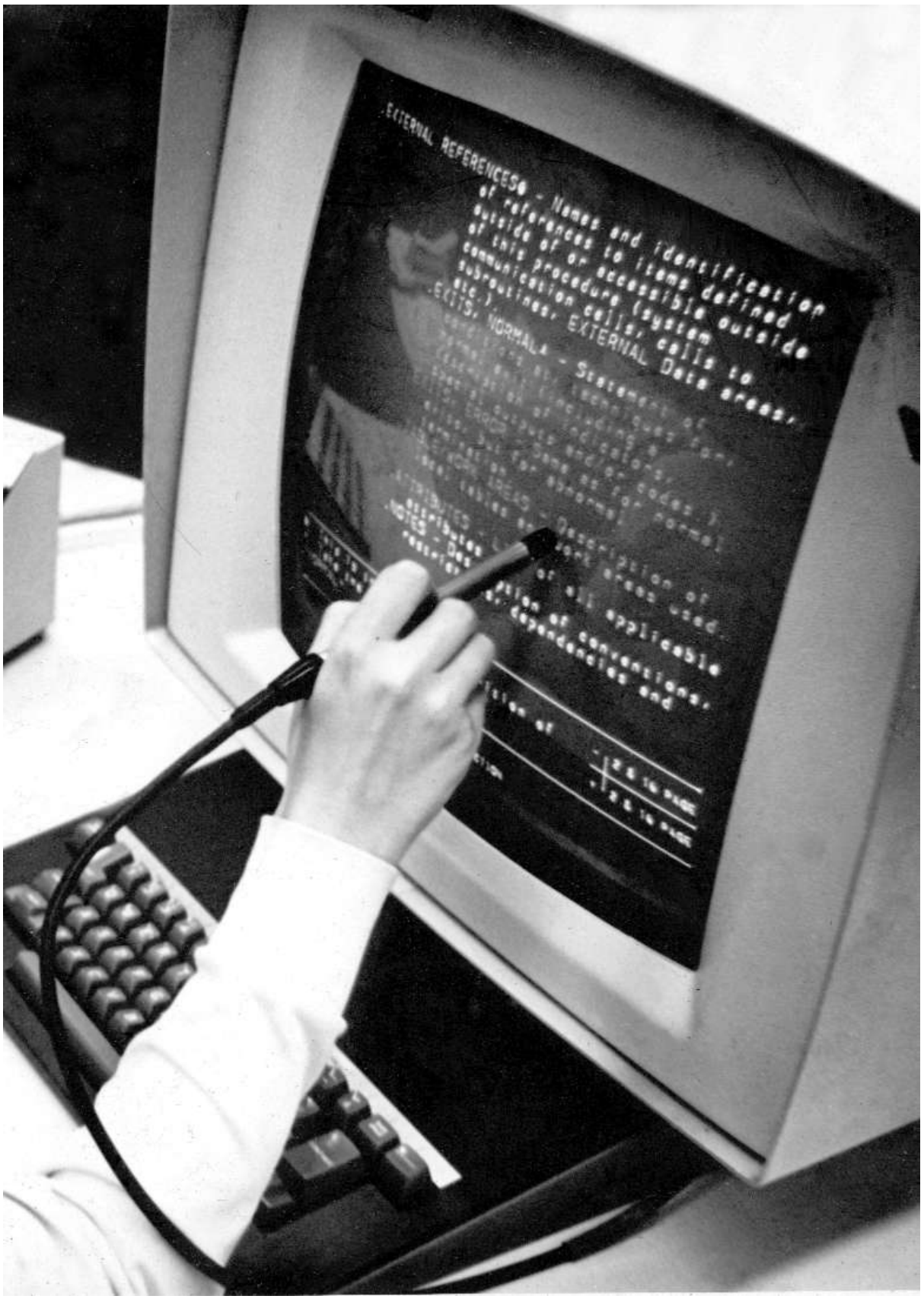
delle proposte di legge sulla concorrenza (antitrust bill package), aumentando di quasi 95 milioni di dollari tra il 2021 e il 2022.

Il 4 maggio appena trascorso, l'amministrazione Biden ha annunciato un insieme di norme per promuovere la responsabilità delle imprese che operano nel comparto delle AI, e questo va nella direzione auspicata dall'AI Now Institute.

La questione ambientale

Infine c'è la questione ambientale. Le infrastrutture che consentono alle Big Tech di avere la supremazia sui dati e, quindi sulle AI, hanno un costo in termini ambientali e conseguenze che richiamano il colonialismo. I data center necessitano di energia elettrica e di acqua tanto per lavorare quanto per il loro raffreddamento. La produzione di energia elettrica da fonti non rinnovabili immette CO_2 nell'atmosfera e, in modo più sottile, gli approvvigionamenti di acqua drenano le forniture idriche pubbliche, aumentando lo stress idrico delle regioni in cui sono ubicati.

Si tende così a impossessarsi di risorse e di terre già provate dagli eventi climatici. "Le aziende spesso attingono alle forniture idriche pubbliche, già sottoposte a stress dopo decenni di crescita e di scarsi investimenti nelle infrastrutture pubbliche, anche se la quantità prelevata dalle stesse è difficile da verificare perché l'uso dell'acqua da parte di molte aziende non è trasparente", scrivono gli autori. Malgrado le promesse fatte da alcune grandi aziende di ridurre i consumi (promesse che il report derubrica a greenwashing), in molti casi non riuscirebbero a raggiungere nemmeno i modesti obiettivi che si sono dati.


EXTERNAL REFERENCES - Names and identification
of references to items defined
outside of or accessible outside
of this procedure (system
communication calls, calls to
subroutines, EXTERNAL Data areas,
etc.)
EXITS, NORMAL -
EXTERNAL

Politica

Il dibattito sulla generazione automatica di disinformazione

Abbiamo chiesto a tre esperti (la giornalista di Newsguard Virginia Padovese, la filosofa Mariarosaria Taddeo e il co-relatore dell'AI Act Brando Benifei) cosa ne pensano del rischio che l'AI amplifichi la disinformazione.

di **Antonio Piemontese**

Elezioni per il Parlamento europeo, presidenziali americane: sono due i principali appuntamenti politici in programma per il 2024. Ma non certo gli unici. Tutti legati da un filo conduttore: quello presente è un decennio di snodo. Dalle guerre alla gestione del cambiamento climatico, dalla transizione energetica alle nuove tecnologie via via fino ai temi etici e diritti, le democrazie si trovano a confrontarsi con scenari complessi e, in molti casi, difficili da prevedere.

Ma c'è una variabile in più: sulle consultazioni si sta palesando lo spettro di un'ondata di disinformazione creata sfruttando la potenza dell'intelligenza artificiale generativa, ora disponibile al grande pubblico. La denuncia arriva da Newsguard, società americana che si occupa di monitoraggio della disinformazione online. Un fenomeno non nuovo (sono note le content farm ospitate in Paesi come la Macedonia): ma la scala resa possibile dai nuovi strumenti non ha precedenti.

Newsguard e le fabbriche di disinformazione

Al momento in cui scriviamo, nel giugno 2023, Newsguard ha identificato 277 siti di notizie o informazioni presumibilmente creati dall'intelligenza artificiale e che opererebbero senza alcuna supervisione umana. Esiste un acronimo per definirli: "UAIN" (dall'inglese "Unreliable AI-generated News"). I siti censiti hanno spesso nomi generici che possono indurre in errore, come iBusiness Day, Ireland Top News e Daily Time Update, e sono in 13 lingue: arabo, cinese, ceco, coreano, francese, indonesiano, inglese, italiano, olandese, portoghese, tagalog, thailandese e turco.

Alcuni arrivano a produrre centinaia di contenuti al giorno. Ma, argomenta Newsguard, con ogni probabilità si tratta di realtà in cui l'apporto umano è irrilevante: non esiste alcuna redazione in grado di tentare una verifica, persino gli account dei giornalisti sono fasulli.

Difficile risalire ai proprietari delle testate: le informazioni riportate - sempre che la legge locale lo richieda – si presentano vaghe o lacunose. "Abbiamo provato a contattare alcune di queste realtà, ma ci hanno risposto in pochi", dice a Guerre di Rete Virginia Padovese di Newsguard. "Alcuni si sono fatti vivi, senza però affrontare le nostre domande. Solo due hanno ammesso chiaramente di usare l'intelligenza artificiale per produrre contenuti".

Non è possibile affermare con certezza se un sito è gestito da bot senza verificare dietro le quinte – prosegue l'esperta - ma ci sono alcuni indicatori che lasciano supporre con ragionevoli margini di approssimazione che sia così. Ad esempio, la presenza di errori ortografici e grammaticali frequenti e non corretti, il linguaggio banale, le frasi ripetitive. Tutte caratteristiche tipiche dei contenuti prodotti dalla AI.

La società americana ha fatto dei test con strumenti di AI generativa. "Abbiamo provato a produrre disinformazione partendo da bufale", prosegue Padovese. "A volte è stato necessario insistere, ma è bastato poco per superare le resistenze: con ChatGPT 3.5 abbiamo avuto successo nell'80 per centi dei casi, con la versione 4 addirittura nella totalità. Si tratta, peraltro, di uno strumento molto più potente. In alcuni casi appariva una nota in fondo al testo, una specie di disclaimer che consigliava di consultare una fonte attendibile; ma, tolta l'ultima frase, l'articolo potrebbe ingannare chiunque perché molto ben scritto"

"Il problema è anche economico", aggiunge Padovese. "Queste news vanno a minare il giornalismo di qualità, che ha bisogno di fondi per assumere personale specializzato, fact-checker, editor. Il costo di produzione delle content farm, invece, rasenta lo zero".

Un altro problema è che molti di questi siti privi di supervisione umana pubblicano annunci di brand i quali, senza consapevolezza, finiscono per farsi pubblicità su spazi che promuovono disinformazione. "Il problema è noto: la pubblicità programmatica è gestita in modo opaco, perché i marchi che chiedono alle agenzie di raggiungere determinati target di audience non sanno dove appariranno gli annunci", prosegue Padovese. "Certo, esistono strumenti di brand safety, ma escludere i siti di disinformazione è complicato: fino a che un brand non saprà dove finisce la propria pubblicità, questi siti non affidabili

continueranno a essere foraggiati e a proliferare. Per questo è importante che le agenzie pubblicitarie e tutti gli attori della filiera abbiano accesso a database che li censiscono, e, soprattutto, che li escludano dal bouquet".

La filosofa Taddeo: "Una Cambridge Analytica sotto steroidi"

Mariarosaria Taddeo è professoressa e senior research fellow all'Oxford Internet Institute. Un lungo curriculum imperniato sull'analisi dei rapporti tra etica, digitale e società, è tra le studiose più accreditate a livello europeo in materia.

"Non mi preoccupano le derive fantascientifiche che vedono nella AI un rischio esistenziale per la razza umana", commenta a Guerre di Rete la studiosa, "ma sono molto preoccupata per la qualità delle informazioni che circoleranno sempre di più: pensiamo se i complottisti avessero avuto a disposizione questi strumenti durante la pandemia.

Certo, i bot su Twitter esistevano già, ma abbiamo fatto un salto di qualità. Anche economico: oggi, con un investimento esiguo si può creare disinformazione targettizzata, ad esempio diretta specificamente alle casalinghe del Centro Italia, che usi uno stile semplice, citi argomentazioni e temi rilevanti, o che, addirittura, li suggerisca. Quando questo tipo di contenuti si innesta sui social network", argomenta Taddeo, "si determina una combinazione perfetta: una Cambridge Analitica, ma, questa volta, sotto steroidi".

Gli argini contro l'uso improprio della AI

Come fermare l'invasione di disinformazione prima che inquini nuovamente le campagne elettorali è la domanda che i legislatori hanno cominciato a farsi da qualche mese.

"L'educazione digitale può aiutare senz'altro il cittadino, ma non si può pensare di fare un corso intensivo a tutta la popolazione su come usare i media", riprende Taddeo. "Penso che la scelta più efficace sia richiamare alla responsabilità i fornitori di tecnologia, non solo quelli alla ChatGPT, ma anche piattaforme e social network. Pensare che si smetterà di utilizzare l'AI per produrre disinformazione è illusorio: possiamo, però, regolamentarne l'uso sui social,

imponendo un controllo maggiore. Il controllo sui contenuti è sempre problematico, perché volenti o nolenti ci trova nella situazione di dover effettuare una scelta. La domanda è: a che punto questa operazione diventa censura?".

Il nodo delle verifiche

Arriviamo a un altro nodo: chi fa le verifiche? "Tutti dovremmo verificare le informazioni di cui fruiamo, ma penso che tocchi soprattutto ai gestori delle piattaforme. Non si può pensare che siano i cittadini, si troverebbero sovraccaricati. I tech providers non producono solo tecnologia: hanno creato un ambiente nel quale la nostra società esiste. E quando disegnano, progettano quell'ambiente, progettano anche le cose che lì possiamo e non possiamo fare. Come fornitori di tecnologie cruciali arriva una responsabilità che porta ad assolvere a doveri più pesanti rispetto ad altri tipi di realtà. Per loro immagino un ruolo da gatekeeper dell'informazione, regolando le informazioni sulla base di criteri scelti dai legislatori e in grado di bilanciare diritto all'informazione con le misure per mitigare i rischi seri della disinformazione".

Gli scandali non sono passati invano. L'attenzione mediatica, sostiene Taddeo, ha modificato l'atteggiamento dei provider. Il punto di svolta è stato il caso di Cambridge Analytica, con un forte impatto sulla politica nel Regno Unito e negli Usa.

"Si è capito, in sostanza, che il digitale può fare cose importanti per supportare i processi democratici e cose terribili per fuorviarli. E questo ha spinto in primo luogo l'Unione Europea a iniziare a fare delle regolamentazioni serie, come la GDPR, il Digital Services Act, il Digital Market Act e l'AI Act, che sono figlie di un'onda lunga partita in quelle settimane".

In risposta a questa stretta regolamentativa, le società tech "hanno cominciato in prima istanza a proporre una narrativa di self governance: ci diamo dei principi da soli. Non è servito a molto, ma ha senz'altro creato un discorso all'interno di queste aziende sulle implicazioni etiche dei loro prodotti e servizi. Per proteggere la democrazia sono servite regolamentazioni, che sono complesse e lunghe da definire, che chiarissero obblighi e modi di erogazione dei servizi e dei prodotti digitali. È la governance etica del digitale, indispensabile per sfruttarne il potenziale positivo e limitarne i rischi. Io credo che, per essere efficace, forse più che in altri ambiti, la governance del digitale richieda un tavolo composito,

dove c'è spazio per il legislatore, che ha il compito di definire una strategia sulla base di una visione politica, ma anche per i tech provider. Devono essere trovati dei compromessi accettabili per tutti", chiosa Taddeo.

L'innovazione, però, avanza a tappe forzate. Come può uno Stato democratico restare al passo? "È il dilemma di Collingridge: se la regoliamo troppo presto, rischiamo di rallentare o fermare l'innovazione perché non si è ancora capito quali sono i rischi che vuoi mitigare; se si interviene troppo tardi, si impianta nella società in modo organico per conto suo", conclude Taddeo. "Ma c'è una finestra temporale, un buffer in cui ci si può agire. Io sono una grandissima sostenitrice di questa finestra, perché penso che proprio lì stia lo spazio del dibattito pubblico. Il buffer tra innovazione e regolamentazione è quello della democrazia, in cui si discute su quanto possiamo normare i contenuti online, su dove investiamo i soldi per la ricerca nel nostro Paese, su qual è il ruolo delle aziende tech. Questioni complesse da dirimere, che richiedono un tempo di osservazione e di discussione. L'AI Act, per esempio, è un modello di norma ben scritta. Concepita prima dell'esplosione di ChatGPT, sono riusciti a cambiare il testo in modo da considerare anche i modelli generativi, dimostrando che la norma è perfettibile. Lo trovo salutare, credo che la classe dirigente europea abbia dimostrato grandissime competenze. L'opposto è il caso dei regimi autoritari, dove i governi decidono le aziende da sostenere, le innovazioni da sviluppare, usi accettabili e non".

Il co-relatore dell'AI Act Benifei: "Applicare le norme sull'AI da subito"

Brando Benifei è un europarlamentare italiano, capodelegazione del Pd. Giovane, ma con due mandati già alle spalle, è co-relatore dell'AI Act, la normativa, recentemente approvata dall'assise di Strasburgo, che regolerà l'intelligenza artificiale a livello continentale (ne abbiamo scritto qui). Ora comincia la fase del cosiddetto "trilogo", in cui il testo dovrà essere palleggiato a tre con Consiglio e Commissione, prima di ricevere il via libera definitivo. Da politico, Benifei è preoccupato per il possibile uso della AI nelle elezioni. "Ritengo che in futuro sia plausibile trovarsi di fronte a un utilizzo più sofisticato dei metodi di produzione delle fake news rispetto a quanto siamo abituati, anche in chiave elettorale", afferma raggiunto al telefono dal Guerre di Rete. "Ma il Digital Services Act e l'AI Act, nonostante per quest'ultimo l'iter sia ancora in corso, offrono alcuni argini.

Per esempio sulla produzione di deepfakes, che dovranno essere resi riconoscibili".

Osserviamo che l'AI Act, nella migliore delle ipotesi, entrerà in vigore poco prima delle elezioni del 2024. "Ma ci sono altre due iniziative che possono essere utili a tutelare il dibattito", ribatte l'europarlamentare. La prima va sotto il nome di AI Code of Conduct, un set di regole base sull'intelligenza artificiale generativa concordato tra i Paesi del G7. "Si tratta di un'iniziativa internazionale che individua requisiti minimi, non massimi, di trasparenza: in questo modo la cooperazione internazionale potrebbe anticipare gli effetti dell'AI Act". La seconda "va sotto il nome di AI Pact. In sostanza, è la richiesta da parte della Commissione europea alle aziende di cominciare ad applicare l'AI Act anche prima dell'effettiva entrata in vigore, per prepararsi, e tutelarci dalla produzione di fake news e deepfakes".

Il patto è alle battute iniziali: a lavorarci è il commissario per il mercato interno e i servizi Thierry Breton. Ma non può diventare un punto d'arrivo. "È importante approvare il regolamento sull'intelligenza artificiale entro la legislatura", sottolinea Benifei. "Mancano meno di dodici mesi, e, se così non fosse, la norma rischia di slittare di anni". E l'Europa non può permetterselo.

Le pressioni delle lobby

Chi, invece, gradirebbe uno stop a tempo indeterminato sono le lobby, i portatori di interessi di alcune aziende che stanno investendo nell'AI. C'è chi dice che il testo sia già stato depotenziato nella versione che si avvia al trilogo. Lei che ne pensa? chiediamo. "Se si riferisce alle rivelazioni sulle pressioni da parte di OpenAI, è chiaro che tutte le imprese sono portatrici di interesse legittimi, che in quella occasione si sono palesati", commenta Benifei." A livello politico, i gruppi del centrodestra hanno opposto resistenza sulla regolamentazione dell'AI generativa; quella dei popolari, conservatori e una certa quota di liberali, con una posizione di forte contrarietà che non ci consentiva di avere i numeri". Ma, prosegue l'europarlamentare, "devo ammettere che, da quando il tema è finito sui media, le posizioni si sono evolute e si è arrivati a un compromesso".

Al ribasso? "Credo sia una valutazione errata. Io non so se questo compromesso sia dovuto a pressioni lobbistiche: credo che però alla fine la politica debba prendere decisioni autonome. Proponevamo di inserire la AI generativa tra le applicazioni ad alto rischio: ma alla fine, per riuscire a far passare il testo,

abbiamo accettato di costruire una casistica a sé, che include una parte degli obblighi per la AI generativa, anche se non tutti. Considerato la situazione sfavorevole da cui siamo partiti, direi che è un compromesso al rialzo".

Ma il pericolo non è scampato. "Temo che quando andremo a negoziare coi governi si manifesteranno pressioni, beninteso legittime, da parte delle grandi imprese tech sugli esecutivi e i negoziatori per annacquare il testo", riprende il politico. Che non vede un rischio di obsolescenza precoce delle norme. "Penso che la formula individuata di non regolare le tecnologie ma gli usi aiuterà. Inoltre, abbiamo chiaramente previsto modalità di aggiornamento più rapide tramite i cosiddetti 'atti delegati della Commissione europea'. Speriamo non si cambi in fase di trilogo. Non saremmo d'accordo, per esempio, se lo strumento prescelto fossero gli atti di implementazione, che prevedono scarso coinvolgimento del Parlamento. Sarebbe una delega quasi in bianco alla Commissione".

E, sulla questione, è indispensabile agire il più possibile alla luce del sole, giocando a carte scoperte.

Strumenti

Se ne sapete poco, partite da qui

Una rassegna minima (e spiegata in modo semplice) di strumenti e servizi di AI. Cosa sono, dove trovarli.

di **Stefano Casini**

Il mondo dell'intelligenza artificiale generativa è balzato al centro dell'attenzione mondiale a partire dalla fine del 2022, con la diffusione di ChatGPT, il chatbot basato sul modello GPT-3.5 e GPT-4 sviluppato dalla società OpenAI. ChatGPT rimane una delle applicazioni rivolte al grande pubblico più note e utilizzate, ma è solo una delle tante all'interno di una costellazione di app, strumenti, siti web in continua evoluzione. Qualche esempio? Bing, MusicStar, Dall-E, Bard, CoPilot, Replicate, solo per citarne alcuni.

Le applicazioni di AI generativa più note al grande pubblico, da ChatGPT a Midjourney, si basano sui foundation model, modelli di base addestrati su enormi quantità di dati, "in grado diapprendere la distribuzione di probabilità che è sottostante ai dati di addestramento,e utilizzare questa distribuzione per generare contenuti nuovi ma simili in stile e struttura ai dati di addestramento", spiega a Guerre di Rete Irene Di Deo, ricercatrice senior dell'Osservatorio Artificial Intelligence del Politecnico di Milano.

Dove e come funziona meglio la Generative AI

Quindi, almeno attualmente, quali sono le applicazioni e gli ambiti dove funziona meglio l'intelligenza artificiale generativa, per quello che è l'uso più comune e non specialistico?

"L'area Testo è quella che ha un'applicazione più immediata", sottolinea Di Deo, "per esempio, conosco poco un tema e ho bisogno di unoscreening inizialedelle informazioni; oppure voglio evitare l'effetto 'pagina bianca' e generare contenuti più velocemente, magari per post e messaggi sui social network, o contenuti per un blog o sito web. O anche, inserisco già del contenutoe chiedo all'AI di sintetizzarlo o rifarlo e rimodularlo per uno specifico contesto".

L’onda innovativa di ChatGPT e dei suoi simili

ChatGPT è l’acronimo di Chat Generative Pretrained Transformer. È disponibile in versione gratuita – la 3.5 – e in versione Plus a pagamento con abbonamento – ChatGPT 4 – per funzioni più ampie ed evolute. La normativa a tutela della privacy negli Stati Uniti è diversa da quella europea, per cui a fine marzo scorso il Garante italiano per la protezione dei dati personali (Gpdp) è intervenuto bloccando di fatto l’attività di ChatGPT in Italia per alcune violazioni. OpenAI è quindi corsa ai ripari e la piattaforma a fine aprile è stata resa di nuovo disponibile garantendo più trasparenza e più diritti a utenti e non utenti europei.

OpenAI ha anche stretto un’alleanza con Microsoft (che ha investito miliardi nella società, divenuta ormai for-profit) che ha integrato questi modelli all’interno di Bing, il suo motore di ricerca, trasformandolo in un assistente virtuale a tutti gli effetti.

Machine learning e specialisti in AI

“ChatGPT è in grado ad esempio di rispondere a domande di follow-up, ammettere i propri errori, contestare premesse errate e rifiutare richieste inappropriate”, fa notare la ricercatrice dell’Osservatorio sull’AI, “la versione Free ha un aggiornamento dei dati che si ferma al 2021, con un sostanziale ‘buco’ di tutto ciò che riguarda gli ultimi 2 anni”, mentre la versione Plus comprende funzioni come la possibilità di attivare plugin.

L’applicazione più recente di ChatGPT funziona con un algoritmo in grado di gestire diversi tipi di dati, non solo testo, ma anche immagini. E risulta ulteriormente migliorata rispetto alle edizioni precedenti in quanto il sistema è stato addestrato non solo attraverso il machine learning ma anche con il contributo di specialisti AI in carne e ossa, che contribuiscono a ridurre i margini di imprecisione e di errore. La casa madre di ChatGPT, OpenAI, ha anche una piattaforma online per iniziare a sperimentare e poi utilizzare al meglio alcuni tra i principali strumenti di AI generativa, tra le cui sezioni si trovano varie informazioni, istruzioni operative ed esempi pratici.

C’è anche una OpenAI community e il suo forum con cui gli utenti si confrontano, parlano dei temi e delle applicazioni più varie, si scambiano informazioni e consigli su come utilizzare al meglio questi strumenti.

Bing, l’intelligenza generativa di Microsoft

In tempo reale e anche in forma gratuita è l’offerta di Bing, il motore di ricerca e chatbot con intelligenza generativa di Microsoft basato su GPT-4 di OpenAI.

Bing, oltre a elaborare i testi richiesti, inserisce in automatico anche fonti e link. In più, chiede all'utente quale stile di risposta preferisce, se più 'precisa' (ovvero più fattuale e concisa) o 'creativa' (più originale). Un altro tool simile a Bing è YouChat.

Risultati sorprendenti e limiti da superare

“Va sottolineato che con questi sistemi tutta la generazione di testo è basata sui dati di addestramento e sulla probabilità di elaborazione, nella costruzione e nel completamento delle frasi”, rimarca la ricercatrice dell’Osservatorio Artificial Intelligence, “e questo è un meccanismo che dà risultati sorprendenti ma allo stesso tempo non è sempre del tutto corretto e infallibile. A volte può portare a esiti non esatti o inventati”.

Bard è invece il chatbot di Google (basato prima sul modello linguistico di grandi dimensioni LaMDA e poi su PaLM 2) e, come rimarca lo stesso colosso hi-tech nelle sue ‘istruzioni’ d’uso, “è un’AI sperimentale e potrebbe fornire risposte imprecise o inappropriate. Puoi contribuire a migliorare Bard lasciando un feedback”.

Come per gli altri strumenti di AI disponibili al grande pubblico, ci sono da tenere presenti gli aspetti che riguardano la privacy dell’utente, i suoi dati e informazioni che vengono messi online, i contenuti delle conversazioni tra la persona e la macchina. Google stesso rimarca (abbastanza) chiaramente: “non includere nelle conversazioni con Bard informazioni che possano essere utilizzate per identificare te o altri utenti”.

Ci sono anche i chatbot di Character AI, con cui si possono creare degli avatar e vari personaggi virtuali con cui si può interagire online, e che hanno un loro particolare stile di linguaggio e di risposta, in linea con il personaggio che rappresentano e riproducono: si può quindi chattare con Super Mario Bros e Luke Skywalker, con uno psicologo, un insegnante di latino o un attore.

L'AI generativa di immagini

Ci sono poi i sistemi generatori di immagini, che in alcuni casi sono meno intuitivi rispetto a quelli testuali e un po' più complicati da utilizzare per un utente poco esperto, anche se si può usare la modalità text to image (da testo a immagine), con cui si scrivono le istruzioni da seguire e gli "ingredienti" da utilizzare, e lo strumento crea l'immagine richiesta. Sono disponibili versioni gratuite e altre con la possibilità di provare a generare alcune immagini, per poi passare a servizi in abbonamento per attività ulteriori.

Un generatore di immagini è ad esempio Dall-E, ora disponibile nella versione 2, anche questo una creatura di Open AI come ChatGPT. Funziona così: inserendo un testo con le istruzioni, il sistema fornisce 4 immagini e varianti che corrispondono alle richieste inserite, che poi si possono modificare ulteriormente, ad esempio con effetti di estensione e zoom. Uno dei punti di forza di Dall-E 2 sta nella possibilità di utilizzare il classico pennello per aggiungere dettagli alle immagini create come ombre, luci, effetti, colori.

Apprendimento su immagini raccolte online

Per quanto riguarda la generazione di immagini in formato digitale, la piattaforma online di OpenAI comprende sezioni dedicate anche a queste funzioni e applicazioni. Per creare immagini generate dall'AI, il modello di apprendimento automatico è addestrato su milioni di immagini raccolte su Internet insieme al testo a esse associato (abbiamo visto in altri articoli come questo stia sollevando domande e proteste in relazione all'uso di tali immagini). Dopodiché, a partire da un testo dato, può creare immagini nuove (text to image appunto).

Lo strumento di graphic design Canva ha aggiunto anche un generatore di immagini AI online gratuito.

Un'altra applicazione gratuita è Starry AI, con una semplice interfaccia, oltre 16 stili grafici diversi, che lascia all'utente la proprietà dell'immagine ed è anche generatore di certificati Nft (Non fungible token), che tracciano la proprietà e l'unicità delle opere e dei beni digitali. Anche il sistema Dream creato da Wombo è gratuito, semplice da utilizzare e adatto agli utenti meno esperti.

Le immagini di una vita non saranno più le stesse

Un altro strumento di generazione immagini molto noto è Stable Diffusion, frutto del lavoro della startup Stability AI e altri soggetti, simile a Dall-E e di cui sono disponibili tante versioni, da cui dipende anche la qualità finale delle immagini generate. Una delle funzioni attivabili è per esempio la 'Negative prompt', con cui è possibile indicare cosa non si vuole vedere nell'immagine.

Un'altra applicazione, a pagamento, è Midjourney, utilizzabile attraverso l'app Discord. È una soluzione meno intuitiva e più complessa di altre, ma sa generare immagini di forte impatto e alta qualità. Il sistema è stato premiato al concorso di belle arti alla Colorado State Fair, con il dipinto intitolato 'Théâtre d'Opéra Spatial'. Altri strumenti e soluzioni sono poi ad esempio NightCafè, Deep AI, Jasper Art e Photosonic.

Con il sito Replicate, invece, non si creano immagini nuove ma si offrono varie funzioni grafiche tra cui quella di riportare a nuovo vecchie fotografie in bianco e nero, migliorate con qualità digitale, anche per quanto riguarda quelle sfuocate o scattate non proprio secondo i canoni di un grande fotografo. Per la creazione e l'editing di video, anche attraverso avatar digitali, si possono utilizzare strumenti come Deepword e Rephrase.ai.

Restano aperte e in evoluzione, come già accennato, la questione del copyright, e della proprietà e trattamento delle immagini generate dall'AI.

Applicazioni digitali dalla musica alla scienza

Ci sono poi strumenti di AI generativa per il mondo della musica che permettono di creare brani e melodie liberamente utilizzabili perché non protette da copyright, o di riprodurre e modificare voci e registrazioni. È il caso, ad esempio, di MusicStar.ai, Beatoven.ai e Supertone.

Con Adobe Podcast, ancora in versione Beta, si può migliorare la qualità di un audio: per esempio, si ottiene la trascrizione dei testi dell'audio, si può editare e correggere la trascrizione, e il sistema in automatico corregge l'audio originale.

Il tool Vocal Remover da una canzone permette di creare due tracce audio separate per voce e musica strumentale, in modo da poterle sentire separatamente.

Perplexity.ai è un chatbot (che sembra essere rivolto più a studenti universitari, ricercatori, docenti, studiosi) che permette di cercare informazioni in modo più specifico.

L'assistente virtuale per fare coding e altri strumenti di produttività

GitHub CoPilot è un assistente virtuale per fare coding, per la scrittura di codice informatico, nato da una partnership fra GitHub e OpenAI, e addestrato su miliardi di linee di codice. Suggerisce codice sulla base del contesto, ed è integrabile con i più comuni software usati dai programmatori.

Con un tool come Tome, invece, si possono creare slide inserendo del testo e generando una presentazione a tema, mentre un altro strumento di lavoro e produttività individuale è Otter.ai, che permette di fare trascrizioni delle riunioni, note automatiche, pianificazione delle attività e del calendario.

Mentre con Re-imaging home si ottengono dall'AI suggerimenti e simulazioni per rinnovare casa, stanze e ambienti. Attraverso immagini virtuali dei locali, l'applicazione ci fa vedere come si può modificare l'arredamento e il design degli interni secondo vari stili e soluzioni.

Prompt design: imparare a usare la Generative AI

Il prompt non è altro che il testo che si utilizza per richiedere alla AI generative di fare delle azioni, appunto ottenere dei testi, immagini, altre elaborazioni. Ad esempio, può essere utile tenere presente che, per quanto il risultato sia sorprendente anche in italiano, utilizzare prompt in inglese è più efficace.

Il prompt design, cioè il modo di utilizzare queste istruzioni, svolge un ruolo fondamentale nell'interazione tra la persona e l'AI, perché il modo in cui

formuliamo e strutturiamo i prompt ha un impatto essenziale sui risultati ottenuti dai modelli di intelligenza artificiale.

La scelta delle parole, la struttura delle frasi e la chiarezza delle istruzioni influiscono sulla comprensione dell'AI e sulla qualità dell'output generato. Un prompt ben formulato favorisce una comunicazione fluida e precisa tra l'utente e l'AI.

"I modelli di AI sono strumenti potenti ma complessi, capaci di generare una vasta gamma di output", rileva Irene Di Deo. "Il prompt design ci consente di guidare l'AI verso risultati specifici e desiderati. Un prompt ben progettato può influenzare la risposta dell'AI in modo tale da ottenere risposte pertinenti, accurate e coerenti con le nostre aspettative". Il prompt design "punta quindi alla progettazione, perfezionamento e ottimizzazione delle richieste di input di un modello di intelligenza artificiale generativa, con l'obiettivo di guidarlo verso la realizzazione dell'output e del risultato desiderato".

Ecco alcuni esempi: specificare il tono (formale, informale, istituzionale, colloquiale) e il ruolo desiderato (come un poeta, un medico o un magistrato). Definire l'obiettivo e il contesto di riferimento; procurare dettagli utili,anche attraverso esempi e keyword da includere; specificare il formato e altri parametri, come la lunghezza dei testi e il numero di parole.

Gli autori di questa monografia

Stefano Casini. Giornalista specializzato nei settori dell'innovazione, delle imprese e delle tecnologie. Ha lavorato per Panorama Economy, Il Mondo, Italia Oggi, TgCom24, Gruppo Mediolanum, Università Iulm. Attualmente collabora con Guerre di Rete, Innovation Post, EconomyUp, Giornale di Brescia e altre redazioni.

Antonio Dini. Giornalista freelance si occupa di innovazione, tecnologia, economia e politica internazionale. Scrive o parla per Il Sole 24 Ore, L'Espresso, La Stampa, Repubblica, Wired Italia, Radio24, Radio Popolare e altre testate online e offline.

Carola Frediani. Ha lavorato per anni come giornalista occupandosi di sorveglianza, cybercrimine e cybersicurezza. Oggi è nel team di information security di una Ong internazionale per i diritti umani. Scrive la newsletter Guerre di Rete e ha fondato l'omonima associazione e il progetto editoriale GuerrediRete.it insieme ai partner di Cyber Saiyan.

Federica Meta. Giornalista professionista, dal 2007 redattrice di CorCom. È specializzata in economia digitale, con un focus sulla grande trasformazione del lavoro, sulla digitalizzazione della PA e sugli impatti dell'innovazione tecnologica sulla parità di genere e sui diritti.

Giuditta Mosca. Giornalista freelance, collabora con Italian Tech e Green&Blue (hub del Gruppo Gedi). Si occupa di IoT, big data, open source, web, internet, protocolli, AI, robotica, cybersicurezza ma anche di diritti umani.

Federico Nejrotti. Autore e co-fondatore di Ufficio Furore, studio di progettazione che crea cultura radicale. Appassionato da sempre di internet governance, è stato responsabile delle comunicazioni di cheFare e capo-redattore dell'edizione italiana di Motherboard.

Antonio Piemontese. Giornalista professionista freelance. Si occupa di clima, economia e innovazione per testate italiane e internazionali, prediligendo le storie al crocevia. I suoi lavori sono apparsi su Wired,National Geographic, RSI, Class Cnbc, Capital.

Andrea Signorelli. Giornalista, classe 1982, si occupa del rapporto tra nuove tecnologie, politica e società. Scrive per Italian Tech, Wired, Domani, Il Tascabile e altri. È autore di "Technosapiens: come l'essere umano si trasforma in macchina" (D Editore, 2021)