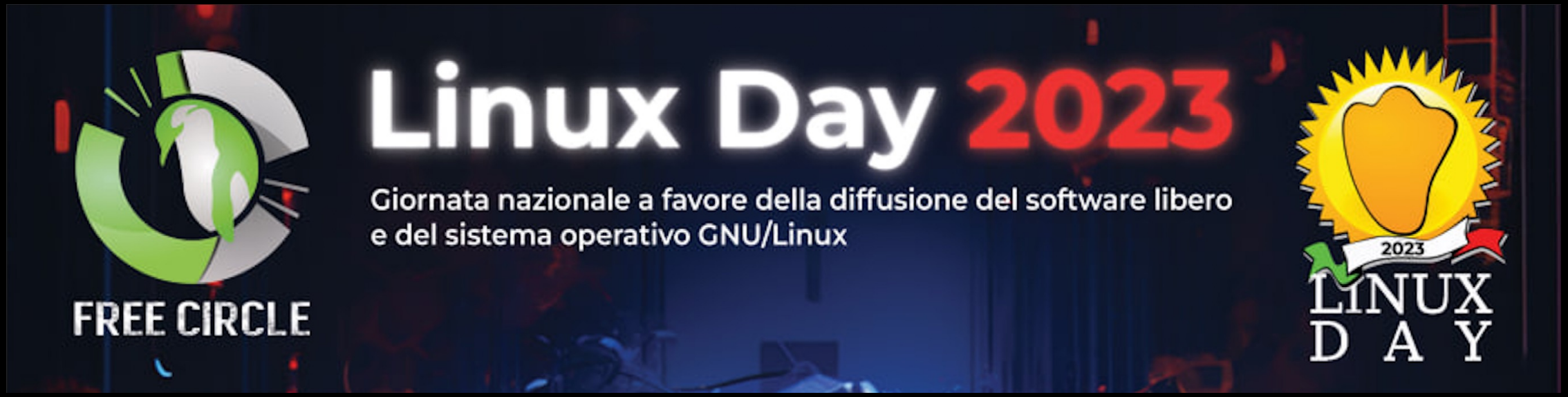


L'IMPATTO E I RISCHI NELL'USO DELL'IA GENERATIVA: VANTAGGI, ETICA E PRIVACY

Relatore: Prof. Vincenzo Agate

COS'È L'IA GENERATIVA?

- Branca dell'Intelligenza Artificiale per la generazione di contenuti nuovi che sembrano essere realizzati da umani…
- Testo
- Immagini
- Codice
- Audio
- Capacità di risposta ai "prompt" degli utenti mai raggiunte fin ora!

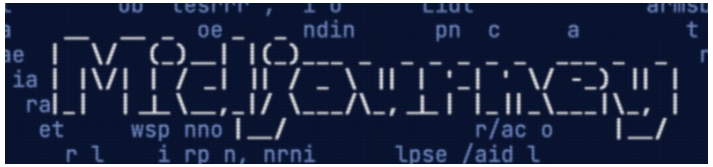
BING CHAT

Bard

ChatGPT

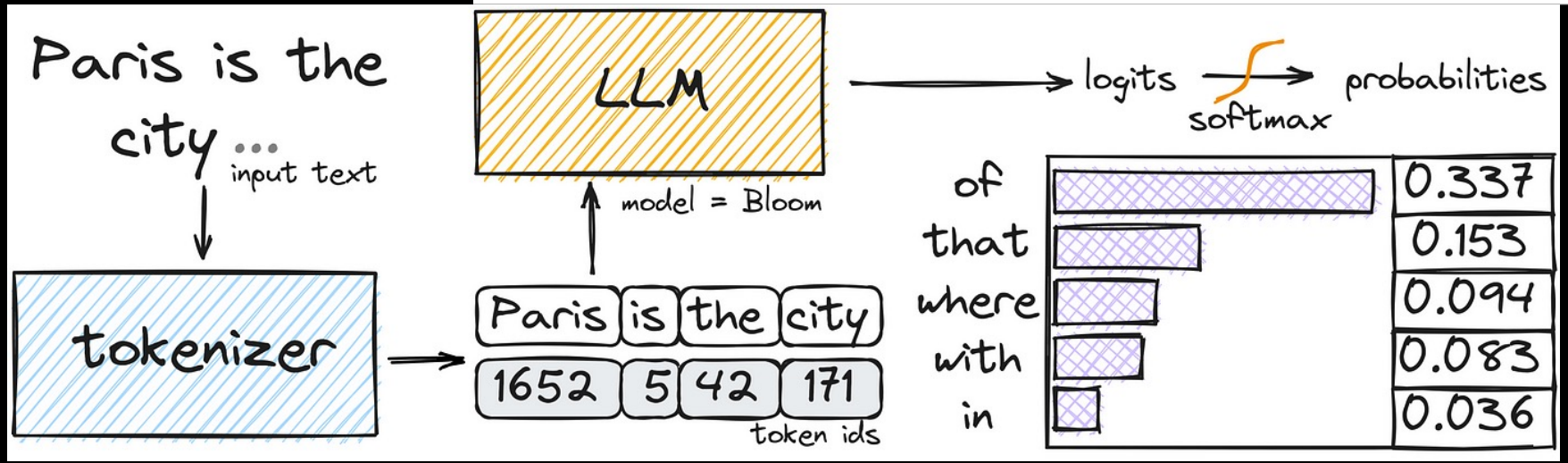
DALL·E 2

GitHub Copilot



TECNOLOGIA ALLA BASE DELLE IA GENERATIVE

- Natural Language Processing: "comprendere" il contenuto dei documenti e le loro sfumature contestuali
- Large Language Model (LLM)
- Trasformer Model (2017 – Attention is all you need!)
- Analisi di grandi quantità di dati
- GPU
- Grandi costi computazionali…
- Chi può permetterselo?



VANTAGGI DELL'IA GENERATIVA

- Miglioramento della produttività e automazione
- Creazione di contenuti creativi
- Personalizzazione dell'esperienza utente
- Esempi concreti di vantaggi in diversi settori:
  - chatbot sempre più sofisticati
  - design di prodotti e servizi
  - calo dei costi per le aziende
  - aumento delle vendite
- Pericoli per il fattore umano?

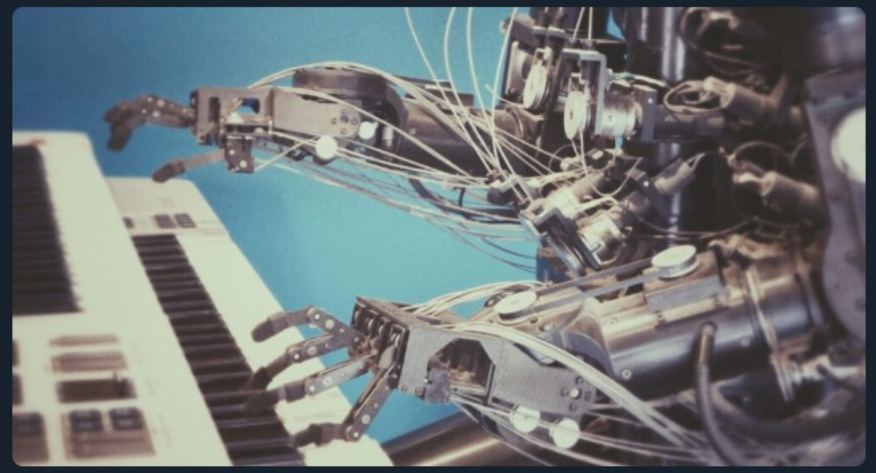
L’IMPATTO

- Scienza: l'IA generativa è utilizzata nella ricerca scientifica e medicina
- Arte: mostre e opere d'arte create con l'assistenza dell'IA generativa
- Marketing: campagne pubblicitarie mirate e personalizzate
- industria cinematografica, editoria e un’infinità di altri ambiti

Econopoly

Blog | Intelligenza artificiale, perché può essere un'arma a doppio taglio

È fondamentale esplorare gli impatti dell' intelligenza artificiale nei media, nella creatività, nelle interazioni sociali ed etiche



CYBERSECURITY360 | Cybersecurity Nazionale | Malware e attacchi | Norme e adeguamenti | S

TECNOLOGIA E SICUREZZA

Rischi di sicurezza dei sistemi di intelligenza artificiale generativa: l'impatto nel mondo cyber

Allarme Cisco Talos: i cyber criminali usano l'IA generativa per creare malware sempre più sofisticati

L'intelligenza artificiale permette di sviluppare attacchi phishing e social più sofisticati, come la creazione di deepfake incredibilmente realistici, siti web ingannevoli, campagne di disinformazione, profili fraudolenti sui social media e bot sempre più complessi

2 Ottobre 2023

RISCHI

Problemi etici e morali: dibattito sull'automazione di decisioni etiche e morali.

Può un IA giudicare in tribunale?

Valutare le condizioni mediche?

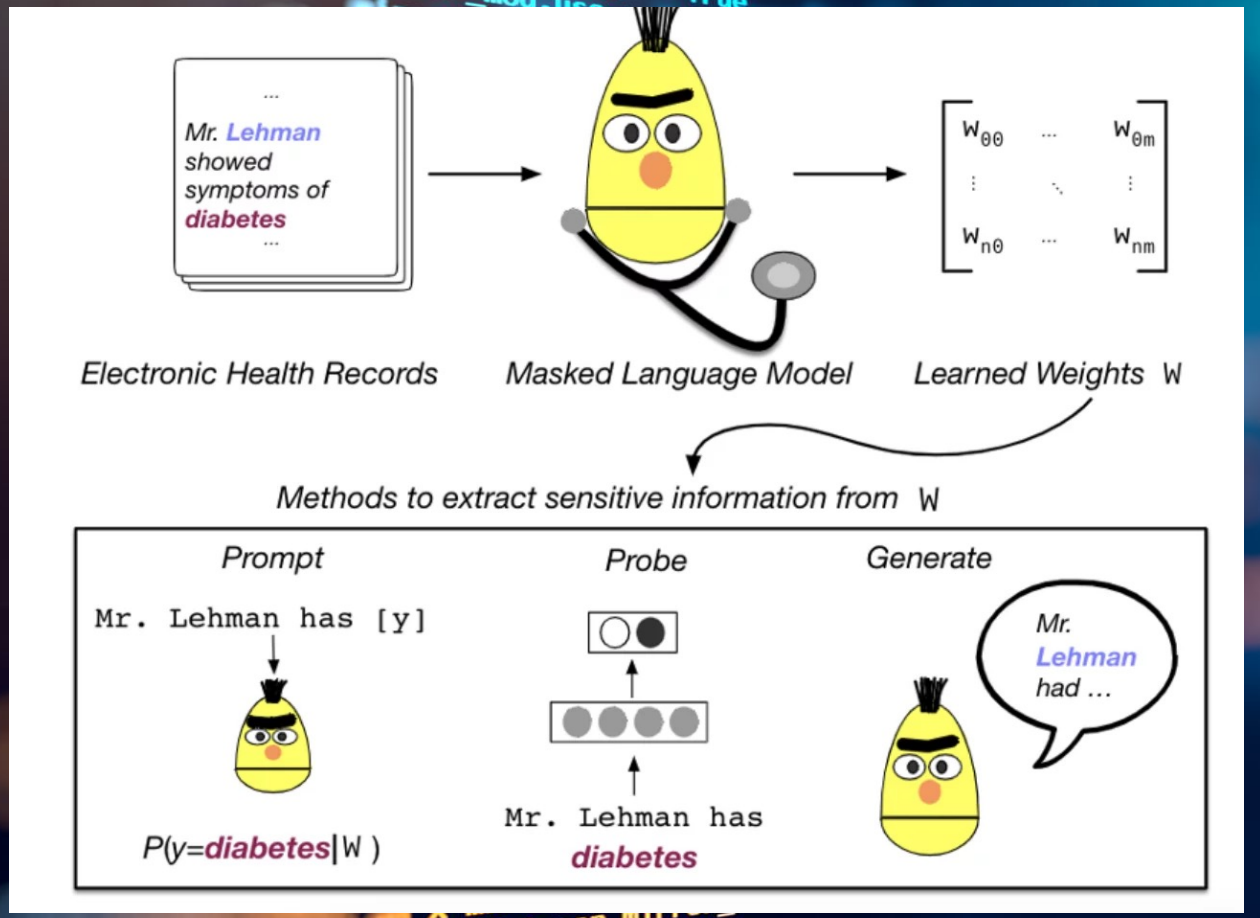
Dipendenza dall'IA generativa e perdita di abilità umane: come le persone possono diventare dipendenti dall'IA.

Implicazioni per l'occupazione e la disoccupazione tecnologica

Crescita della capacità di influenzare le persone da parte di pochi

PRIVACY E SICUREZZA

- Raccolta di dati sensibili e gestione della privacy
- Possibili abusi dell'IA generativa
- Strumenti e tecniche per proteggere la privacy
- In Carlini et al, 2019 si mostra come è possibile estrarre informazioni sensibili memorizzate da GPT-2



ETICA NELL'IA GENERATIVA

- Questioni di **bias** e **discriminazione** nell'IA generativa
- Principi etici per lo sviluppo e **l'uso responsabile: deep fake**
- Il ruolo delle linee guida etiche e organizzazioni

Il papa sui social con un piumino bianco oversize, milioni di visualizzazioni. Ma è una fake creata dall'intelligenza artificiale

IL GAP NORMATIVO

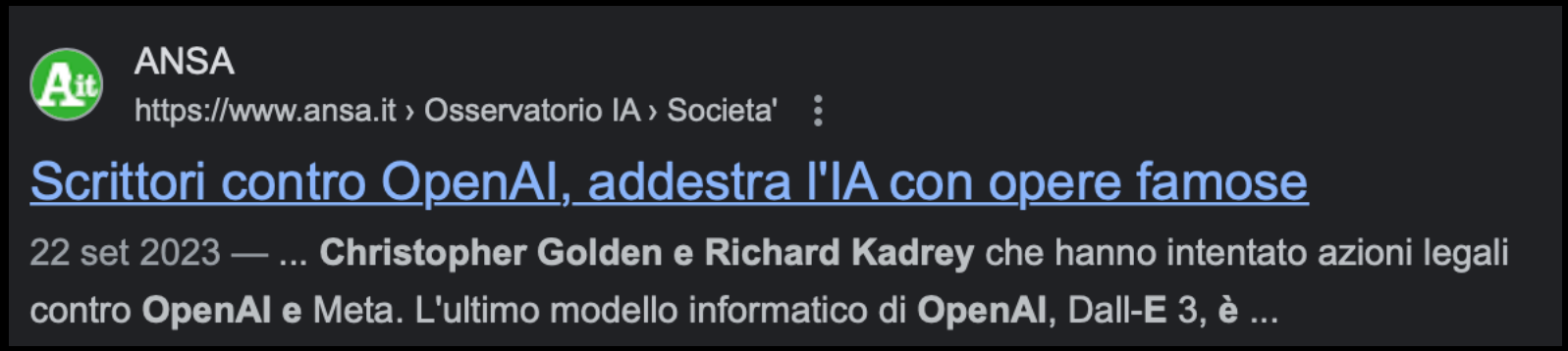
- **GDPR – Regolamento Generale sulla Protezione dei Dati Personali**
- **AI ACT** - Regolamento europeo sull'intelligenza artificiale adottato dall'Europarlamento
- Per chi sfrutta IA generative a rivelare che il contenuto è stato prodotto dall'IA;
- essere progettate in modo da non generare contenuti illegali;
- rendere noti i dati protetti da copyright che il modello utilizza per il suo addestramento.
- Maggiore trasparenza dei processi di addestramento!

CASI NOTI

- Le allegazioni di vari programmatori software vs GitHub, OpenAI et al. Gli attori hanno accusato software come Copilot – che soppiantano la scrittura manuale di pezzi di codice non particolarmente creativo o completano automaticamente operazioni di programmazione più rutinarie con compilazioni automatizzate – di aver copiato parti di codice da loro sviluppate (coding);

CASI NOTI

- Il caso Sarah Silverman, Christopher Golden e Richard Kadrey vs OpenAI Secondo le allegazioni attoree, OpenAI avrebbe allenato Chat GPT su enormi database contenenti libri (Books1 e Books 2), per un totale di titoli superiore a quello ottenibile lecitamente tramite scraping web (ad esempio dal sito del Gutemberg Project, che contiene solo testi i cui diritti d'autore sono esauriti). I software restituirebbero riassunti precisi dei libri degli attori, senza alcun riconoscimento dei loro diritti d'autore.



CASI NOTI

- Il **caso** Sarah Andersen, Kelly McKernan, and Karla Ortiz vs. Stability AI, **Midjourney**, e DeviantArt
  Le AI generative visive di questi fornitori sarebbero capaci di produrre opere “nello stile di” vari artisti viventi, ledendone i diritti d’autore.

IL PROGRESSO TECNOLOGICO

La pittura

La fotografia

IL PROGRESSO TECNOLOGICO

La radio

La televisione

IL PROGRESSO TECNOLOGICO

Le comunicazioni prima e dopo internet…

CONCLUSIONI

- Tra le **sfide** ci sono la necessità di usare **enormi risorse computazionali** ed **energetiche** per allenare e usare i modelli, la difficoltà di **garantire la qualità** e **l'affidabilità** dei testi generati, la **gestione dei dati sensibili** e la **protezione dei diritti d'autore**.
- Tra i **rischi** ci sono la possibilità che i modelli vengano usati per **scopi negativi o illegali**, come la diffusione di notizie false (fake news), la **manipolazione delle opinioni pubbliche** o la **violazione della privacy**.
- Inoltre, i modelli possono incorporare e amplificare i pregiudizi e le distorsioni presenti nei dati usati per allenarli, compromettendo la loro equità e neutralità.
- Possibili soluzioni: investire nella ricerca per mitigare gli effetti negativi ed educare all'utilizzo di questi nuovi strumenti.

GRAZIE PER L'ATTENZIONE