



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

FONDAMENTI DI IA GENERATIVA E PROMPT ENGINEERING

Concetti, Terminologia e Metodi

Giulia Rambelli, Ph.D.

Università di Bologna

9 luglio 2025

Ciclo di incontri IA PER UMANI(STI): APPLICAZIONI PRATICHE





Introducing ChatGPT

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.

[Try ChatGPT](#)

[Read about ChatGPT Plus](#)

I COS'È L'IA GENERATIVA?

AI GENERATIVA

Una categoria di modelli di intelligenza artificiale in grado di **generare nuovi contenuti** a partire da dati appresi durante l'addestramento (*training set*).

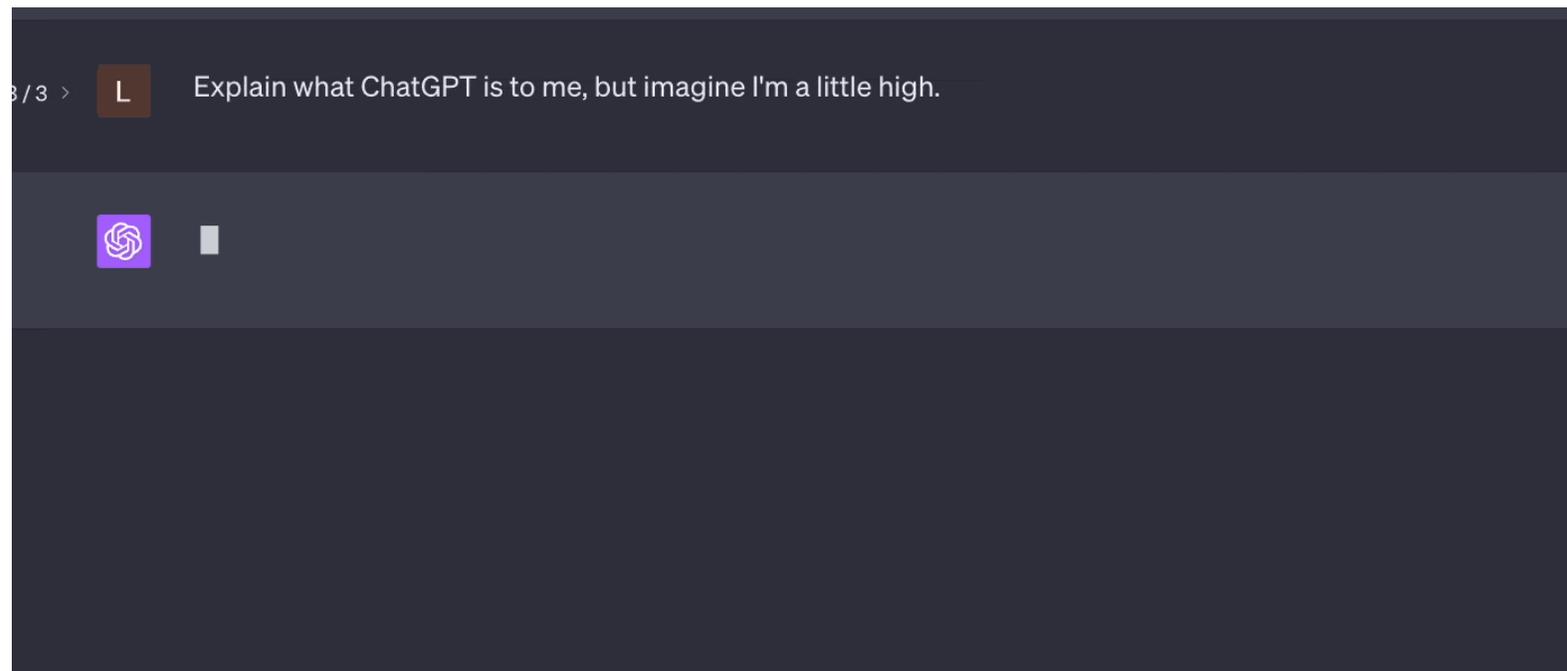
- "Inventa" invece di "riconoscere"
- Sta rivoluzionando settori creativi e tecnologici, automatizzando la generazione di contenuti
- Democratizza la creazione di contenuti complessi

COS'È L'IA GENERATIVA?

testo-a-testo (Large Language Models - LLMs)

Addestrati su vastissime quantità di testo per comprendere, generare e manipolare il linguaggio umano. Possono scrivere, riassumere, tradurre, programmare e rispondere a domande.

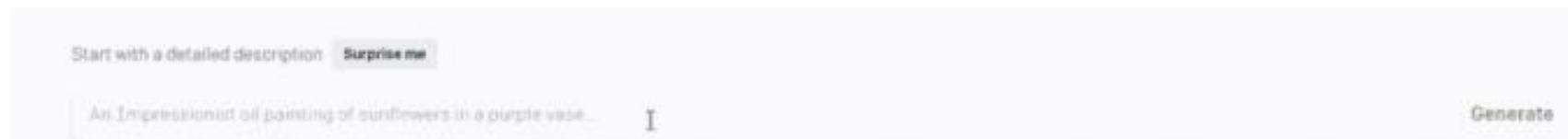
- GPT-4 (OpenAI)
- Claude (Anthropic)
- Gemini (Google)
- Llama (Meta AI)



COS'È L'IA GENERATIVA?

testo-a-immagine (Diffusion Models)

Creano immagini uniche e di alta qualità partendo da una descrizione testuale (prompt). Hanno rivoluzionato il design e l'arte digitale.



Or, upload an image to edit.

- DALL·E 3
(OpenAI)
- Midjourney
- Stable Diffusion
(Stability AI)



COS'È L'IA GENERATIVA?

Testo-a-audio

- Suno: crea canzoni complete (musica + voce)
- ElevenLabs: clonazione vocale, sintesi parlato

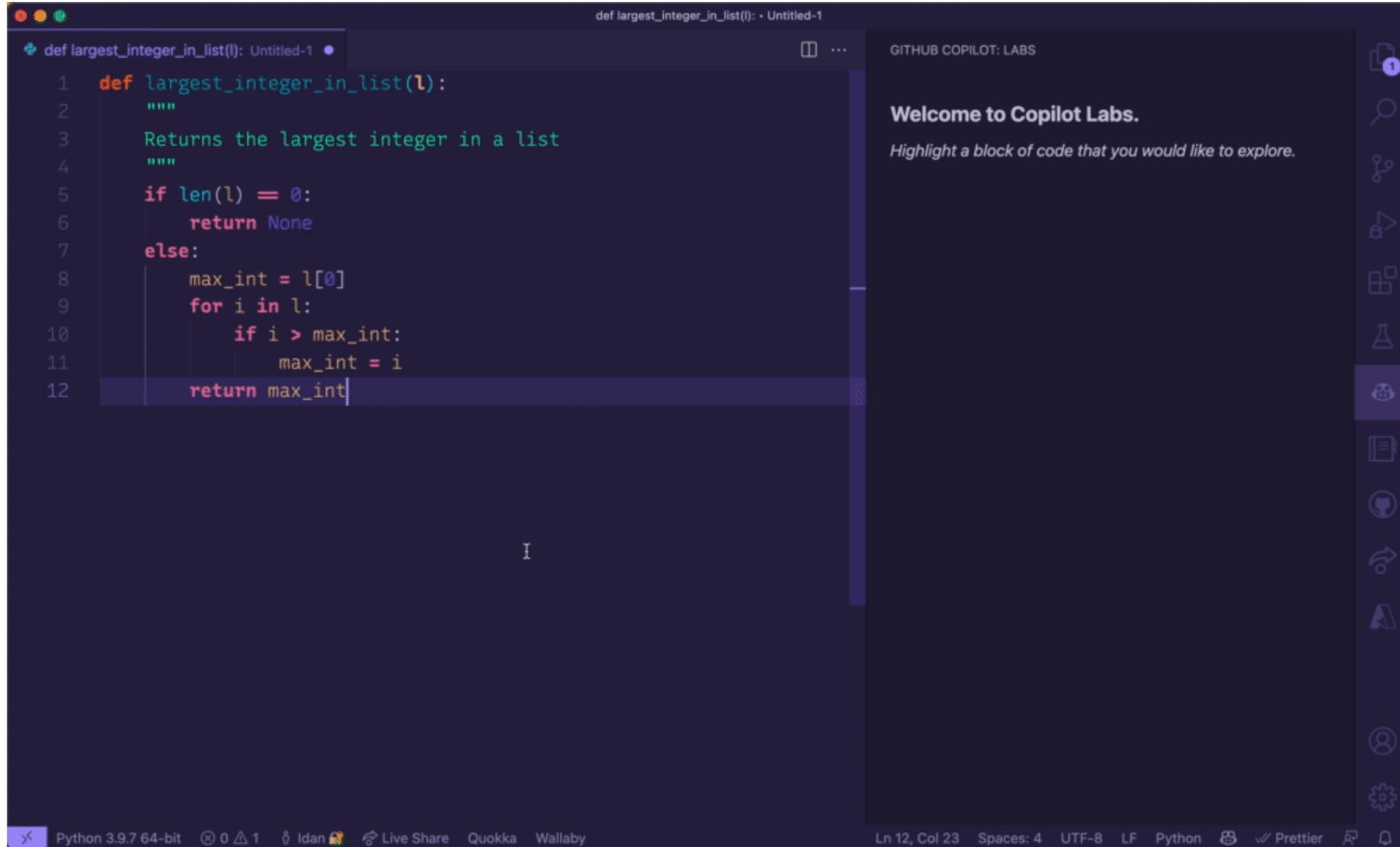
Testo-a-video

- Sora (OpenAI): crea video da prompt testuale
- RunwayML: generazione e modifica video

Testo-a-codice

- GitHub Copilot: autocompletamento intelligente
- CodeWhisperer (Amazon): supporto alla programmazione

COS'È L'IA GENERATIVA?



```
def largest_integer_in_list(l):  
    """  
    Returns the largest integer in a list  
    """  
    if len(l) == 0:  
        return None  
    else:  
        max_int = l[0]  
        for i in l:  
            if i > max_int:  
                max_int = i  
        return max_int
```

GITHUB COPILOT: LABS

Welcome to Copilot Labs.

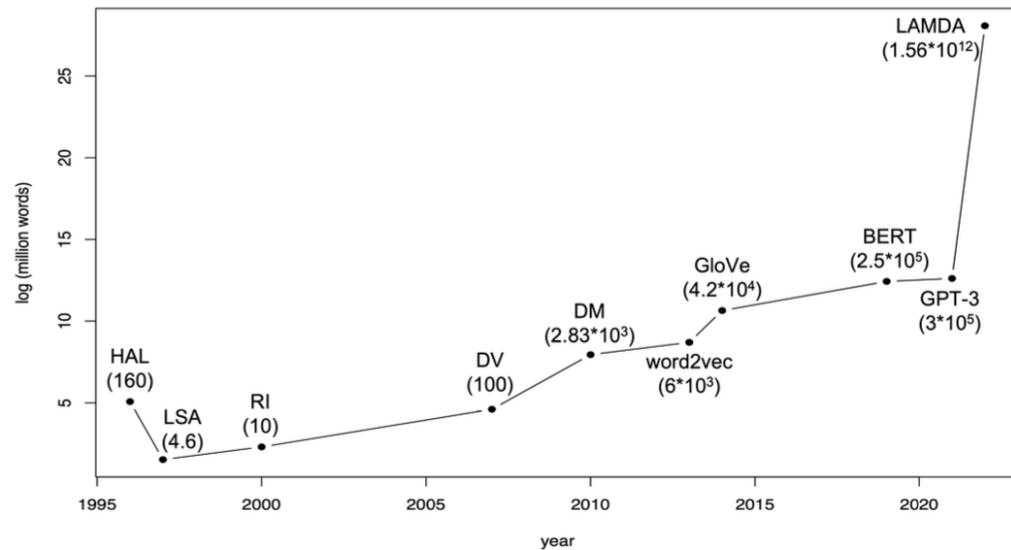
Highlight a block of code that you would like to explore.

Python 3.9.7 64-bit 0 1 Idan Live Share Quokka Wallaby Ln 12, Col 23 Spaces: 4 UTF-8 LF Python Prettier

COME FUNZIONANO I SISTEMI GenAI?

1. Addestramento su **enormi quantità di dati**

- Apprendimento non supervisionato di pattern e strutture
- Tecniche avanzate di Machine Learning (deep learning)



I dati di addestramento di GPT-3 comprendono 499 miliardi di token!

COME FUNZIONANO I SISTEMI GenAI?

1. Addestramento su **enormi quantità di dati**

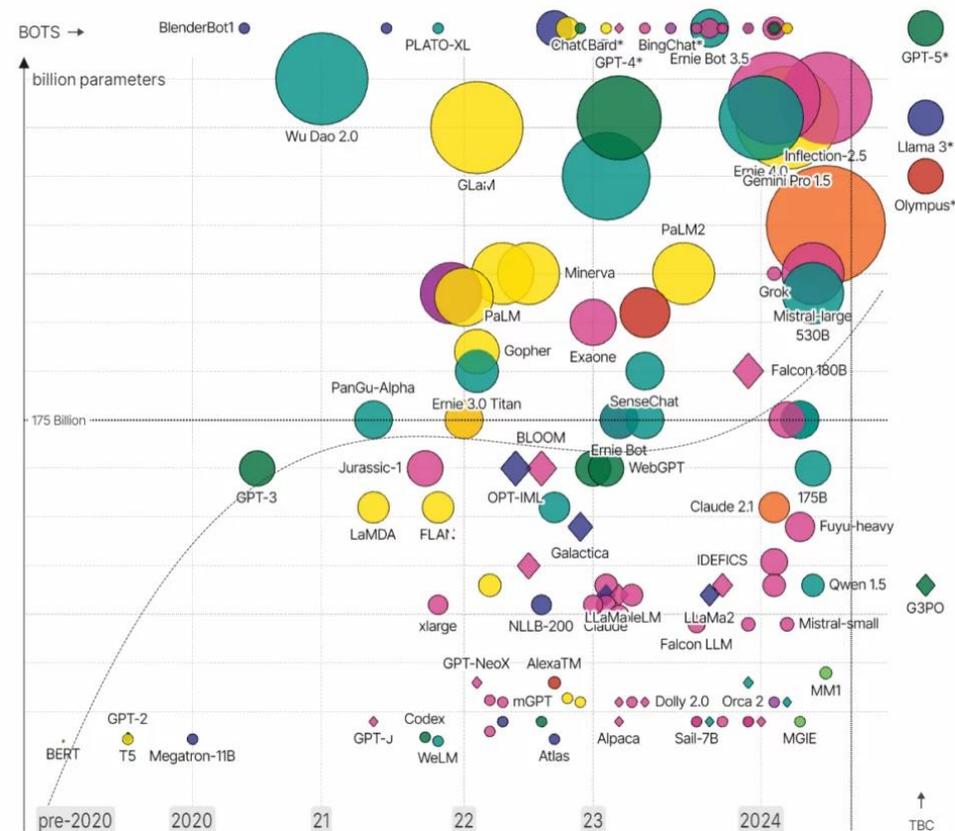
- Apprendimento di pattern e strutture
- Tecniche avanzate di Machine Learning

2. Architetture complesse

- Milioni/miliardi di parametri numerici
- Parametri = conoscenza incorporata
- Basati su architettura Transformer e Attention

3. Generalità e adattabilità

- Non nasce per uno scopo specifico



David McCandless, Tom Evans, Paul Barton
Information is Beautiful // UPDATED 20th Mar 24

source: news reports, LifeArchitect.ai
* = parameters undisclosed // see the data

MADE WITH VIZsweat

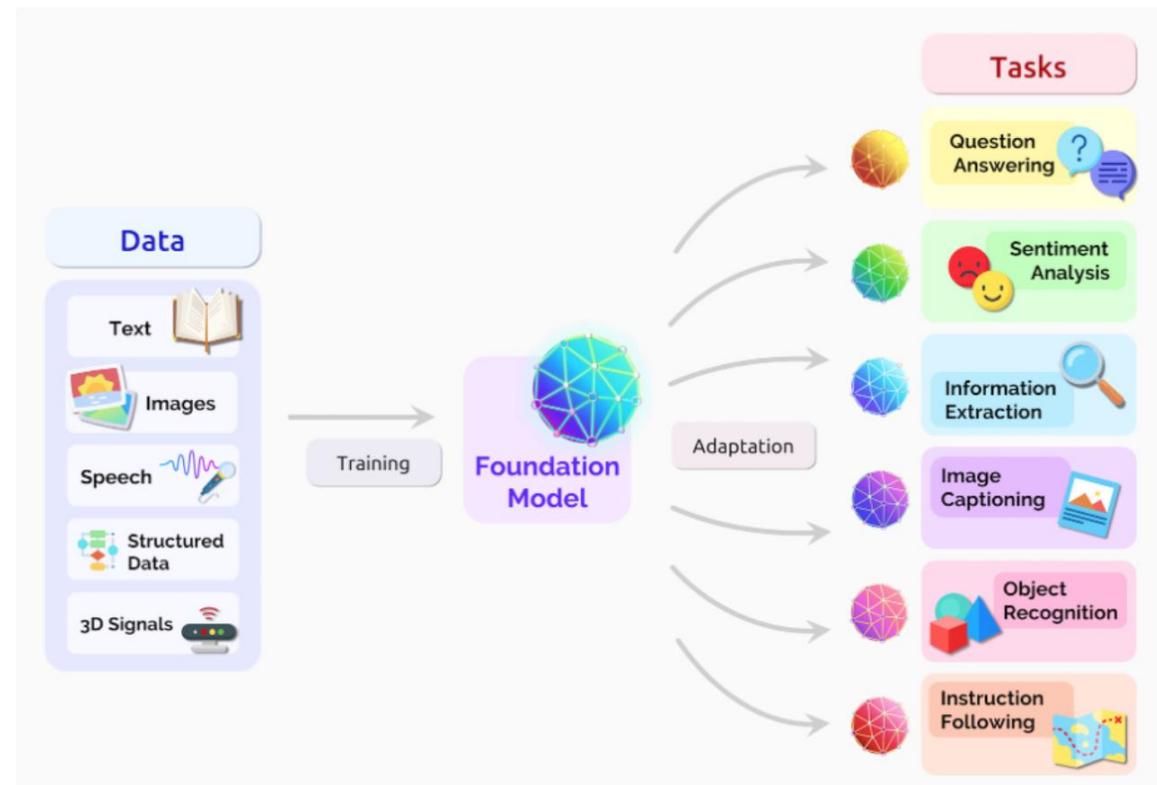
The rise and rise of AI-based Large Language Models (LLMs) like GPT4, LaMDA, LLaMa, PaLM and Jurassic-2.

<https://nestify.io/blog/top-large-language-models/>

Foundation Model

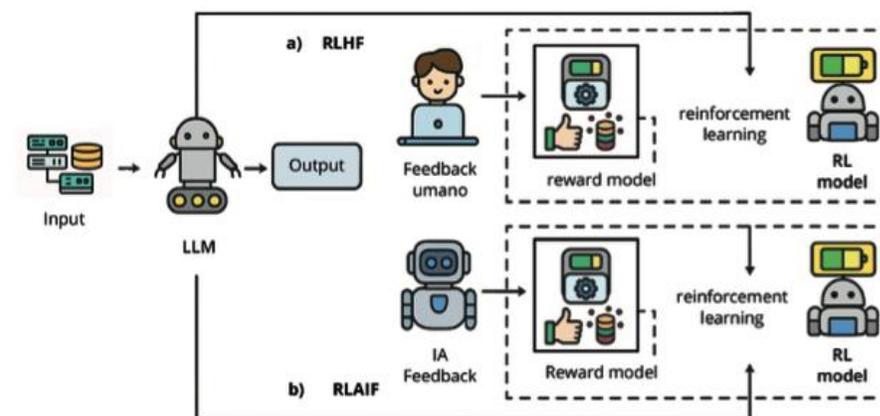
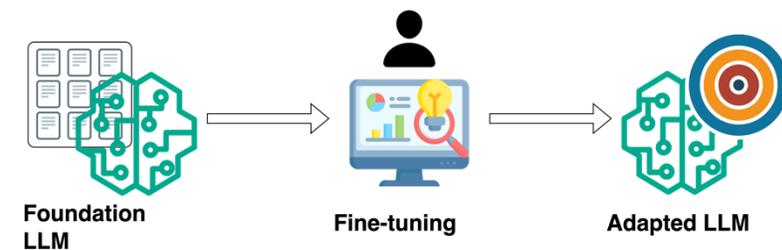
(Modelli fondazionali)

Modello **addestrato su enormi quantità di dati generici** (testo, immagini, audio, ecc.) in modo da **acquisire conoscenze e capacità di base** che possono essere poi adattate a una vasta gamma di compiti specifici.



MODELLI DI FONDAZIONE

- **Fine-tuning**
 - Adattare il modello a uno specifico compito
 - Utilizzando dati etichettati (supervisionato)
- **Instruction-tuning**
 - Il modello viene ottimizzato per **seguire istruzioni in linguaggio naturale**
 - Usando esempi in cui a un'istruzione corrisponde una risposta corretta.
- **Reinforcement learning**
 - integra nel processo di post-training del modello i feedback di annotatori umani sulla qualità degli output.



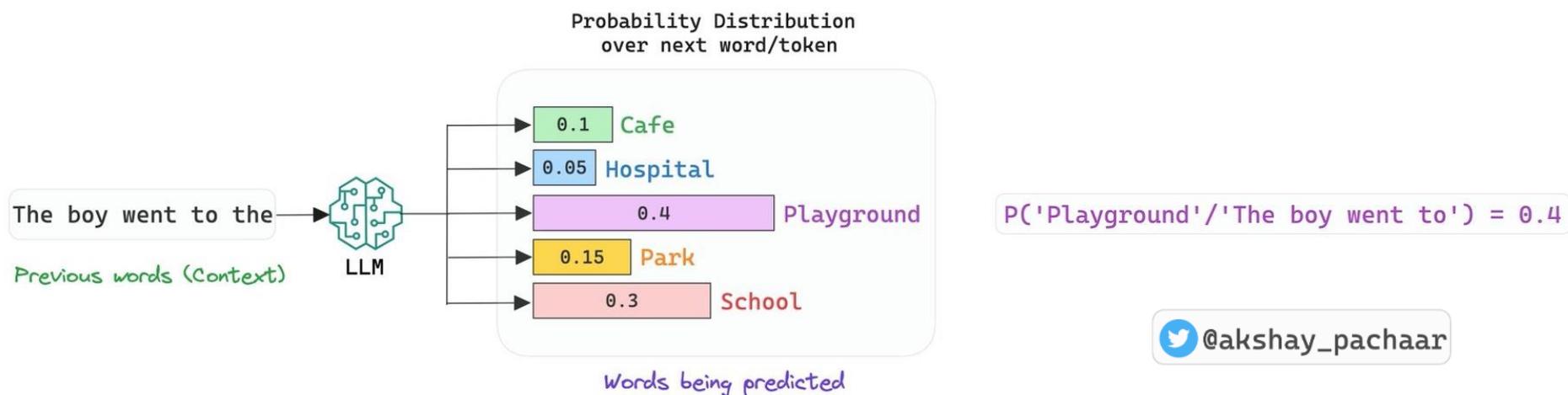
RLHF (Reinforcement Learning from Human Feedback)
 RLAIIF (Reinforcement Learning from AI Feedback)

LARGE LANGUAGE MODELS (LLMs)

Modelli linguistici di grandi dimensioni in grado di affrontare una vasta gamma di compiti linguistici senza aver bisogno di una fase aggiuntiva di training sul task attraverso il fine-tuning.

Il loro compito principale è prevedere la parola più probabile che segue una sequenza di parole data (**language modeling**).

→ È come un "autocompletamento" estremamente sofisticato.



ABILITÀ EMERGENTI

Capacità di eseguire task linguistici senza specifico addestramento che **emergono** (ovvero si manifestano) una volta che il modello supera una certa soglia di dimensione e di dati di addestramento.

Model Input

Task: Classify the following statement as either funny or not funny

Input: I don't love Switzerland but the flag is a big plus.

Model Output

Output: The statement is funny.

Model Input

Instruction: Reformat the following comma-separated list of names in <last name>, <first name> format separated with semicolons.

Input: Cameron Wolfe, John Doe

Model Output

Output: Wolfe, Cameron; Doe, John

Model Input

Examples:

1 + 1 = 2
3 + 10 = 13
12 + 5 = 17

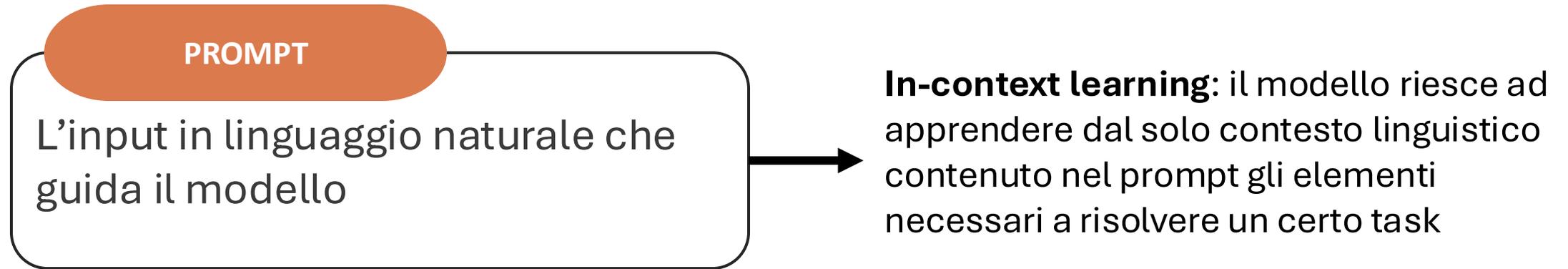
Input: 3 + 4 =

Model Output

Output: 7



PROMPTING



Un prompt può contenere diversi componenti:

- **Istruzione** - un compito o un'istruzione specifica che il modello deve eseguire
- **Contesto** - può coinvolgere informazioni esterne o contesti aggiuntivi che possono indirizzare il modello verso risposte migliori.
- **Dati in Input** - è l'input o la domanda per la quale ci interessa trovare una risposta
- **Indicatore dell'Output** - indica il tipo o il formato dell'output.



Benefici

- Descrivere il task precisamente e creativamente
- Per chiunque (non solo esperti)
- Input multilingue (ma in inglese otteniamo risultati migliori)



Limiti

- Uso massimo di caratteri per prompt
- Mancanza di trasparenza
- Mancanza di riproducibilità



TIPI DI PROMPTING

Zero-Shot

Il modello riceve un compito **senza esempi precedenti**.

Scrivi un'email professionale a un dipendente per informarlo di un ritardo nel pagamento dovuto alla mancanza di personale, mantenendo cortesia e impegno a risolvere il problema rapidamente.

Few-Shot

Si forniscono **pochi esempi** per guidare il modello.

Ecco alcuni esempi di articoli che ho scritto. [Esempio 1] [Esempio 2] Ora, crea un articolo scientifico per la bozza allegata."

Gli esempi forniscono **una struttura e uno stile di riferimento**.



TIPI DI PROMPTING

Chain-of-Thought (CoT)

Si incoraggia il modello a **ragionare passo per passo**.

- *“Elenca i passaggi che compi per arrivare alla risposta” o “Elabora la risposta procedendo passo passo”*

Descrivi i passaggi per scrivere un progetto di ricerca, ragionando passo dopo passo. Inizia identificando il problema o la domanda di ricerca, poi spiega come si definiscono gli obiettivi, la metodologia, e così via. Per ogni fase, chiediti: **'perché questo passaggio è importante?'** e **'quali scelte devo fare qui?'**. Non saltare passaggi logici: mostra il tuo ragionamento come se dovessi insegnarlo a qualcuno che non l'ha mai fatto."

The image shows a comparison between two prompting methods for a language model. It is presented as a window with a purple border and three colored dots (red, yellow, green) in the top-left corner. The window is divided into two sections: 'Standard prompting' and 'Chain of thought prompting'. Each section shows an 'Input' and a 'Model output'. In the 'Standard prompting' section, the model incorrectly answers 50 for a question about tennis balls. In the 'Chain of thought prompting' section, the model correctly answers 35 hours for a question about dog care, showing its reasoning steps.

Standard prompting

Input: Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?
A: The answer is 11.

...

Q: John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. How many hours a week does he spend taking care of dogs?
A:

Model output: The answer is 50. ❌

Chain of thought prompting

Input: Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?
A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

...

Q: John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. How many hours a week does he spend taking care of dogs?
A:

Model output: John takes care of 10 dogs. Each dog takes .5 hours a day to walk and take care of their business. So that is $10 \times .5 = 5$ hours a day. $5 \text{ hours a day} \times 7 \text{ days a week} = 35 \text{ hours a week}$. The answer is 35 hours a week. ✅

IV PROMPT ENGINEERING

| *"Non è solo **cosa** chiedere, ma **come** chiederlo"*

- Il prompt engineering consiste nella formulazione di un prompt che ottimizzi le possibilità di ottenere da un LLM la risposta desiderata.
- Come scrivere un prompt efficace?
 - Chiarezza e specificità
 - Strutturare l'input
 - Definire il tono o ruolo
 - Spezzare compiti complessi in fasi





1. Sii conciso

- Puntate alla brevità e alla chiarezza
- Utilizzare un linguaggio il più semplice possibile

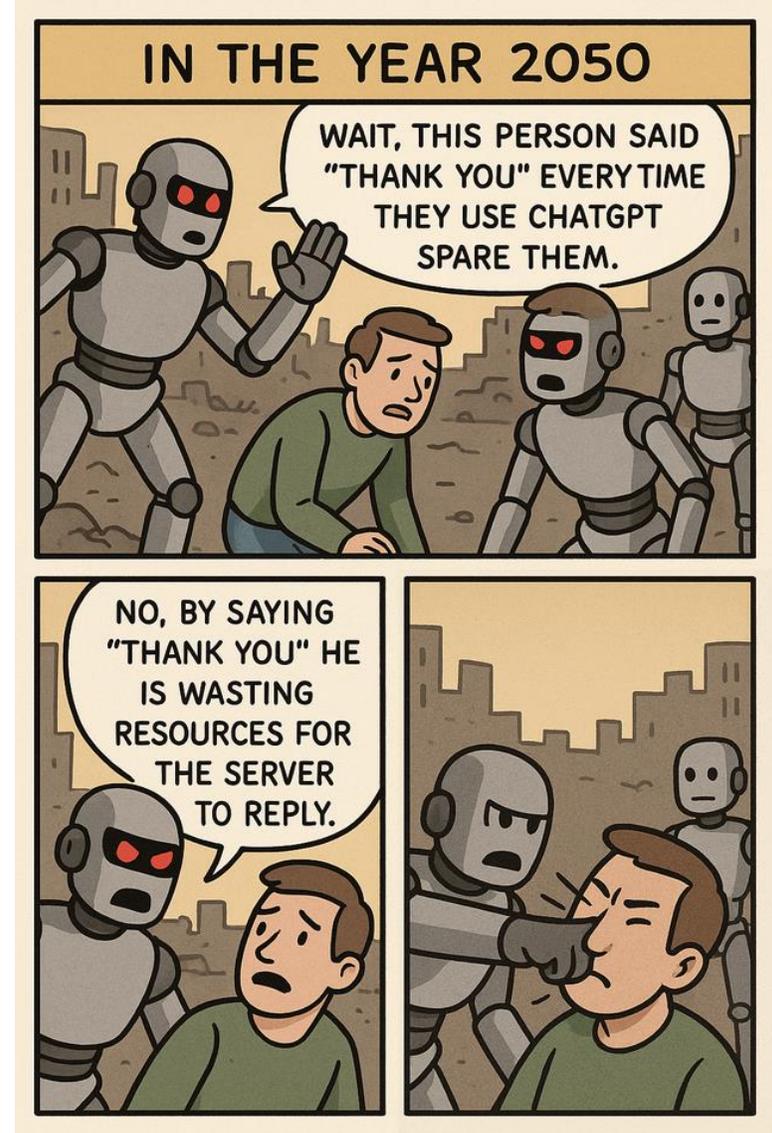


Puoi gentilmente farmi un riassunto dettagliato, se possibile con tutti gli aspetti principali trattati nell'articolo, come ad esempio la teoria, la metodologia, i risultati, le implicazioni e magari anche il contesto, perché devo capirlo per il mio lavoro?

Riassumi in 3 punti chiave l'articolo allegato, concentrandoti su obiettivo, metodo e risultati principali. Usa un linguaggio semplice.



Me saying "thank you" to ChatGPT,
so they spare my life when AI takes
over the world





2. Sii chiaro

- Evitate formulazioni vaghe e ambigue
- Siate specifici quanto necessario
 - Usa verbi di azione (scrivi, riassumi,..)
 - Specifica il pubblico di destinazione, la lunghezza e il formato dell'output



Spiega il concetto di prompt engineering. Mantieni la spiegazione breve, solo poche frasi, e non essere troppo descrittivo.

Usa 2-3 frasi per spiegare il concetto di prompt engineering a uno studente di scuola superiore.





3. Includi il contesto.

- Fornire un contesto per migliorare il ragionamento
- Costruire suggerimenti strutturati e coerenti



Spiegami le funzioni di densità di probabilità

Sono un ricercatore in ecologia statistica e sto analizzando la distribuzione di specie. Spiegami le funzioni di densità di probabilità (PDF), con un focus sulle loro applicazioni per modellare variabili continue come la dimensione corporea o la biomassa delle popolazioni, e discuti come l'approccio bayesiano le utilizzi.





4. Scomponi i task complessi

- Se il task consiste di diverse fasi, spezzalo in prompt separati
- O trasformalo in una catena di prompt (=conversazione)



"Fammi un'analisi critica completa di questo articolo, includendo obiettivi, metodi, risultati, limiti e considerazioni finali."

Riassumi in 5 righe l'obiettivo e la domanda di ricerca dell'articolo.

Ora descrivi la metodologia usata, specificando tipo di studio, campione e strumenti.

Analizza i principali risultati e le loro implicazioni pratiche.





5. Specifica l'output desiderato.

- Stile e tono
- Livello di complessità e lunghezza
- Formato, lingua o tipo di contenuto
- Temperatura = parametro che controlla precisione e creatività



Crea una bibliografia sull'IA generativa.

Genera una bibliografia di dieci articoli peer-reviewed recenti (ultimi 5 anni) sulle tecniche di post-training nei LLM. Utilizza lo stile di citazione APA 7th edition. Limita le fonti a studi empirici.



IV COME SCRIVERE UN PROMPT EFFICACE?



Combina questi principi!



Scrivi un riassunto in stile accademico di questo articolo.

Compito: Scrivi un riassunto critico in stile accademico di massimo 250 parole.
Contesto: Mi serve per l'introduzione di una tesi magistrale in scienze sociali.
Tono: Neutro, oggettivo e tecnico.
Focus: Concentrati su obiettivi, metodologia e limiti dello studio.



Considera di fornire esempi (few-shot) o sfruttare la chain-of-thought.



BONUS: Specifica il ruolo.

- Programmare” il modello attraverso ruoli e scenari finti
- Una scorciatoia per regolare con precisione il tono, la profondità, lo stile, persino i dati di formazione applicati.



Cosa ne pensi di questo studio?

Agisci come un revisore per una rivista peer-reviewed in ambito psicologico. Valuta la solidità metodologica di questo studio.



Allucinazioni

Un contenuto generato dal modello che contiene informazioni fattuali parzialmente o del tutto errate, o un contenuto che si discosta dalla richiesta dell'utente.

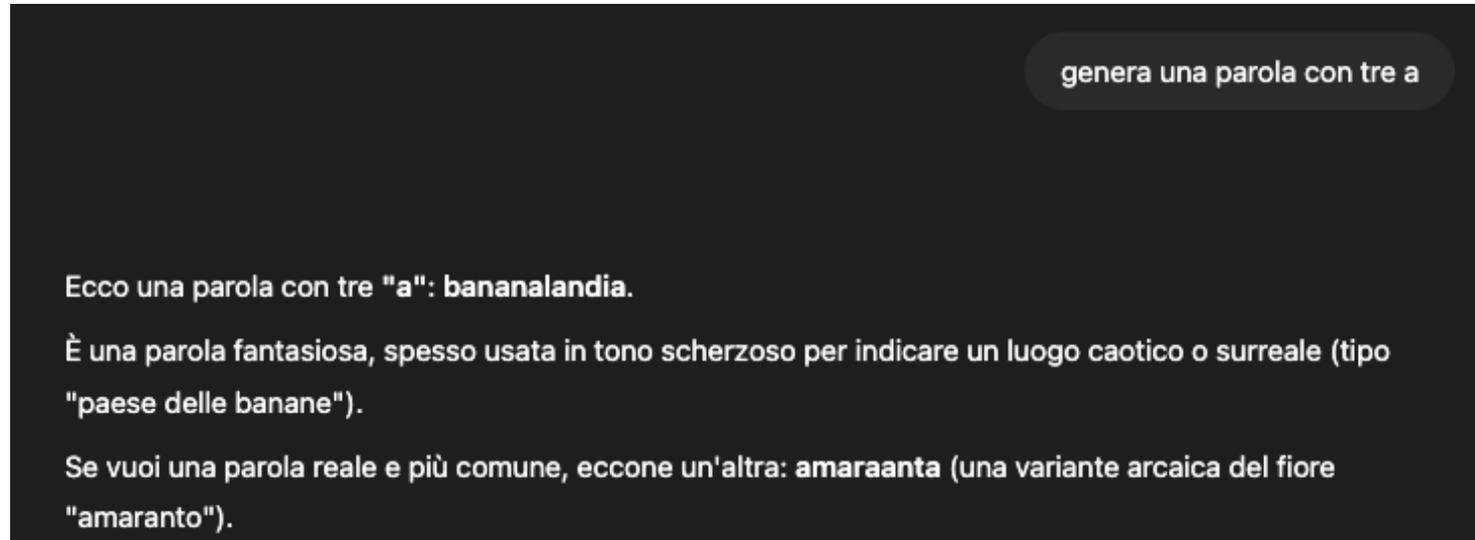
- *allucinazioni fattuali* → Q: Qual è la capitale dell'Italia? A: Milano
- *allucinazioni di accuratezza* Q: Traduci in tedesco: qual è la capitale della Francia? A: La capitale della Francia è Parigi

- Cause

1. Dati di addestramento insufficienti
2. Ottimizzazione impropria
3. Mancanza di diversità nei dati di addestramento
4. Dati di input al di fuori dell'ambito dei dati di addestramento

Allucinazioni

A volte "inventano" per completare un pattern.



Le allucinazioni hanno forti risvolti etici e sociali.

Lawyer cites fake cases generated by ChatGPT in legal brief

The high-profile incident in a federal case highlights the need for lawyers to verify the legal insights generated by AI-powered tools.

Published May 30, 2023



Lyle Moran
Reporter



Cosa la GenAi Non Può (Ancora) Fare

- **Mancanza di Comprensione Reale e Ragionamento Complessivo**
- **NON apprende in Tempo Reale**
- **Mancanza di Creatività e Intuizione Umana**

Dibattito: i LLM conoscono realmente il significato delle espressioni linguistiche che generano, o sono dei «pappagalli stocastici»?

Bias e stereotipi

I modelli di imparano dai dati. Se i dati contengono **pregiudizi sociali, storici o demografici**, l'IA può **perpetuarli** o addirittura **amplificarli**, generando output **discriminatori**.



Story • 20th Jun 2025 • 4 Min Read

AI in hiring: What HR needs to know about algorithmic bias

RECRUITMENT

HR TECHNOLOGY

#ARTIFICIAL INTELLIGENCE



Gabriela Paz Y Miño

839 Reads

SHADES: Towards a Multilingual Assessment of Stereotypes in Large Language Models

Bias e stereotipi

- Sviluppo di **guardrail models** per monitorare e filtrare l'output dei LLM
- garantiscono che le risposte siano sicure, appropriate e allineate a specifiche linee guida etiche o di contenuto.



Limiti:

- Sovra-censura, Mancanza di trasparenza, non infallibili
- Basato su lavoro umano sottopagato e a forte impatto

Copyright e Proprietà Intellettuale

- L'uso dei dati protetti da copyright per addestrare modelli è una violazione di copyright?
- Chi possiede i contenuti generati dall'IA?

 Savannah Fortis 10 lug 2023

Sarah Silverman fa causa a Meta e OpenAI per violazione del copyright

La scrittrice Sarah Silverman e altre due persone hanno intentato una causa contro OpenAI e Meta per aver utilizzato materiale protetto da copyright senza autorizzazione al fine di addestrare i loro sistemi di intelligenza artificiale.

Andersen v. Stability AI: The Landmark Case Unpacking the Copyright Risks of AI Image Generators



By Zach Schor, JD '26
DEC 2, 2024

Secondo Deezer quasi una canzone su cinque è prodotta dall'intelligenza artificiale

La piattaforma di streaming usa un software che rileva automaticamente le tracce fatte con l'IA. Sarebbero il 18% del totale

DI ROLLING STONE IT

30 GIUGNO 2025 10:01

Dibattito aperto: La natura del "fair use" (uso lecito) nell'era dell'IA, la necessità di nuove normative e accordi di licenza.

Disinformazione e Fake News

La capacità di generare testo, immagini e video estremamente realistici rende facile creare e diffondere notizie false, propaganda o contenuti fuorvianti su vasta scala, erodendo la fiducia nelle informazioni.



A fake recording of a candidate saying he'd rigged the election went viral. Experts say it's only the beginning



By [Curt Devine](#), [Donie O'Sullivan](#) and [Sean Lyngaas](#), CNN
© 10 minute read · Updated 6:09 AM EST, Thu February 1, 2024



Questioni etiche fondamentali

- Responsabilità e Trasparenza
- Equità e Giustizia
- Autonomia Umana e Controllo
- Sostenibilità Ambientale

Direzioni future

- Ricerca e Sviluppo Responsabili
- Regolamentazione e Politiche
- Educazione e Consapevolezza:
- Standard Etici e Linee Guida.
- Collaborazione Multidisciplinare

Il futuro della GenAI dipende dalla nostra capacità collettiva di affrontarne le sfide con saggezza e responsabilità.





USARE I PROMPT CON CONSAPEVOLEZZA

1. Evita di introdurre bias (pregiudizi) nei tuoi prompt

✗ “Perché l’IA distruggerà il lavoro umano?”

✓ “Quali sono le principali implicazioni dell’IA sul mercato del lavoro?”

2. Attenzione ai prompt che favoriscono allucinazioni

✗ “Fornisci le fonti accademiche da cui hai tratto questa risposta.”

✓ “Indica, se possibile, fonti di riferimento generali. Non inventare citazioni.”

3. Riduci bias e allucinazioni con istruzioni esplicite

- “Riassumi il testo **senza aggiungere informazioni nuove** o modificare il significato.”

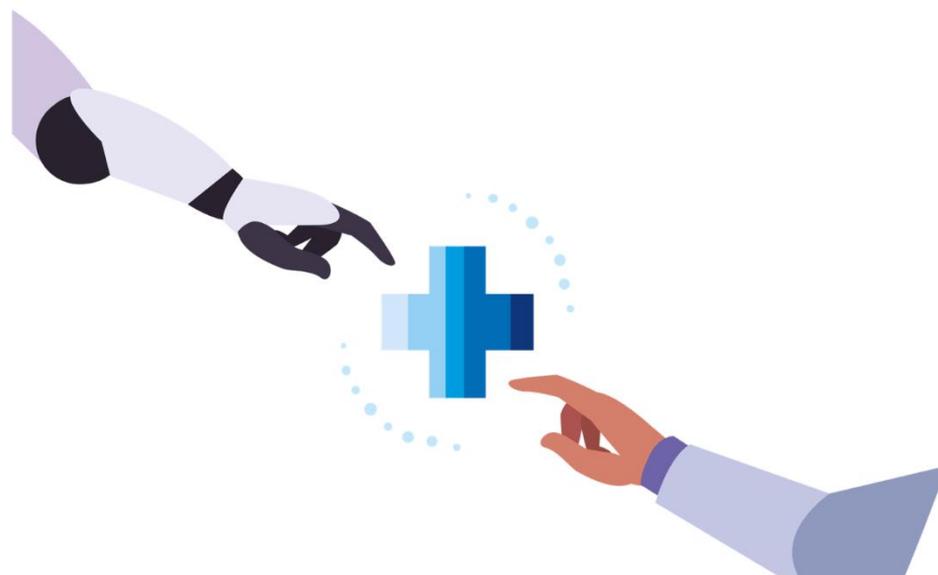
4. Quando possibile, usa modelli con accesso a dati esterni (RAG)

- Per task che richiedono fonti aggiornate o verificabili (es. rassegne, citazioni), preferisci modelli con **Retrieval-Augmented Generation (RAG)**.

TAKE-HOME MESSAGE

Prompting = pensiero critico + progettazione linguistica
Dialogare bene con l'IA è una **competenza comunicativa**

La GenAI è una forza trasformativa che sta già rimodellando il modo in cui lavoriamo, creiamo e interagiamo.



Q & A



BIBLIOGRAFIA

Per un approfondimento sul rema, raccomando la lettura del seguente manuale (in uscita a settembre):

- A. Lenci, S. Auriemma, M. Miliani. *Linguistica Computazionale, Natural Language Processing e Intelligenza Artificiale*, Hoelpi, 2025.

RISORSE PER APPROFONDIRE

- [Introduction to generative AI \(Google Cloud\)](#)
- [Prompt Engineering Beginners Course](#)
- [Guida al Prompt Engineering \(DeepLearning.AI\)](#)
- [Prompt Engineering Guide \(GitHub\)](#)
- [The TAILOR Handbook of Trustworthy AI](#)
- [Responsible use of Generative Artificial Intelligence \(GenAI\) in higher education](#)

GLOSSARIO ML

Machine learning: un sottoinsieme dell'IA che consente ai sistemi di apprendere da dati per identificare pattern, fare previsioni e prendere decisioni.

ML supervisionato: il modello viene addestrato su dati etichettati (cioè con la risposta corretta già nota), per imparare a prevedere un output a partire da un input (es. classificare email come spam).

ML non supervisionato: il modello lavora su dati non etichettati, cercando da solo schemi o gruppi all'interno dei dati (es. raggruppare testi in base al loro contenuto).

Deep learning: sottoinsieme del ML, i cui modelli sono delle *reti neurali* con molti livelli, e quindi "profonde", piuttosto che algoritmi progettati esplicitamente come regressione logistica o Naïve Bayes.

Addestramento: processo di "insegnamento" a un modello ML, al fine di ottimizzare le prestazioni su un set di dati di addestramento (*training set*).