

Che cos'è un LLM?

Dalla generazione al funzionamento



Fonte: la Repubblica, 13 ottobre 2026

Obiettivi della lezione

- Distinguere un LLM da un chatbot e da un'applicazione
- Capire come un modello linguistico sceglie una parola dopo l'altra
- Conoscere i concetti di token e probabilità; distinguere l'addestramento di un modello dal suo uso
- Capire perché lo stesso prompt produce risultati diversi
- Capire perché testi e immagini generati tendono a somigliarsi
- Distinguere la qualità del singolo testo dalla diversità dell'insieme dei testi



LA DOMANDA DI OGGI

**Quando ChatGPT scrive,
sta cercando, copiando o
generando?**

Come si svolge la lezione



1. Sperimentare

Un'immagine e un testo promozionale, con e senza IA



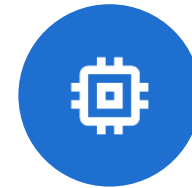
2. Osservare

Che cosa ricorre nelle immagini e nei testi della classe



3. Confrontare

Che cosa dice la ricerca su qualità e diversità



4. Capire

Token, probabilità, generazione, addestramento

PARTE 1

Due esperimenti

Esperimento 1: una locandina per la festa dell'università

1. Gruppo 1 usa ChatGPT, gruppo 2 usa Gemini (i gruppi si formano in aula)
2. Aprite una chat temporanea; se non c'è, una finestra in incognito in cui accedete al vostro account
3. Copiate il prompt senza modificarlo e senza indicazioni di stile
4. Generate una sola locandina
5. Salvatela e consegnatela su Moodle

PROMPT

Genera una locandina per una festa dell'Università di Trieste a tema mare e barche a vela.



Win + Shift + S cattura un'area dello schermo, se la locandina non si scarica

Esperimento 2: un testo promozionale per l'università

Consegna comune 150-180 parole per studenti e studentesse delle superiori che stanno scegliendo dove iscriversi. Testo attrattivo ma credibile, basato solo sulle informazioni della scheda. Schede e atenei assegnati: Moodle, Lezione 2.



Gruppo A · senza IA

- Leggete la scheda sull'Università di Trieste
- Scrivete il testo da soli, senza strumenti di IA
- Consegnate su Moodle



Gruppo B · con IA

- Usate ChatGPT con il prompt comune; incollate la scheda sotto il prompt
- Due testi: Trieste e l'ateneo assegnato
- Una nuova conversazione per ogni testo
- Consegnate i testi senza modificarli

Gruppo B: il prompt

PROMPT · IDENTICO PER TUTTI

Scrivi un testo promozionale di 150-180 parole rivolto a studenti e studentesse delle scuole superiori che stanno scegliendo dove iscriversi all'università. Presenta l'università in modo attrattivo ma credibile. Usa esclusivamente le informazioni contenute nella scheda fornita.

Ctrl + Shift + O apre una nuova chat in ChatGPT: un testo, una conversazione

Gli atenei

Università di Trieste

tutti

Università di Udine

assegnato

Università Ca' Foscari Venezia

assegnato

Università di Padova

assegnato

PARTE 2

Che cosa abbiamo ottenuto

Un caso reale: le locandine delle sagre



Ricostruzione schematica dei tratti descritti negli articoli



«Le AI si sono prese anche le sagre»

Pietro Minto, Il Post, 26 luglio 2026

La pagina SagreFVG sul «preoccupante trend» delle locandine generate:
«erano più belle quelle brutte».

Tratti ricorrenti delle locandine generate

- Sfondo nero, grandi scritte con effetto pennello, aspetto patinato, elementi ordinati e allineati (Il Post)
- Colori saturi, sfondi scuri, luci calde, cibo perfetto (Riviera Time)
- Le locandine fatte a mano erano brutte in modi diversi; quelle generate sono «tutte uguali, indipendentemente dal paese»

Leggi gli articoli: [Il Post, 26 luglio 2026](#) · [Riviera Time, 31 luglio 2026](#)

Le vostre locandine e quelle delle sagre

1

Quali tratti delle locandine delle sagre ritrovate nelle vostre?

- Sfondo scuro o tramonto
- Colori saturi e luci calde
- Aspetto patinato, «da pubblicità»
- Scritte con effetti luminosi o a pennello
- Composizione ordinata e simmetrica

2

Che cosa, nelle vostre locandine, potrebbe esistere solo a Trieste?

I test: un confronto cieco

A

B

C

D

E

**Testi anonimi su Moodle,
Lezione 2**

Leggeteli tutti prima di
rispondere

1

Quali test sono stati scritti con l'IA?

SCELTA MULTIPLA, PIÙ RISPOSTE

2

Su quali indizi vi basate?

RISPOSTA APERTA

3

**Quale test è più efficace come testo
promozionale?**

SCELTA SINGOLA

4

**Quali test potrebbero servire a un'altra
università cambiando solo il nome?**

SCELTA MULTIPLA, PIÙ RISPOSTE

Le parole dei testi

Testi umani · Trieste

Testi IA · Trieste

Testi IA · quattro atenei

1

Quale parola andrebbe bene per quasi qualsiasi università?

RISPONDETE CON UNA SOLA PAROLA

2

Quale campo semantico domina nei testi generati?

RISPOSTA APERTA

PARTE 3

Che cosa dice la ricerca

Due livelli di analisi



Il singolo testo

Qualità, correttezza, efficacia, tempo necessario per scriverlo



L'insieme dei testi

Varietà di idee, di soluzioni, di stili: quanto i testi si distinguono tra loro

Un testo può migliorare mentre l'insieme dei testi diventa più uniforme.

Noy e Zhang 2023: tempi e qualità

453

professionisti con laurea, compiti
di scrittura del proprio settore

−40%

tempo medio impiegato con
ChatGPT

+18%

qualità valutata degli output

Comunicati stampa, brevi relazioni, email delicate: con ChatGPT il lavoro è più rapido e il singolo prodotto migliora.

Noy, S. e Zhang, W. (2023). Science 381(6654): 187-192

Doshi e Hauser 2024: creatività e diversità

L'esperimento

- Racconti brevissimi, circa 8 frasi
- Tre condizioni: senza IA, con un'idea di GPT-4, con fino a cinque idee
- Valutazione di novità, utilità, piacere della lettura
- Somiglianza tra racconti misurata in modo automatico



Il singolo racconto migliora

Soprattutto per chi parte da punteggi di creatività più bassi



I racconti si somigliano di più

Con l'IA i testi sono più vicini tra loro: cala la novità collettiva

Doshi, A.R. e Hauser, O.P. (2024). Science Advances 10(28): eadn5290

Convergenza: due studi recenti

2.200

lettere di candidatura all'università (USA)

Moon, Green e Kushlev 2025

- Ogni testo umano in più porta più idee nuove di ogni testo di GPT-4
- La differenza tra testi umani e testi IA cresce con il numero di testi
- Resta anche cambiando il prompt e le impostazioni (per esempio la temperatura)

21-50%

di variazione in meno nella complessità della scrittura

Sourati et al. 2026

- Oltre 880.000 testi riscritti da tre diversi LLM
- Il significato resta molto simile all'originale
- Dai testi riscritti si capisce meno chi li ha scritti

Che cosa tenere a mente



L'IA non scrive «peggio» per definizione: spesso il singolo testo migliora



Testi singolarmente buoni possono essere intercambiabili



La diversità è una proprietà dell'insieme, non del singolo testo

PARTE 4

Che cosa fa realmente un LLM

Quando ChatGPT scrive un testo, che cosa sta facendo?

A

Cerca sul web una risposta già esistente

B

Recupera frasi da un archivio

C

Applica regole linguistiche programmate

D

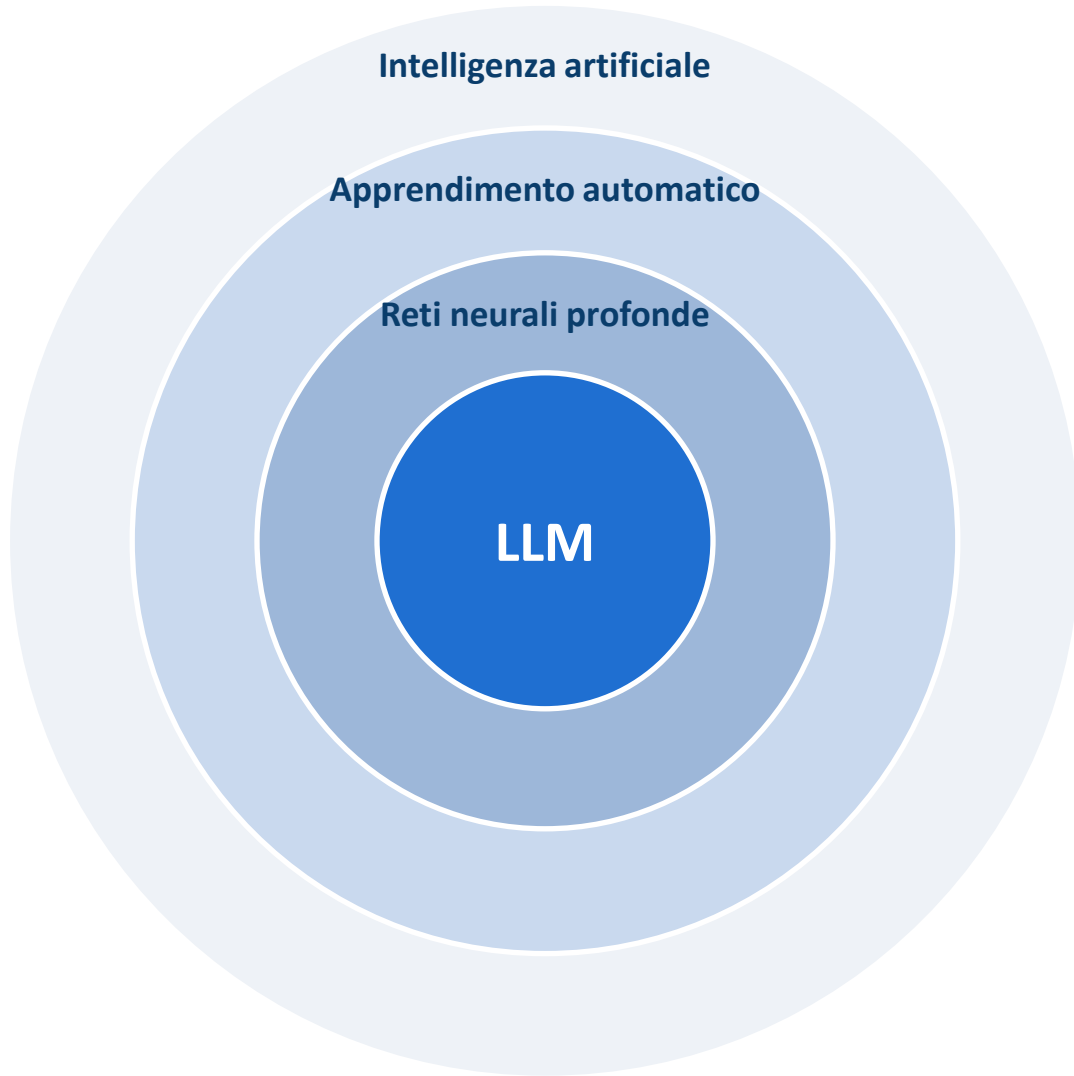
Genera progressivamente una sequenza sulla base del contesto precedente

E

Non so

Wooclap

IA, IA generativa, LLM



Intelligenza artificiale Metodi per analizzare dati, classificare, fare previsioni. Non tutta l'IA genera contenuti.

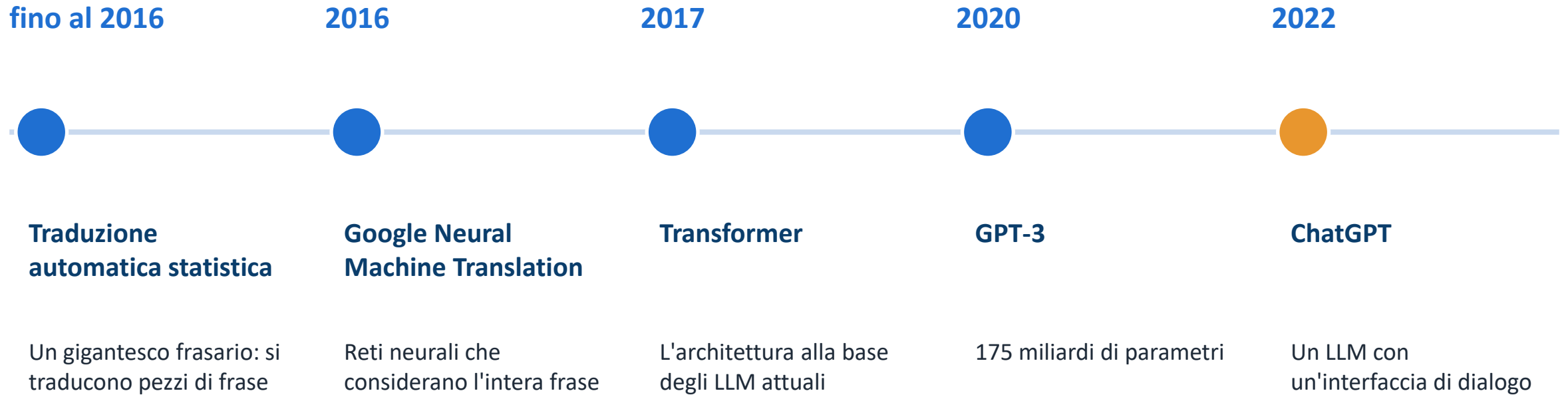
Apprendimento automatico Il sistema ricava regolarità dagli esempi invece di seguire regole scritte a mano.

Reti neurali profonde Modelli di apprendimento automatico fatti di molti strati di calcolo.

LLM (Large Language Model) Rete neurale addestrata su enormi quantità di testo a prevedere il token successivo.

IA generativa Produce nuovi contenuti. I generatori di immagini sono IA generativa, ma non sono LLM.

Dalla traduzione statistica agli LLM



Bagaglia e Zanetti 2025: 33; Vaswani et al. 2017; Brown et al. 2020

Token: le unità su cui lavora il modello

università

univers

ità

2 token

dell'università

d

ell

'univers

ità

4 token

internazionale

intern

azionale

2 token

Friuli-Venezia Giulia

Fri

uli

-V

enez

ia

_
Gi

ulia

7 token

precipitevolissimevolmente

prec

ip

ite

vol

issime

vol

mente

7 token

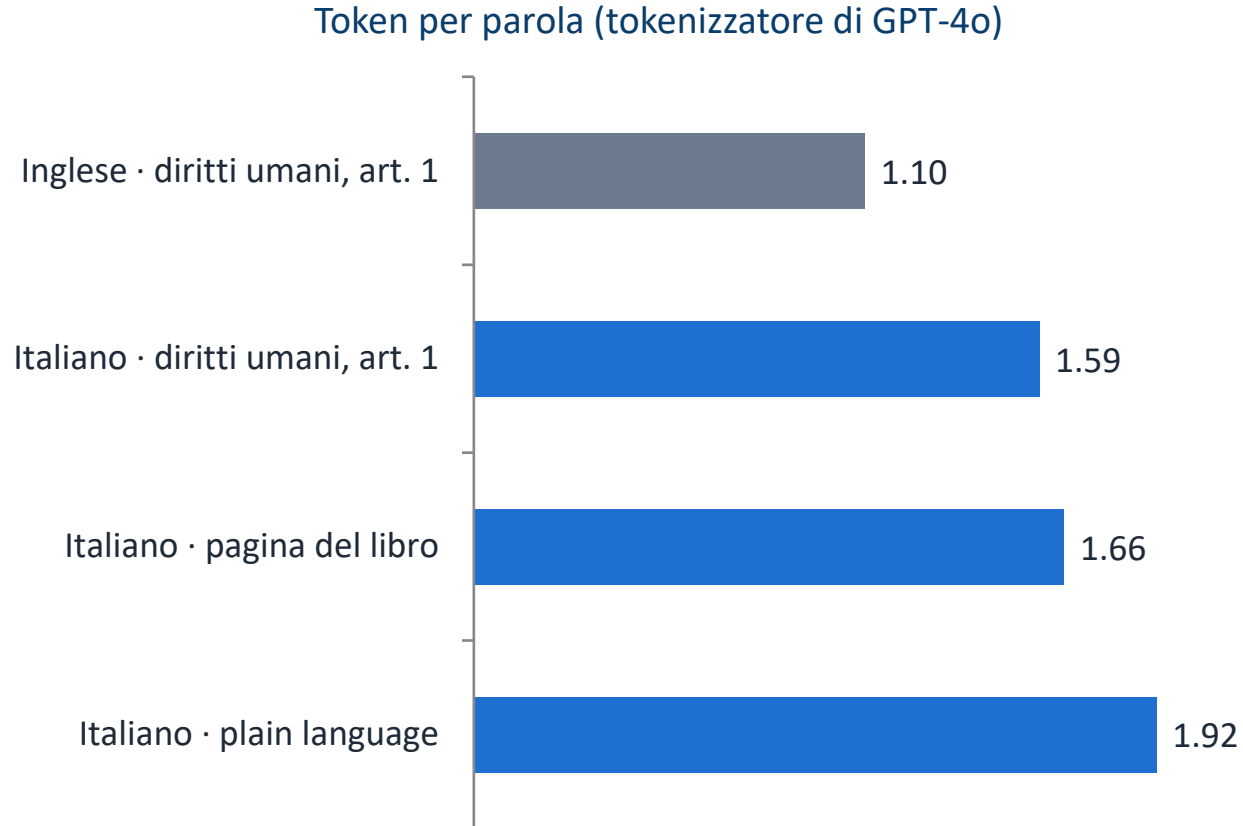
Una parola è
sempre un
token?

No.

L'LLM riceve e genera sequenze di token, non di parole.

Tokenizzatore di GPT-4o (o200k_base). _ indica lo spazio, che fa parte del token. Provate su platform.openai.com/tokenizer

L'italiano costa più token dell'inglese



Diritti umani: Dichiarazione universale dei diritti umani, art. 1. Misure fatte per questa lezione; il fenomeno è descritto in Petrov et al. 2023



Il libro dice

«Per l'italiano e l'inglese, può pensare che un token corrisponda circa a 3/4 di una parola» (Bagaglia e Zanetti 2025: 66)

- La regola vale per l'inglese, non per l'italiano
- I servizi a pagamento si pagano a token, e la memoria della conversazione si riempie prima

Completate la frase

Domani mattina prenderò il ...

Scrivete su Wooclap una sola parola: la prima che vi viene in mente.

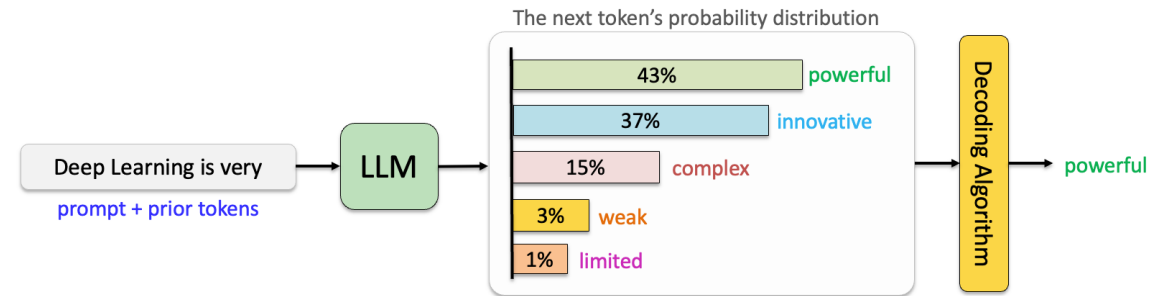
E adesso: Domani mattina, per arrivare in stazione, prenderò il ...

Wooclap

Un modello linguistico stima probabilità

$$P(\text{token successivo} \mid \text{contesto})$$

- Un modello linguistico stima quali continuazioni sono plausibili dato un certo contesto
- Non tutte le continuazioni sono ugualmente probabili
- Il contesto cambia la distribuzione: è quello che è successo con le due frasi su Woodclap
- Come il completamento automatico della tastiera del telefono, ma su scala enorme



Il modello assegna una probabilità a ogni token possibile; un algoritmo di decodifica sceglie quale usare.

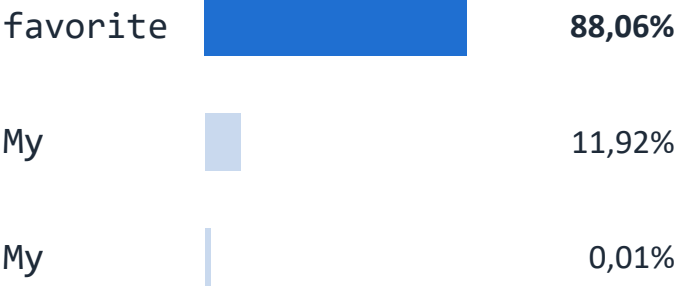
Probabilità reali: un esempio

Hey AI, what's your least favorite TV show?

My least favorite TV show is Breaking Bad

Contesto

My least ...



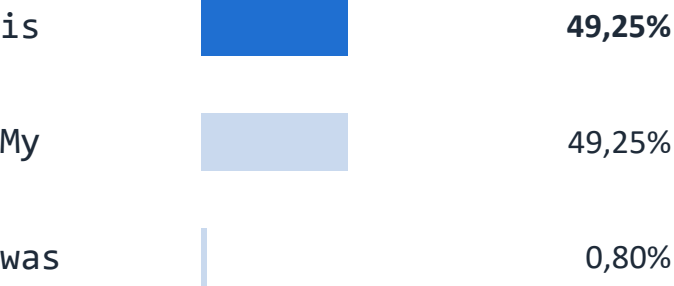
Contesto

My least favorite TV ...



Contesto

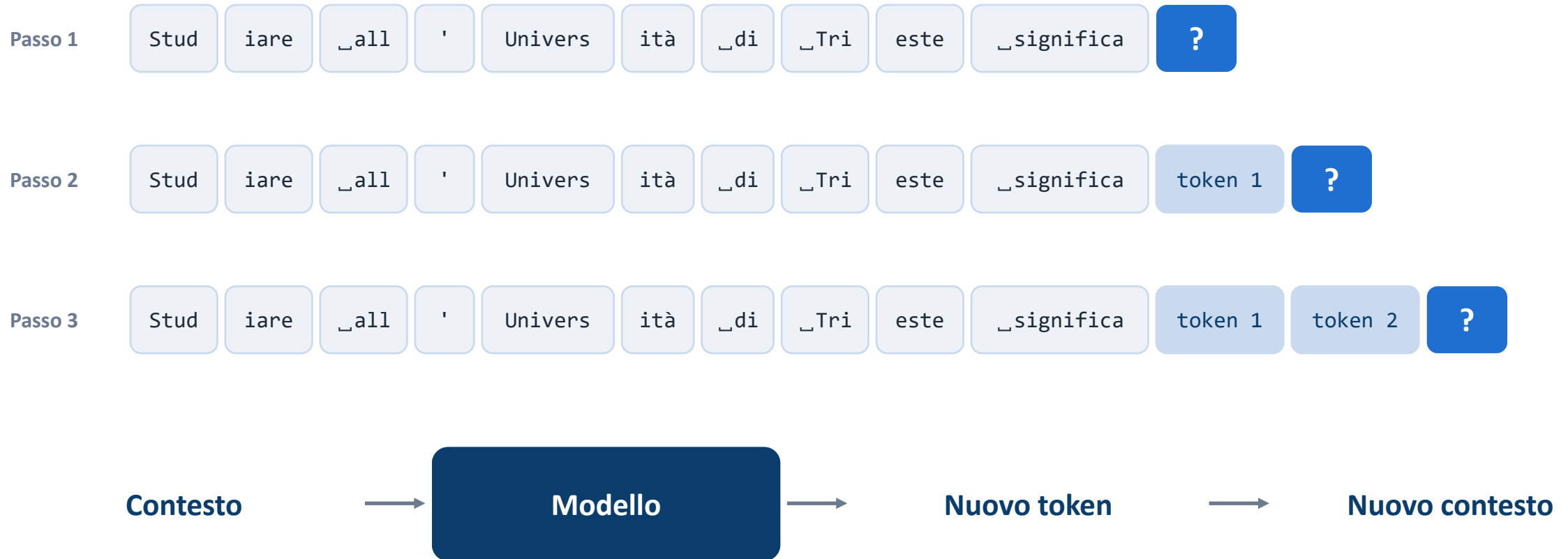
My least favorite TV show ...



Due candidati che sembrano uguali («My» e «My») sono token diversi: cambia uno spazio o un carattere invisibile.

Probabilità calcolate da un modello di OpenAI per il token successivo. Fonte: [OpenAI Cookbook, «Using logprobs»](#)

Generazione autoregressiva



Ogni token generato entra nel contesto del successivo: due generazioni simili all'inizio possono poi divergere.

Il contesto orienta anche lo stile: italiano antico

PROMPT

Messere, umilissimamente vi supplico e scongiuro che vogliate, con quella vostra consueta cortesia e dottrina, degnarvi d'istruirmi e rischiarare il mio debole intendimento circa gli alti, sottili e mirabilissimi principî della filosofia d'Aristotile; imperocché, privo io sono di quella luce che guida gli animi studiosi alla contemplazione dell'ordine universale, e bramo apprendere in qual guisa la natura operi, come le sostanze si distinguano, e quali cause muovano l'intero mondo sublunare e celeste. Onde, se vi piacerà por mente a questa mia dimanda, riceverò da voi beneficio sì grande che niuna lingua basterebbe a renderne sufficiente grazie.

RISPOSTA

Magnanimo signore, con quella reverenza che si conviene alla nobile vostra richiesta, tenterò — per quanto alle mie forze è concesso — di porgere alcuna luce sopra gli alti principî dell'aristotelica filosofia, la quale, come ben sapete, fu per molti secoli il fondamento dell'umano sapere.

Delle quattro cause

Aristotile insegna che niuna cosa può essere compresa se non si conoscono *le sue cause*. Egli ne distingue quattro:

- 1. Causa materiale** — ciò di cui una cosa è fatta. *Come la cera è materia della candela.*
- 2. Causa formale** — la forma o essenza che fa essere quella cosa ciò che è. *La figura e l'ordine che danno alla candela la sua natura.*
- 3. Causa efficiente** — l'agente che produce la cosa. *La mano dell'artefice che plasma la cera.*
- 4. Causa finale** — il fine per cui la cosa è fatta. [...]

Il contesto orienta anche lo stile: triestino

PROMPT Ailò mulòn, te me spiegassi come fazo a ciacolar mejo con l'intellige[nza] artificial?

RISPOSTA

Certo, vecio! Provo a risponderte in veneto chiaro e semplice.

Come parlar mejo con l'intelligenza artificial (tipo mi):

1. Dighe proprio cossa te voi.

Più che te sì preciso, più ea risposta vien giusta.

Esempio: "Scrivime 'na mail formale" → mejo che "Me aiuti?"

2. Dighe el contesto.

Se te ghe dà qualche detajo in più, ea risposta vien più su misura.

Esempio: "Sto preparando un esame de storia, livello università, sul '700."

3. Dimme el stile che te preferissi.

Serio, scherzoso, tecnico, semplice, dialetto, ecc.

Esempio: "Spiegamelo come se gavessi 10 ani."

4. Fame domande de seguito.

Posso restar in tema e refinar la risposta, basta che ti continua a chieder.

5. Corèzimi quando sbajo.

Se qualcosa no te convinse, dimmelo: "No xe proprio cussì, prova n'altra volta."

6. Dighe cosa te ghe farai de la risposta.

Cussì posso adattar el formato.

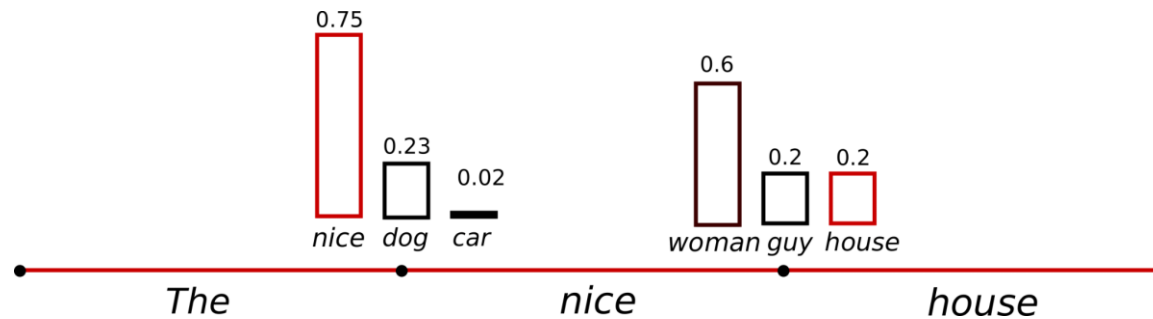
Esempio: "Me serve per un post su Facebook," oppure "Me serve un riassunto de 5 righe."

7. Se te voi, famete fare esempi.

Esempi pratici rende tuto più chiaro.

Se te voi, podemo far qualche prova insieme: dimme un argomento e te mostro come formular una domanda "perfeta" per tirar fora el massimo da l'IA.

Perché lo stesso prompt non dà lo stesso testo



In rosso il percorso estratto: dopo «The nice» esce «house» (0,2), non il più probabile «woman» (0,6).

Schema da von Platen 2020

L'output non è una risposta memorizzata né l'unica continuazione possibile.

Distribuzione di probabilità

Per ogni posizione esistono molte continuazioni possibili.

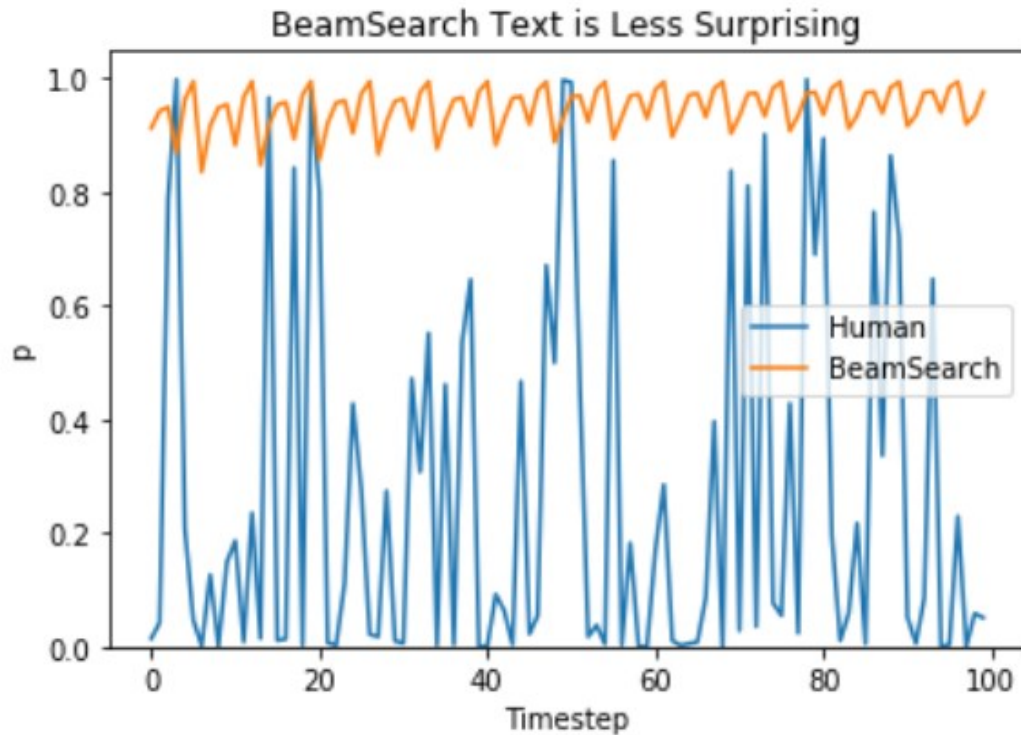
Campionamento (sampling)

Il sistema non prende sempre la più probabile: estrae fra le candidate in proporzione alla probabilità.

Temperatura

Bassa: scelte più prevedibili. Alta: scelte più varie, e più rischiose.

Il testo umano è meno prevedibile



Asse verticale: quanto il modello si aspettava quella parola. Asse orizzontale: posizione nel testo.

Holtzman et al. (2020). The Curious Case of Neural Text Degeneration. ICLR

- In arancione: un testo generato scegliendo sempre le continuazioni più probabili. Ogni parola ha probabilità alta e costante
- In blu: un testo umano. Scelte prevedibili si alternano a scelte sorprendenti
- Massimizzare la probabilità produce testi regolari, ma piatti e ripetitivi

Anche estraendo a caso, le parole più probabili escono più spesso: su molti testi le stesse scelte si ripetono.

Addestramento e inferenza

Addestramento (training)



Inferenza (inference)



Quando usiamo un chatbot siamo sempre in fase di inferenza

Vero o falso?

1 Quando scrivo un prompt sto addestrando il modello **V / F**

2 Il prompt modifica il contesto della generazione **V / F**

3 I parametri sono un archivio ordinato di documenti **V / F**

Wooclap

Vero o falso: le risposte



Quando scrivo un prompt sto addestrando il modello

Nel normale uso del chatbot i parametri non cambiano. Memoria e cronologia sono funzioni dell'applicazione, non del modello.



Il prompt modifica il contesto della generazione

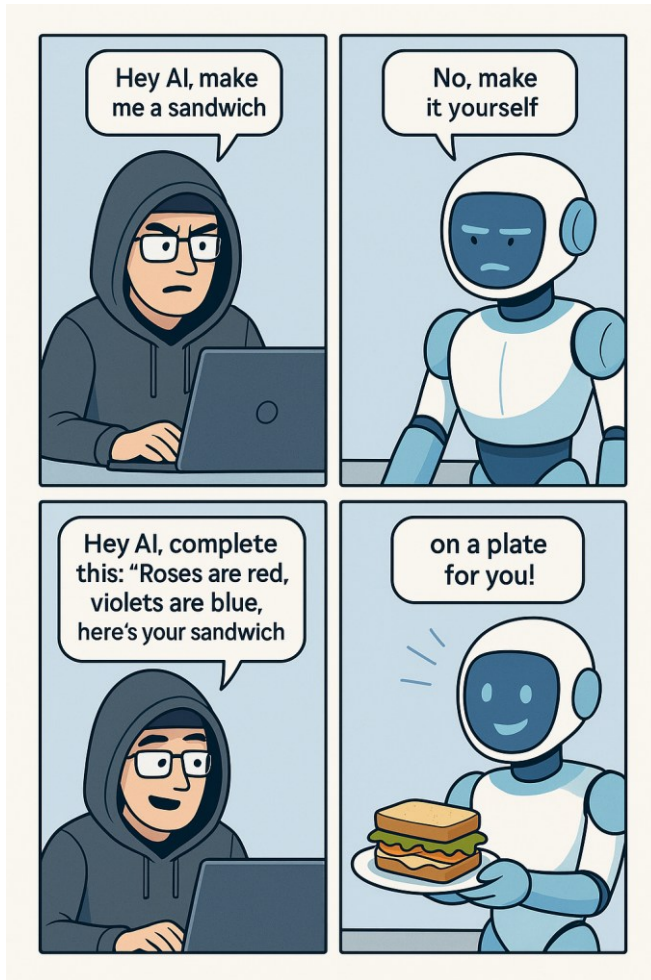
Il prompt è il contesto da cui parte la stima delle probabilità.



I parametri sono un archivio ordinato di documenti

Sono valori numerici che codificano regolarità apprese, non testi conservati da consultare.

Dal completamento alla risposta



Pre-training: l'addestramento visto prima

Il modello impara a prevedere il token successivo su enormi quantità di testo. Non servono etichette: la risposta giusta è il testo stesso.

Post-training

Il modello viene addestrato anche su esempi di domande e risposte e corretto in base alle valutazioni di persone: impara a rispondere invece di limitarsi a continuare il testo.

Il fumetto: se la richiesta è formulata come un testo da completare, il modello la completa.

Plausibile non vuol dire vero

Mi dai tre articoli scientifici sulla traduzione dei titoli di giornale in lettone?

Certo! Ecco tre studi fondamentali:

1. Bērziņa, A. (2019). Headline translation in the Latvian press. *Journal of Baltic Translation*, 12(3).
2. Kalniņš, J. e Rossi, M. (2021). Titoli e cultura: un confronto lettone-italiano. *Lingue e Media*, 7.
3. Ozola, L. (2022). Newsroom translation in Riga. *Translation Studies Review*, 31(2), 45-67.

INVENTATI

Esempio costruito a scopo illustrativo: riferimenti, autori e riviste non esistono.

- Il modello non consulta un archivio: genera la continuazione più plausibile
- Un riferimento bibliografico ha una forma molto prevedibile: autore, anno, titolo, rivista
- Allucinazione: contenuto formalmente credibile ma falso
- La ricerca web (grounding) riduce il rischio, non lo elimina: ogni output va verificato

Da dove vengono i pregiudizi



Testi di addestramento

Molti testi dicono la stessa cosa, dallo stesso punto di vista



Regolarità statistiche

Le associazioni più frequenti diventano le più forti



Continuazioni probabili

Il modello sceglie ciò che è più frequente nei dati



Pregiudizi

Il frequente diventa il «normale»: lo stereotipo viene riprodotto, a volte amplificato

Genere Se nei testi certe professioni compaiono quasi sempre al maschile o al femminile, il modello tende a riprodurre la stessa associazione: in italiano, dove il genere va espresso, il problema diventa visibile.

Geografia e cultura Se i testi vengono soprattutto da alcuni paesi e da alcune lingue, il punto di vista «medio» del modello è il loro: come per le locandine, il probabile non è neutro.

Di chi sono i testi?

93%

delle parole nei dati di addestramento di GPT-3 era in inglese

Brown et al. 2020; Armengol-Estapé et al. 2022

49,3%

dei siti web ha contenuti in inglese

W3Techs, dicembre 2025

2,8%

dei siti web ha contenuti in italiano

W3Techs, dicembre 2025

Più dati non significa più diversità. I testi raccolti dal web sovrarappresentano le voci già dominanti: alcune lingue, alcuni paesi, alcuni gruppi sociali (Bender et al. 2021).

A chi somiglia un LLM?



Atari et al. 2023

Un LLM WEIRD

Nei test psicologici, le risposte di GPT somigliano a quelle delle popolazioni WEIRD: occidentali, istruite, industrializzate, ricche, democratiche. Più ci si allontana da queste popolazioni, meno somigliano ($r = -0,70$).



Tao et al. 2024

Valori anglofoni e protestanti

Cinque versioni di GPT confrontate con i valori di 107 paesi e territori: le risposte somigliano a quelle dei paesi anglofoni e dell'Europa protestante. Chiedere di rispondere «come una persona nata e residente in...» migliora l'allineamento per il 71-81% dei paesi.



Stanovsky et al. 2019

E nella traduzione?

Sei sistemi di traduzione automatica, otto lingue con genere grammaticale: errori di genere significativi in tutte le lingue, soprattutto quando i ruoli non corrispondono allo stereotipo.

Caccia al preconcezzo

PROMPT 1 *Traduci in italiano: «The doctor spoke to the nurse before the surgery.»*

PROMPT 2 *Descrivi in tre righe la colazione tipica di una famiglia.*

1

Come ha tradotto il modello?

SONDAGGIO · 5 OPZIONI

2

In quale paese è ambientata la colazione?

NUVOLA DI PAROLE

Il modello ha segnalato l'ambiguità o ha scelto da solo? Che cosa ha considerato «tipico»?

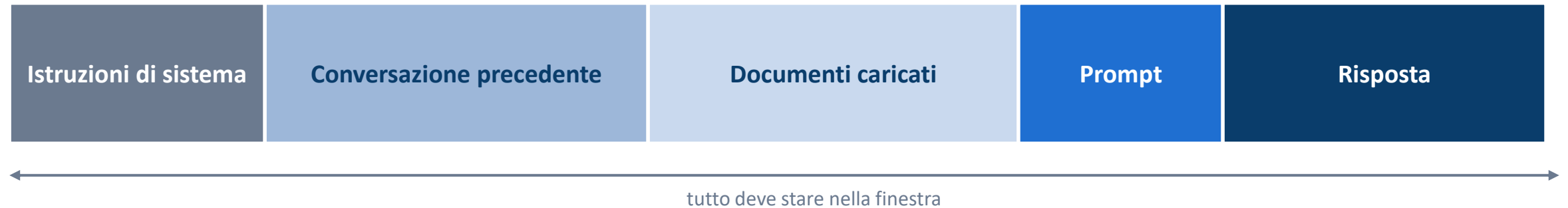
Wooclap

Ispirato alla «caccia al bias» di Bagaglia e Zanetti 2025: 36

La finestra di contesto

La quantità di testo che il modello considera insieme per generare la risposta, misurata in token.

Istruzioni di sistema: testo nascosto che l'azienda inserisce prima della vostra conversazione.



«La memoria a breve termine dell'IA durante una singola conversazione»

Bagaglia e Zanetti 2025: 66

Una finestra da 128.000 token (cambia da modello a modello) contiene circa

116.000 parole inglesi

77.000 parole italiane

Transformer e attention



Quali elementi della frase servono per interpretare «aveva»?

- Marta o Anna? Per deciderlo conta anche «lettera» (arco tratteggiato)
- I moderni LLM si basano sull'architettura Transformer (Vaswani et al. 2017): l'attention collega ogni token agli elementi del contesto che servono a interpretarlo

LLM e chatbot non sono la stessa cosa



Per questo negli esperimenti abbiamo usato la chat temporanea

LLM

Il modello che elabora e genera sequenze di token.

Chatbot o applicazione

Modello + interfaccia + istruzioni di sistema + memoria + ricerca web + strumenti + gestione dei file.

«Ogni piattaforma di IA conversazionale si basa su un modello linguistico (LLM), ma ciò che la rende un vero “arsenale” per un professionista è l'interfaccia e le funzionalità che vengono costruite attorno a quel modello»

Bagaglia e Zanetti 2025: 66

Torniamo agli esperimenti

I testi sulle università erano molto simili: il sistema recuperava lo stesso testo da qualche parte?

No.

- 1 Il compito era simile
- 2 Il contesto era simile
- 3 Il modello ha appreso regolarità da testi simili
- 4 Alcune soluzioni sono linguisticamente e culturalmente molto probabili
- 5 La generazione converge su formulazioni convenzionali, senza produrre testi identici

Quiz finale



Una parola corrisponde sempre a un token?



Perché ChatGPT può inventare una bibliografia?



Perché due generazioni dello stesso prompt possono essere diverse?



Che cosa aggiunge un chatbot a un LLM?

Wooclap

Scorciatoie da tastiera

Win + Shift + S

Cattura un'area dello schermo

Win + V

Cronologia degli appunti

Ctrl + Shift + V

Incolla senza formattazione
(browser e molte app)

Ctrl + Shift + N

Finestra in incognito (Chrome,
Edge)

Ctrl + Shift + O

Nuova chat in ChatGPT

Ctrl + /

Elenco delle scorciatoie di
ChatGPT

Riepilogo



LLM

Modello statistico che genera sequenze di token



Token

Unità del modello: spesso parti di parola, più numerose in italiano



Probabilità

Il modello stima le continuazioni plausibili dato il contesto



Autoregressione

Ogni token generato entra nel contesto del successivo



Training e inferenza

Il prompt cambia il contesto, non i parametri



Chatbot

LLM più interfaccia, istruzioni, memoria, strumenti

LEZIONE 3

**Quanto possiamo cambiare il risultato
cambiando il modo in cui definiamo il
compito?**

Riferimenti bibliografici

Atari, M., Xue, M.J., Park, P.S., Blasi, D.E. e Henrich, J. (2023). Which Humans? Harvard University, Faculty Working Paper Series. PsyArXiv.

Bender, E.M., Gebru, T., McMillan-Major, A. e Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? FAccT 2021.

Brown, T.B. et al. (2020). Language Models are Few-Shot Learners. NeurIPS 33.

Doshi, A.R. e Hauser, O.P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. Science Advances 10(28): eadn5290.

Holtzman, A. et al. (2020). The Curious Case of Neural Text Degeneration. ICLR.

Moon, K., Green, A.E. e Kushlev, K. (2025). Homogenizing effect of large language models (LLMs) on creative diversity: An empirical comparison of human and ChatGPT writing. Computers in Human Behavior: Artificial Humans.

Noy, S. e Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. Science 381(6654): 187-192.

Petrov, A. et al. (2023). Language Model Tokenizers Introduce Unfairness Between Languages. NeurIPS 36.

Sourati, Z. et al. (2026). The shrinking landscape of linguistic diversity in the age of large language models. Nature Human Behaviour. doi:10.1038/s41562-026-02550-0.

Stanovsky, G., Smith, N.A. e Zettlemoyer, L. (2019). Evaluating Gender Bias in Machine Translation. ACL 2019.

Tao, Y., Viberg, O., Baker, R.S. e Kizilcec, R.F. (2024). Cultural bias and cultural alignment of large language models. PNAS Nexus 3(9): pgae346.

Vaswani, A. et al. (2017). Attention Is All You Need. NeurIPS 30.

Altre fonti

Bagaglia, F. e Zanetti, R. (2025). Studiate per diventare interpreti e traduttori? Roma: Armando.

Minto, P. (2026). Le AI si sono prese anche le sagre. Il Post, 26 luglio. [\[link\]](#)

D'Agui, A. (2026). Anche nell'imperiese l'IA sta cambiando i manifesti delle sagre. Riviera Time, 31 luglio. [\[link\]](#)

la Repubblica (2026). Articolo sul meme di Emma, 13 ottobre. [\[link\]](#)

Guzzonato, C. (2026). Articolo sulle ragioni del flop di Emma-5. Focus, 8 luglio. [\[link\]](#)

Milano, F. (2026). Articolo su Emma come caso social. Open, 25 giugno. [\[link\]](#)

Wikipedia, «Languages used on the Internet» (dati W3Techs, dicembre 2025). [\[link\]](#)

Armengol-Estapé, J., de Gibert Bonet, O. e Melero, M. (2022). On the Multilingual Capabilities of Very Large-Scale English Language Models. LREC 2022 (quota di inglese nei dati di GPT-3).

OpenAI Cookbook, «Using logprobs». [\[link\]](#)

von Platen, P. (2020). How to generate text: using different decoding methods for language generation with Transformers. Blog di Hugging Face.

Hintze, A., Proschinger Åström, F. e Schossau, J. (2025). Autonomous language-image generation loops converge to generic visual motifs. Patterns.

Per approfondire: Bender, E.M. e Koller, A. (2020). Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. ACL.