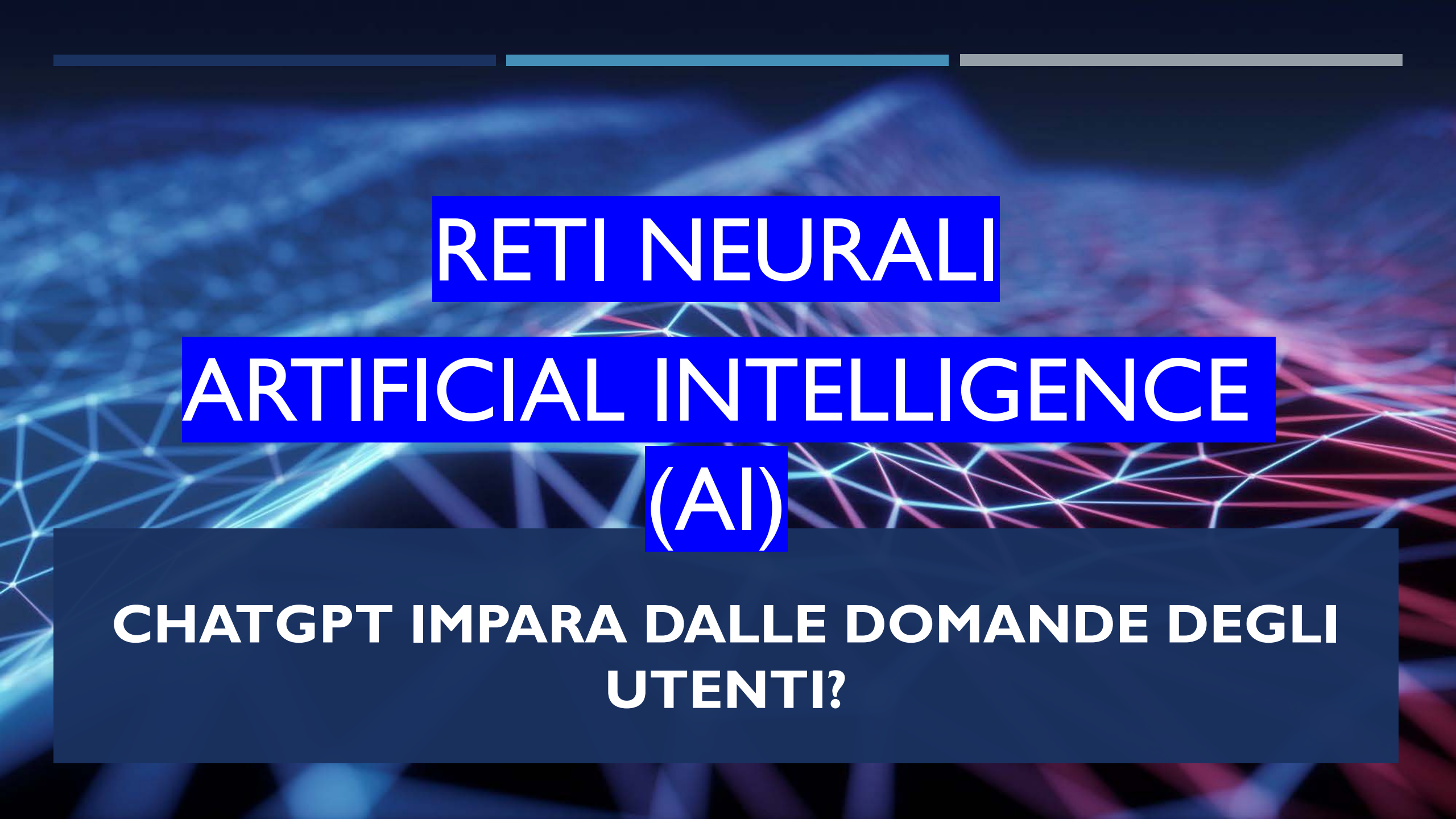




RETI NEURALI ARTIFICIAL INTELLIGENCE (AI)

**CHATGPT IMPARA DALLE DOMANDE DEGLI
UTENTI?**

Introduzione

- 🔗 **Cos'è ChatGPT (Chat Generative Pre-trained Transformer)?**
- ✓ Un'architettura neurale che elabora sequenze tramite meccanismi di attenzione, permettendo il processamento parallelo senza passo-per-passo, usata nei modelli GPT di OpenAI.
 - ✓ Basato su *Large Language Models* (LLM), cioè reti neurali addestrate su enormi quantità di testi.
 - ✓ Obiettivo: generare risposte naturali e utili a domande e richieste degli utenti.

🔗 **Domanda centrale della presentazione:**

“Quando interagiamo con ChatGPT,
lui impara davvero da noi?”

COME FUNZIONA UN LLM (A GRANDI LINEE)

Addestramento iniziale:

- Viene alimentato con miliardi di testi presi da libri, siti web, articoli, forum, ecc.
- Impara a riconoscere schemi e probabilità nelle sequenze di parole.

Fine-tuning (rifinitura):

- Dopo l'addestramento, viene migliorato con dati selezionati e feedback umano (*Reinforcement Learning with Human Feedback* - RLHF).

Risultato:

- Non “capisce” come un umano, ma genera testi molto coerenti basandosi su probabilità linguistiche.

CHATGPT IMPARA IN DUE MODI DIVERSI

A) MEMORIA PERSONALE (PER IL SINGOLO UTENTE) 1/3

A) Memoria personale (per il singolo utente)

- Di default **non ricorda nulla** dopo una sessione.
- Se l'utente attiva la funzione *Memoria*:
 - ❖ Può ricordare informazioni personali (es. nome, interessi, stile di scrittura).
 - ❖ L'utente può sempre:
 - vedere ciò che è stato memorizzato,
 - modificare o cancellare,
 - disattivare del tutto la memoria.
- Serve a rendere l'esperienza più personalizzata e utile.

Cos'è la Memory?

È un sistema che permette a ChatGPT di ricordare alcune informazioni generali e non sensibili che tu decidi di condividere, così da:

- evitare di ripetere ogni volta le stesse cose,
- adattare lo stile delle risposte alle tue preferenze,
- tenere conto di progetti o attività che segui nel tempo.

CHATGPT IMPARA IN DUE MODI DIVERSI

A) MEMORIA PERSONALE (PER IL SINGOLO UTENTE) 2/3

■ 🔍 Cosa può ricordare

- le tue preferenze di scrittura (es. “scrivi in modo formale”),
 - hobby, interessi o competenze,
 - informazioni utili per progetti in corso,
 - il nome della tua azienda (se vuoi),
 - il formato preferito dei documenti,
 - il tipo di risposte che preferisci (brevi, lunghe, tecniche...).
- 👉 **Non memorizza mai automaticamente dati sensibili**, come informazioni finanziarie, sanitarie, password, dati personali privati, opinioni politiche profonde, ecc.

🚫 Cosa NON ricorda

- Documenti che carichi.
- Conversazioni passate parola per parola.
- Informazioni sensibili o private.
- Cose che non hai autorizzato o che non sono appropriate per la memoria.

CHATGPT IMPARA IN DUE MODI DIVERSI

A) MEMORIA PERSONALE (PER IL SINGOLO UTENTE) 3/3

Come controllarla

- Puoi:
- Attivarla o disattivarla in qualsiasi momento.
- Vedere cosa ChatGPT ha memorizzato** (è tutto trasparente).
- Modificare o cancellare singole memorie.**
- Cancellare tutta la memoria** con un clic.

Esempio semplice

Tu: *“Preferisco che le risposte siano sempre molto concise.”*

ChatGPT: *“Vuoi che lo ricordi per il futuro?”*

Se rispondi sì → da quel momento risponderà sempre in modo conciso.

Privacy

La Memory è:

- **Disattivata di default**
- **Controllata interamente da te**
- **Separata dai dati usati per migliorare i modelli**
(altra opzione indipendente)
- **Cancellabile totalmente quando vuoi**

CHATGPT IMPARA IN DUE MODI DIVERSI

B) ADDESTRAMENTO DEI MODELLI FUTURI (PER TUTTI GLI UTENTI) 1/3

I) L'apprendimento del modello (training)

Questo è l'“imparare” vero e proprio dell'IA.

Succede **solo nei laboratori OpenAI**, non nelle conversazioni con te.

Non imparo automaticamente da te durante la chat.

Questa conoscenza rimane fissa finché OpenAI non rilascia una nuova versione del modello.

CHATGPT IMPARA IN DUE MODI DIVERSI

B) ADDESTRAMENTO DEI MODELLI FUTURI (PER TUTTI GLI UTENTI) 2/3

1. Pre-addestramento (Pretraining)

- ☐ Il modello legge **enormi quantità di testo** da libri, articoli, siti web, ecc.
- ☐ L'obiettivo è **predire la parola successiva** in una frase.
- ☐ Durante questo processo, il modello impara **strutture linguistiche, conoscenze generali e schemi statistici**.
- ☐ Qui non “impara” esempi specifici per rispondere a te personalmente, ma costruisce una sorta di **modello statistico del linguaggio**.

2. Addestramento con rinforzo (RLHF – Reinforcement Learning from Human Feedback)

- ☐ Umani valutano risposte generate dal modello e indicano quali sono “migliori”.
- ☐ Il modello usa queste valutazioni per **aggiustare i pesi della rete neurale**, cioè per dare più probabilità a risposte giudicate corrette o utili.
- ☐ Questo lo rende più “umano” nel rispondere, ma sempre generalizzato, non personale.

CHATGPT IMPARA IN DUE MODI DIVERSI

B) ADDESTRAMENTO DEI MODELLI FUTURI (PER TUTTI GLI UTENTI) 3/3

3. Memoria persistente opzionale

OpenAI può abilitare una **memoria tra chat**:

- ☐ In questo caso, alcune informazioni possono essere salvate in un database separato.
- ☐ Non sono aggiornamenti dei pesi del modello: è più come un **quaderno digitale** da cui il modello può leggere quando risponde.
- ☐ Tu puoi visualizzare, aggiornare o cancellare queste informazioni.

4. Cosa non succede

- ☐ Il modello **non modifica il suo comportamento globale** sulla base delle singole conversazioni ordinarie.
- ☐ Non “ricorda” fatti personali a meno che la memoria persistente sia attiva.
- ☐ Non accumula conoscenza come un umano o un database dinamico in tempo reale.

In pratica: ChatGPT **apprende statisticamente durante l'addestramento**, ma in chat “normali” funziona solo come se stesse usando **una finestra temporanea di memoria** per rispondere coerentemente.

ESEMPI PRATICI

Caso 1: Memoria personale

- o Tu chiedi: “Ricordati che sto preparando la tesi in biologia marina.”
- o Nelle conversazioni future ChatGPT ti proporrà esempi o spiegazioni più adatte al tuo campo.

Caso 2: Addestramento generale

- o Molti utenti chiedono informazioni su una nuova tecnologia uscita di recente.
- o ChatGPT oggi può non saperlo.
- o Se questi dati vengono usati per addestrare il modello successivo, la nuova versione sarà aggiornata e più competente.

TABELLA RIASSUNTIVA

Tipo di “apprendimento”	Dove avviene	Quando	Cosa memorizza	Controllo utente
Training del modello	OpenAI	Prima che il modello venga rilasciato	Linguaggio, conoscenza generale	 Nessun controllo
Memory	Nel tuo profilo	Durante l’uso	Preferenze, contesto personale	 Totale controllo

DIFFERENZE CHIAVE

Aspetto	Memoria personale	Addestramento generale
Scopo	Aiutare un singolo utente	Migliorare il modello per tutti
Tempistica	Immediata (nella tua esperienza)	Differita (nuove versioni future)
Controllo utente	Totale: puoi attivare/disattivare	Nessun controllo diretto
Tipo di dati	Info che scegli di condividere	Conversazioni anonime e aggregate

VANTAGGI E RISCHI

Vantaggi:

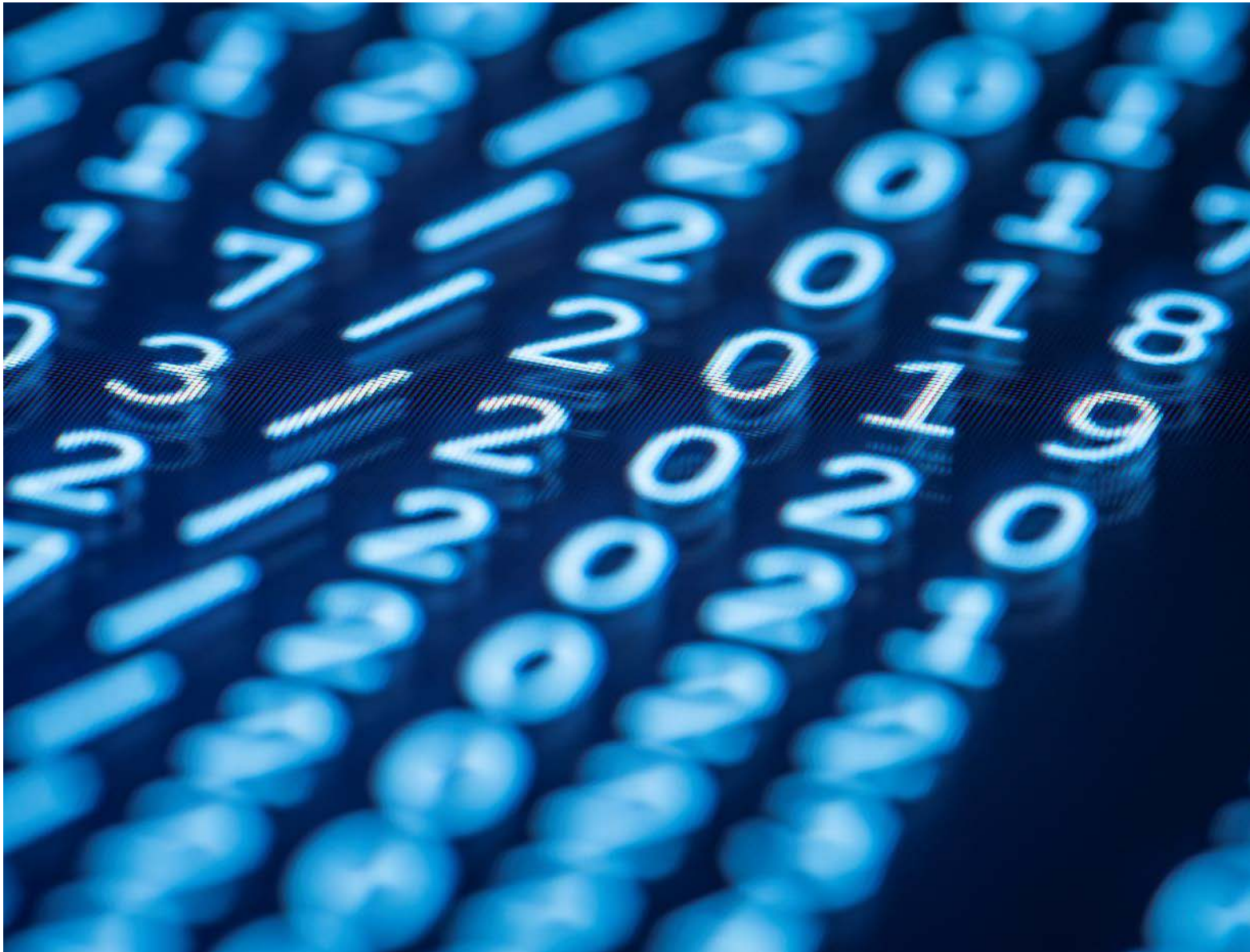
- Risposte più personalizzate (con memoria attiva).
- Modelli sempre più aggiornati e competenti.
- Migliore comprensione del linguaggio e riduzione di errori.

Rischi e precauzioni:

- Necessità di protezione della privacy.
- Possibili bias nei dati usati per l'addestramento.
- Importanza della trasparenza sul funzionamento della memoria e sull'uso dei dati.

CONCLUSIONE

1. ChatGPT **non impara sul momento** dalle tue domande.
2. Ha **due livelli di apprendimento**:
 - ✓ **Memoria personale** (solo se attiva → per migliorare l'esperienza del singolo utente).
 - ✓ **Addestramento generale** (dati usati in forma anonima → per migliorare i modelli futuri).
3. Risultato: un equilibrio tra utilità, evoluzione continua e tutela della privacy.



GRAZIE

ING.BRUNACAMPAGNA@
GMAIL.COM