

Tecno - *IA, Anthropic scopre il "J-Space": Claude ha uno spazio interno che elabora concetti come la mente umana*

Roma - 07 lug 2026 (Prima Notizia 24) **I ricercatori individuano un'area separata dalla "catena di pensiero" visibile all'utente. Il modello trattiene idee in background senza scriverle.**

Importante passo in avanti nella comprensione del funzionamento profondo dei modelli linguistici. I ricercatori di Anthropic – la società guidata da Dario e Daniela Amodei, al centro dell'attenzione mediatica per le recenti controversie con l'amministrazione Trump e attesa al debutto a Wall Street nell'autunno 2026 – hanno individuato una specifica area interna alla piattaforma Claude. Questo spazio è in grado di trattenere ed elaborare idee e concetti in background senza tradurli immediatamente in parole o output visibili all'utente, con una dinamica che ricorda il modo in cui gli esseri umani archiviano o accedono consapevolmente ai propri pensieri. La scoperta è stata illustrata dall'azienda attraverso un video pubblicato sul proprio profilo X. Anthropic ha battezzato questo ambito "J-Space", un nome che deriva dalla tecnica matematica dello "Jacobiano" (Jacobian). All'interno di quest'area distinta, il modello pianifica strategie che prescindono dal compito immediato richiesto e che rimangono separate dalla classica "catena di pensiero" (il processo di ragionamento lineare) condivisa con l'interlocutore. Gli scienziati non hanno dimostrato che l'IA provi sensazioni o esperienze soggettive, ma hanno rilevato una netta distinzione tra le informazioni strettamente utilizzate per formulare una risposta e la mole ben più vasta di elaborazioni sottostanti. Per spiegare il fenomeno, Anthropic ha tracciato un parallelismo con l'attività cerebrale umana: "Proprio come gli esseri umani possono pensare a una cosa mentre ne fanno un'altra, Claude è in grado di attivare concetti ed elaborazioni nel suo J-Space che non sono correlati ai suoi output". A supporto della tesi, i ricercatori hanno fornito un esempio empirico: hanno chiesto a Claude di pensare al Golden Gate Bridge in California mentre, contemporaneamente, eseguiva il compito di copiare una frase del tutto slegata da quell'argomento. Nello svolgimento del test, mentre l'IA era impegnata a trascrivere il testo richiesto nell'output visibile, dietro le quinte il suo J-Space mostrava un'attivazione concettuale differente, in cui figuravano chiaramente i temi "ponte" e "California". Secondo le conclusioni della società, l'osservazione e il monitoraggio delle dinamiche interne al J-Space potrebbero rivelarsi fondamentali in chiave di sicurezza informatica, in particolare per individuare tempestivamente eventuali disallineamenti del modello o l'elaborazione di intenti malevoli non esplicitati nel testo. Data la complessità e la portata della scoperta, che offre nuovi spunti al dibattito sui confini della "coscienza della macchina", Anthropic ha ufficialmente invitato esperti di neuroscienze e della filosofia a condividere le proprie impressioni e analisi sulla ricerca.

(Prima Notizia 24) Martedì 07 Luglio 2026

PRIMA NOTIZIA 24

Sede legale : Via Costantino Morin, 45 00195 Roma
E-mail: redazione@primanotizia24.it