

***Large Language Models* tra diritto e filosofia: i limiti e le conseguenze delle informazioni sintetizzate**

Gianluca Fasano ed Emanuele Bottazzi Grifoni
Istituto di Scienze e Tecnologie della Cognizione (LOA)



14 maggio 2024, h 10:00 – 12:00 online

link per la registrazione: <https://forms.office.com/e/V4J6fJ2RbD>

I *large language models* sono modelli di intelligenza artificiale in grado di generare testi plausibili e contestualmente rilevanti, per mezzo di un'attività che va oltre la semplice riproduzione di informazioni esistenti introducendo un elemento di novità definito, a seconda delle prospettive, come capacità di *reasoning*, abilità creatività ovvero originalità dei contenuti.

Qui offriamo una differente prospettiva che, grazie al metodo di indagine del tardo Wittgenstein, ci conduce ad affermare che i LLM non raggiungono una comprensione linguistica né realizzano una effettiva conversazione con le persone: l'idea di produrre un linguaggio genuino basato su relazioni statistiche tra parole non riesce a catturare un aspetto essenziale della grammatica del nostro linguaggio, ovvero la contraddizione e le sue implicazioni nel discorso.

Ciò nonostante, la singolarità (l'avvento di un'intelligenza superiore a quella umana) è già qui con noi. A dispetto delle apparenze essa non è dovuta ad un salto tecnologico ma ad una simulazione che la tecnologia produce, facendo leva sull'atteggiamento che abbiamo nei confronti dei nostri artefatti. La singolarità è nelle nostre menti, non nelle macchine, per cui dobbiamo chiederci quali conseguenze produce sulle nostre libertà.

Per mezzo di questa simulazione la macchina prende parte all'esercizio della libertà degli esseri umani. Non si ritiene, sia detto a evitare equivoci, che la macchina eserciti un diritto di libertà ma, egualmente all'esercizio della libertà di espressione dei vivi, ad essa deve corrispondere l'assunzione di doveri e responsabilità, non solo in termini di qualità, veridicità e correttezza, ma sulla base di nuovi modelli di autorialità connessi alla creazione di 'informazioni sintetizzate'.