



Multilingual, Open-text Unified Syntax-independent SEMantics

Risultati in breve

Un progetto apre nuove strade all'intelligenza artificiale per creare il «DNA del linguaggio»

Grazie a nuovi usi dell'apprendimento automatico, la ricerca offre rappresentazioni dei testi indipendenti dalla lingua in grado di affrontare le allucinazioni dell'intelligenza artificiale.



RICERCA DI BASE



Le allucinazioni dell'intelligenza artificiale (IA), ovvero la generazione di contenuti falsi e non basati su dati reali, sono ormai un argomento di tendenza da quando sono stati resi pubblici corposi modelli linguistici come ChatGPT e Bard. Oggi, grazie a un progetto di ricerca finanziato dall'UE, gli specialisti informatici possono avvicinarsi a una perfetta elaborazione del linguaggio naturale (o NLP, dall'inglese Natural Language Processing) indipendentemente dalla lingua in questione,

evitando che l'IA fornisca informazioni false.

Il progetto [MOUSSE](#) (Multilingual, Open-text Unified Syntax-independent SEMantics) ha esplorato nuove direzioni di lavoro per migliorare le capacità di parsing semantico multilingue, senza l'onere di annotare i dati per ogni singolo idioma.

«Per quanto potenti e dalle capacità impressionanti, i modelli linguistici di grandi dimensioni, come ChatGPT o Bard, hanno ancora difficoltà a replicare la sicurezza e

il buon senso che caratterizzano gli esseri umani. MOUSSE getta le basi per raggiungere questo ambizioso obiettivo. Il progetto offre un enorme archivio di conoscenze multilingue utilizzabile come base per i ragionamenti e i prodotti di questi modelli, nonché per affrontare il problema dell'allucinazione», afferma [Roberto Navigli](#), responsabile di [Sapienza Natural Language Processing Group](#) e coordinatore del progetto MOUSSE.

Il multilinguismo come risorsa

L'ampio archivio sviluppato da MOUSSE è descritto da Navigli come «il DNA del linguaggio», poiché fornisce le fondamenta per costruire frasi significative in molte lingue diverse.

Questo è possibile grazie al principale risultato raggiunto dal progetto: la capacità di creare l'equivalente computazionale delle rappresentazioni mentali dei testi generate dal cervello umano, ma indipendentemente dal linguaggio. Navigli spiega: «Il computer sviluppa un'idea del significato di una frase che astrae dal linguaggio e dalla forma superficiale, cioè dalle parole attraverso le quali il significato è espresso.»

Più lingue il team utilizza per esprimere la semantica, più può corroborare la qualità della rappresentazione appresa. Inoltre, una volta ottenuta una rappresentazione da una frase in una lingua, è possibile produrre frasi in lingue diverse per esprimere il medesimo significato. «È una tecnica molto simile alla traduzione automatica, ma con un passo avanti: fornisce una prova formale e strutturata di ciò che la macchina ha capito», spiega Navigli.

Sfruttando il multilinguismo, MOUSSE contribuisce a uniformare il campo della ricerca nell'NLP per tutte le lingue dell'UE e per centinaia di altre. L'archivio multilingue può essere utile anche agli studenti di lingue per migliorare il vocabolario e per apprendere basandosi più sul significato che sulle singole parole.

Sfruttare l'IA per i migliori risultati

Le capacità sviluppate da MOUSSE sono ottenute in quattro fasi principali: disambiguazione del senso delle parole, collegamento delle voci, etichettatura dei ruoli semantici e parsing semantico. Il team ha raggiunto risultati di ottimo livello non solo grazie all'apprendimento profondo, ma anche facendo sì che il modello e i suoi prodotti fossero sempre interpretabili e manipolabili.

In sintesi, il progetto è stato in grado di mettere in collegamento la conoscenza simbolica e le reti neurali, producendo un approccio neuro-simbolico innovativo. Pertanto, secondo Navigli, il sistema coglie il meglio dei due mondi: da una parte le alte prestazioni e l'efficacia dei modelli neurali, dall'altra l'interpretabilità, la

manipolabilità e l'indipendenza dal linguaggio ricavate dalla parte simbolica.

Le conoscenze simboliche sono fornite da grafici di conoscenze multilingue come BabelNet, un estesissimo dizionario enciclopedico computazionale multilingue nato dal progetto [MultiJEDI](#), anch'esso coordinato da Navigli. L'ingegneria e la sostenibilità dei risultati di entrambi i progetti sono state rese possibili da [Babelscape](#) , la sua azienda universitaria spin-off di grande successo.

Parole chiave

[MOUSSE](#)

[IA](#)

[intelligenza artificiale](#)

[allucinazione dell'IA](#)

[modello linguistico di grandi dimensioni](#)

[elaborazione del linguaggio naturale](#)

[NLP](#)

[apprendimento automatico](#)

[apprendimento profondo](#)

[disambiguazione del senso delle parole](#)

[etichettatura dei ruoli semantici](#)

[parsing semantico](#)

[IA simbolica](#)

Scopri altri articoli nello stesso settore di applicazione



migliorare la raccolta e la gestione dei dati nell'ambito della scienza dei cittadini

29 Settembre 2022





L'elaborazione quantistica aumenta progressivamente e balza fuori dalla teoria alla realtà

29 Maggio 2020 



Una ricerca rivela i meccanismi che guidano le affascinanti transizioni di fase dall'isolante al metallo

26 Marzo 2021 



Insegnare alle reti neurali a seguire il flusso

26 Febbraio 2021 

Informazioni relative al progetto

MOUSSE

ID dell'accordo di sovvenzione: 726487

[Sito web del progetto](#) 

DOI

[10.3030/726487](https://doi.org/10.3030/726487) 

Progetto chiuso

Finanziato da

EXCELLENT SCIENCE - European Research Council (ERC)

Costo totale

€ 1 497 250,00

Contributo UE

€ 1 497 250,00

Coordinato da

Data della firma CE

27 Febbraio 2017

UNIVERSITA DEGLI STUDI DI
ROMA LA SAPIENZA

 Italy

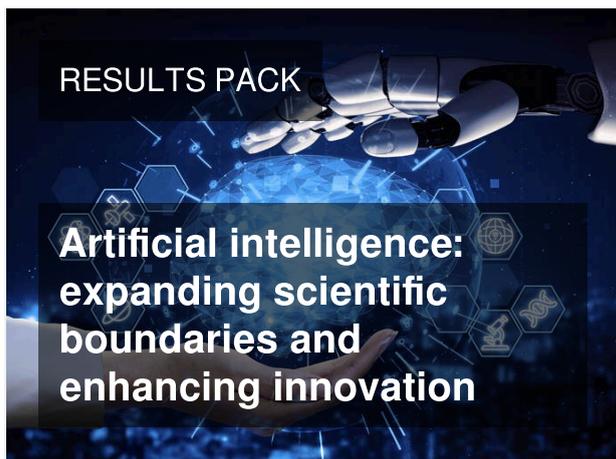
Data di avvio

1 Giugno 2017

Data di
completamento

31 Maggio 2023

Questo progetto è apparso in...



5 Settembre 2023



Ultimo aggiornamento: 25 Agosto 2023

Permalink: <https://cordis.europa.eu/article/id/446038-project-breaks-new-grounds-in-ai-to-create-dna-of-language/it>

European Union, 2025