Special Issue

Atoms of Representation in Natural Language Processing

Message from the Guest Editor

As language models take center stage not only in NLP but in a vast array of scientific applications, the question of how it is best to map natural language in textual form into vector space gains more and more interest. While most popular models still use subword tokens as their atomic units, "token-free" methods including character-level, byte-level, and encoding of visual text rendering have been making promising progress. Still, development and analysis of tokenization and untokenization methods is advancing at a slower rate than research in model architecture and optimization technologies, mostly due to the early stage at which representation is applied, which makes evaluation of new algorithms and techniques particularly challenging. Fundamental insights into the effect of representation atomicity on morphological modeling, on multilingual and crosslingual applications, on computation efficiency, on representations of groups in society, and on other aspects, are still being gained, making this research topic ripe for aggregation and integration of findings and methodologies.

Guest Editor

Dr. Yuval Pinter

Department of Computer Science, Ben-Gurion University of the Negev, Beer Sheva, Israel

Deadline for manuscript submissions

31 January 2025

MDPI

Applied Sciences

mdpi.com/journal/
applsci

About the Journal

Message from the Editor-in-Chief

As the world of science becomes ever more specialized, researchers may lose themselves in the deep forest of the ever increasing number of subfields being created. This open access journal Applied Sciences has been started to link these subfields, so researchers can cut through the forest and see the surrounding, or quite distant fields and subfields to help develop his/her own research even further with the aid of this multi-dimensional network.

Editor-in-Chief

Prof. Dr. Giulio Nicola Cerullo
Dipartimento di Fisica, Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milano, Italy

Author Benefits

Open Access:
free for readers, with article processing charges (APC) paid by authors or their institutions.

High Visibility:
indexed within Scopus, SCIE (Web of Science), Inspec, CAPlus / SciFinder, and other databases.

Journal Rank:
JCR - Q1 (Engineering, Multidisciplinary) / CiteScore - Q1 (General Engineering)