

Tokens

Tokens refer to the smallest units of a language, typically words or characters, used for text analysis and processing. They serve as the building blocks for natural language understanding tasks such as parsing, tokenization, part-of-speech tagging, and named entity recognition. Tokens enable computers to break down and process textual data into meaningful units, facilitating various natural language processing tasks. They play a fundamental role in text analysis, information retrieval, and machine learning applications involving text data.