

# **Understanding Text Splitters**

Text splitters are an important component in modern AI and NLP systems, especially in Retrieval-Augmented Generation (RAG) applications.

## **Why Text Splitters Are Needed**

Large documents cannot be directly processed by language models due to token limits. Text splitters break large text into smaller, manageable chunks while preserving meaning.

## **How Text Splitters Work**

A text splitter divides text based on characters, sentences, paragraphs, or tokens. Some splitters also use overlap to maintain context between chunks.

## **Example**

If a document has 10,000 characters and the splitter size is 1,000 characters with an overlap of 200 characters, the document will be split into multiple chunks with shared context.

## **Benefits of Text Splitters**

- Improves retrieval accuracy
- Handles large documents efficiently
- Preserves semantic context
- Optimizes performance of vector databases