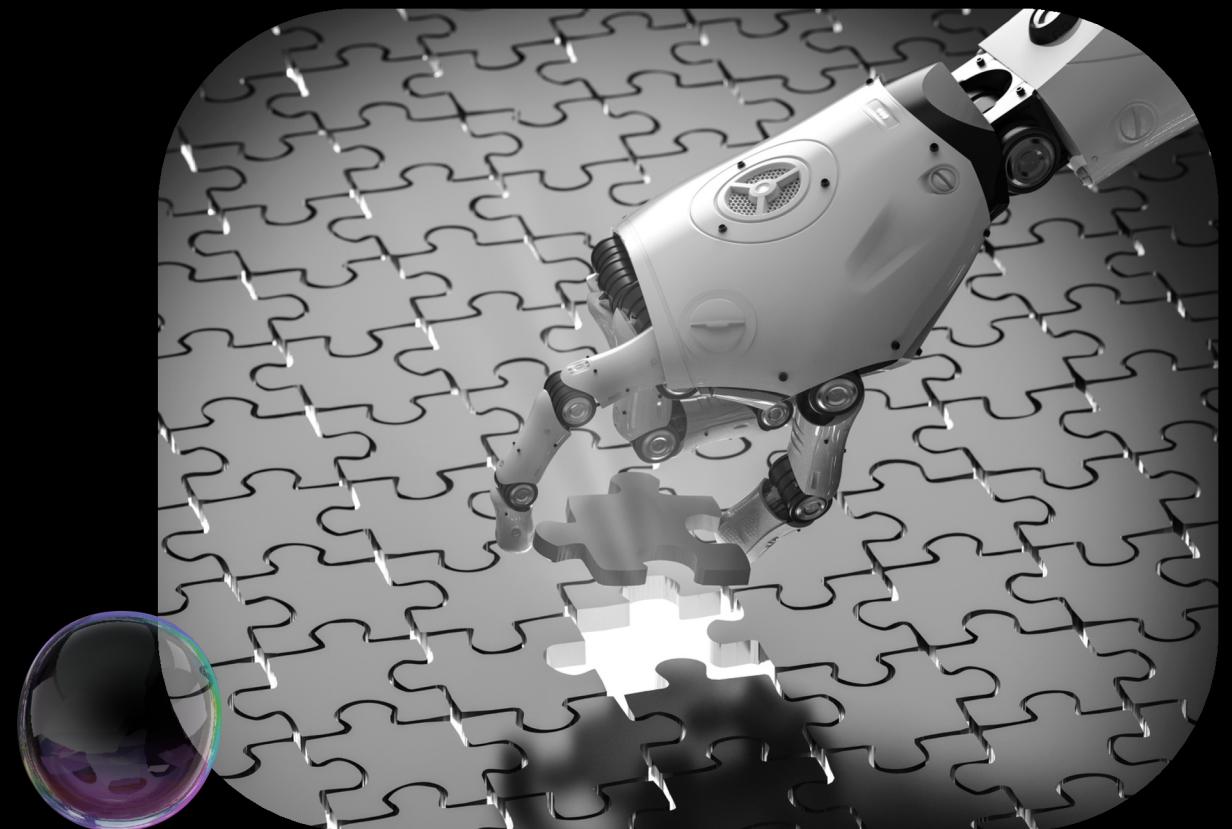


Azure Mastery: Building Smarter Bots

Chunking & Vectorization



Created by
Katerina Chernevskaya





Chunk

Definition

A “chunk” is a small section of text taken from a larger document, designed to be easily understood on its own and used in processing or retrieval tasks.



Created by
Katerina Chernevskaya





Chunking

Definition

Chunking is the process of dividing large texts into smaller, more manageable pieces or chunks.

In the context of RAG (Retrieval-Augmented Generation), using these smaller chunks instead of whole documents allows for more precise retrieval, fewer input tokens, and better context for the language model to generate responses.



Created by
Katerina Chernevskaya



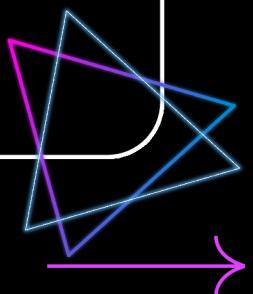
Chunking

Purpose

- **Length Constraints:** Some models, like transformers (e.g., GPT), have a limit on the number of tokens they can process at once. Chunking breaks down large documents into smaller segments that fit within these limits.
- **Context Preservation:** Chunking helps maintain the context of the text within manageable sections, ensuring that the information remains coherent and relevant during processing.
- **Optimization:** This method enables more efficient document processing, preventing system overload and enhancing performance.



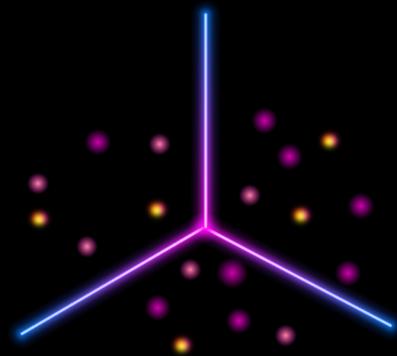
Created by
Katerina Chernevskaya



Vector

Definition

In AI, a vector is a representation of data, typically words or phrases, in a format that a computer can process - usually as points in a multidimensional space.



Created by
Katerina Chernevskaya





Vectorization

Definition

The process of converting text, images, or other types of content into numerical vectors.

These vectors are mathematical representations created by machine learning models, such as those provided by Azure OpenAI or other embedding models.



Created by
Katerina Chernevskaya





Vectorization

Purpose

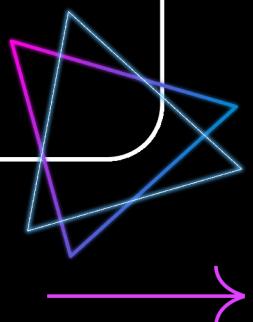
Capture the semantic meaning of the content, enabling efficient similarity searches within a vector space.

Vector search is a way to find information by using numbers instead of plain text. These numbers represent the content in a way that lets the search match things based on:

- **Similar meanings** (like “dog” and “canine”, which mean the same thing even though they’re different words).
- **Different languages** (like “dog” in English and “hund” in German).
- **Various types of content** (like finding a picture of a dog and the word “dog” in text).



Created by
Katerina Chernevskaya



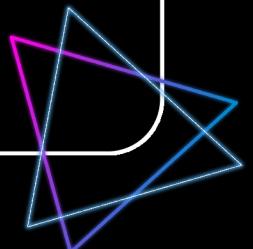
Azure AI Search: Import and vectorize data



Purpose

The “Import and Vectorize” wizard in Azure AI Search makes it easy to bring in text data and automatically turn it into vectors, helping you to efficiently handle and search large amounts of information.

For text vectorization, often uses Azure OpenAI models like **text-embedding-ada-002**, which are popular choices for this task.



Created by
Katerina Chernevskaya



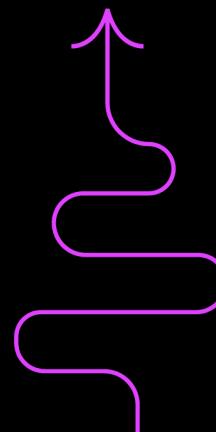
Azure Mastery: Building Smarter Bots

Wanna make
your chatbots
smarter



**REGISTER TODAY,
SECURE YOUR SEAT,
AND SAVE 50%!**

www.balticsummit.pl



BALTIC SUMMIT

UNLEASH THE POWER OF MICROSOFT SOLUTIONS

PPNT GDYNIA, POLAND

13 SEPTEMBER 2024

50% OFF WITH PROMO CODE: KATERINA_BPPC

ENHANCING CHATBOT
INTELLIGENCE WITH
AZURE AI STUDIO'S
PROMPT FLOW

REGISTER NOW

KATERINA CHERNEVSKAYA

www.balticsummit.pl



Azure Mastery: Building Smarter Bots

Key Terms

Follow for more!



Created by
Katerina Chernevskaya