Semantic search is a powerful technique that enables machines to understand the meaning behind words rather than relying solely on keyword matching.

By using embeddings, large language models and vector databases can represent text as high-dimensional vectors. This allows for more accurate and intuitive information retrieval.

For example, when a user searches for "how to cook pasta," the system may return documents containing "boil spaghetti" or "prepare Italian noodles" even if the exact phrase isn't present.

This approach is at the heart of modern AI-powered applications like recommendation engines, chatbots, and document classification systems.

Semantic similarity is often measured using cosine similarity between vector representations of text chunks.

The effectiveness of such systems relies heavily on the quality of the embeddings and the structure of the chunked documents.