Contribution ID: **89** Type: **not specified**

Integrating Knowledge Retrieval Gen AI in Financial Services

Retrieval-Augmented Generation (RAG) offers a robust way to enhance large language models (LLMs) with domain-specific knowledge via external information retrieval. In banking—where precision, compliance, and accuracy are vital—optimizing RAG is crucial. This study explores how various document parsing, chunking, and indexing techniques influence the performance of RAG systems in banking contexts. Our evaluation framework measures their effects on retrieval accuracy, contextual relevance, and output quality, offering practical insights for building more reliable and effective RAG solutions.

Classification

Both methodology and application

Special/ Invited session

Keywords

Retrieval-Augmented Generation (RAG), Large Language Models (LLM), Artificial Intelligence (AI), Parsing, Chunking, Indexing, Document Preprocessing, Information Retrieval, Vector Search

Primary author: MAKRIS, MICHAIL (UNIPI)

Co-authors: Dr BESIMIS, Sotirios (UNIPI); Dr PLAGIANAKOS, Vasileios (University of Thessaly)

Presenter: MAKRIS, MICHAIL (UNIPI)

Track Classification: Machine Learning