

PHAROS Training Series - Course 9 "RAG End-to-End: Architecture, Retrieval, Generation and Evaluation"

Contribution ID: 2

Type: **not specified**

RAG architectures: Naïve, Advanced and Modular

Tuesday, 7 July 2026 11:15 (25 minutes)

This theory session presents the main architectural paradigms used in Retrieval-Augmented Generation systems. It starts with Naïve RAG as the baseline pipeline: document ingestion, chunking, embedding generation, vector indexing, retrieval and grounded answer generation. It then explains how Advanced RAG improves this baseline through better preprocessing, metadata enrichment, query rewriting, hybrid retrieval, reranking, context compression and evidence filtering. The session also introduces Modular RAG, where routing, rewriting, retrieval, fusion, generation and verification can be organised as configurable components rather than a fixed linear chain. Participants will learn how each architecture changes the quality, reliability, latency and maintainability of a RAG application. The session prepares participants to recognise which design pattern is appropriate for a simple educational prototype, a more robust domain assistant or a production system that must adapt to different question types and knowledge sources. It also clarifies how the same concepts appear later in the hands-on notebook exercises.

Presenter: DROSATOS, George (ATHENA RC)