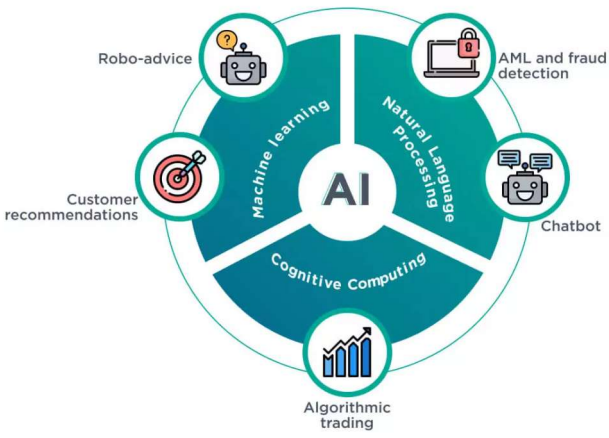

EmergeFlow
SMALL DIFFERENCE. LARGE IMPACT.

Case study:

AI-enabled Financial Market Knowledge Intelligence platform using Graph RAG and MCP server

Client Background

The client is a product-focused IT company headquartered in India, with operations spanning the Nordics and Dubai. Established in 2013, the company specializes in developing advanced software solutions for Wealth Management and Insurance sectors. Its flagship products include a comprehensive portfolio and asset management system for wealth managers, discretionary mandates, private banking, and family offices, and Huston Insurance, catering to the insurance industry.

NEED OF THE CLIENT

- Videos
- Social Media
- Podcasts
- Reports
- Blogs
- News Updates
- Articles

The client required an end-to-end automated system to collect financial news from multiple sources, extract meaningful market events from the raw content, and expose the structured intelligence through a query able API and MCP layer. The key requirements were:

- Continuously scrape financial news from multiple reliable sources on a daily basis
- Automatically identify and extract market-relevant events, entities, and relationships from raw article content using AI
- Store the structured intelligence in a graph database to preserve causal relationships between events, sectors, regions, companies, and people
- Expose the knowledge base through natural language APIs so portfolio managers could query market intelligence without reading raw articles

BEFORE IMPLEMENTATION

Before this system was built, the client relied on manually reading financial news and summarizing market events. There was no structured mechanism to track which sectors, regions, or companies were affected by specific events, nor any way to quantify the sentiment or magnitude of those effects. Identifying patterns across hundreds of articles was time-consuming, inconsistent, and heavily dependent on individual analyst bandwidth. Querying historical market intelligence required manual effort and produced inconsistent results.

AFTER IMPLEMENTATION

After implementing the AI-enabled Market Knowledge Engine, a four-phase production pipeline was delivered covering data ingestion, knowledge extraction, and natural language querying.

Phase 1 built a robust multi-source scraping pipeline collecting articles daily from financial news sources, with hash-based deduplication, content cleaning, and three-layer logging ensuring data integrity and operational visibility.

Phase 2 implemented an LLM-based event extraction engine that reads every ingested article, identifies market-relevant events, and writes structured nodes and relationships into a Neo4j graph database. Every relationship in the graph carries analytical properties, sentiment scores ranging from -1.0 to +1.0, impact reasons, and confidence scores - enabling precise querying across sectors, regions, companies, and named individuals. The schema was deliberately designed to support reliable AI-generated graph traversal queries, with all entities connected through event nodes to preserve causal context.

Phase 3 delivered a hybrid Graph RAG query layer combining LLM-generated Cypher graph traversal with vector similarity search, exposed through a secure JWT-authenticated REST API.

Further implementation added an MCP server with tools which perform the tasks as the API. For a given question, instead of accessing each endpoint to gain specific information, the LLM decides the most appropriate tool for the task, runs the RAG pipeline and presents the results as given by the execution.

RESULTS

The system successfully automated the end-to-end pipeline from raw news ingestion to structured market intelligence, eliminating manual summarization effort entirely. The hybrid retrieval approach, combining precise graph traversal for structured analytical queries with semantic vector search for broader contextual coverage, ensured reliable and accurate responses across diverse query types. The graph database schema preserved causal relationships across hundreds of articles, enabling queries that span multiple events, sectors, regions, and time periods simultaneously, something no manual process or traditional database approach could support efficiently.

The delivered system provides the client's portfolio managers with on-demand, plain English access to a continuously updated financial market knowledge base, significantly reducing the time and effort required to monitor market developments and assess portfolio impact.