DBA5102 Business Analytics Capstone Project

Roy Yeo Fu Qiang (A0280541L)

Investment Context Engine (ICE): A Lightweight Graph-Based Context-Aware AI Backbone for Hedge Fund Workflows

1. Executive Summary

Problem: Lean hedge funds drown in fragmented, unstructured data (filings, transcripts, notes, news), but lack the infrastructure and manpower of large investment funds. This causes missed early signals, low insight reuse, inconsistent context, and manual triage bottlenecks, leaving small funds at structural disadvantage versus larger peers.

Opportunity: Frontier AI (LLMs, RAG) is commoditised; the true edge is <u>context engineering</u>—encoding the fund's proprietary knowledge, drivers, and mental models into a reusable, explainable, and compounding system that can ground the LLMs or AI tools into giving more meaningful outputs.

Solution: ICE — a graph-native, MCP-ready context engine integrating external + internal data into a domain-specific knowledge graph queried via Lazy Graph-RAG + Hybrid Retrieval**.**

- Grounds LLM reasoning in fund-specific context.
- Delivers traceable**,** causal**,** decision-ready insights for research, portfolio monitoring, and risk.

2. Key Differentiators

Spine, not muscle: Build proprietary cognitive core, turning fragmented noise into a compounding, explainable knowledge assets that powers sharper and more confident investment decisions; buy external tools (plug in via MCP).

Lazy Graph-RAG: On-demand, sparse, high-signal subgraph retrieval; minimal upfront KG build.

Hybrid Retrieval: Vector + keyword + HyDE + graph traversal

Full Traceability: Every claim has timestamp, source, evidence snippet; unsupported facts dropped.

Single deterministic agent: Consistent extract → reason → synthesize pipeline.

3. MVP Scope (~4 to 6 months)

Coverage: S&P 500 equities; focus on **portfolio monitoring** + **fundamental research**.

<u>Edge Types (MVP)</u>

1. **Ownership Context Edge:** *Portfolio → holds → Company*
2. **Thematic Edge:** *Company → exposed_to → Theme*
3. **Anchor KPI Edge:** *Company → key_driver → KPI*

Hop Depth: Incremental 1-hop → 2-hop → 3-hop retrieval with forward/reverse traversal, temporal/confidence scoring.

UI Modules:

- Ask ICE (QA w/ causal paths + citations)
- Per-Ticker Intelligence Panel ("What Changed", drivers, exposures)
- Mini Subgraph Viewer (1-3 hops)
- Daily Brief Table (portfolio/watchlist)

4. Business Value

Time savings: ≥30 min/day per analyst → faster triage and more focus on alpha generation.

Alpha: Minutes-earlier signal capture can drive 1% profit difference, amounting to ~$1M for $200M book; studies show +3–5% annualized abnormal returns from GenAI adoption.

Risk Control: Human-in-the-loop, evidence-backed insights → reduces costly hallucinations.

Institutional Memory: Knowledge compounds with every query, note, and data ingestion.

5. Build Plan & Feasibility

Phased Roadmap:

1. Basic RAG MVP
2. Hybrid Retrieval + Lazy Graph Links
3. Context Processing + Synthesis Refinement
4. Deployment (batch briefs + UI)
5. Batch KG Enrichment

Tech Stack: Python + FAISS + NetworkX + local embeddings + GPT (selective) + Streamlit UI.

Zero-ops: Laptop-scale, open-source, no backend infra needed for MVP.

6. Risks & Mitigation

Risk	Mitigation
Lack of Data	Start with public, free corpora. Ingest portfolio CSVs, utilize synthethic data for development and to showcase business value.
Sparse / Stale KG ***	Seed a minimal high-signal schema (e.g. portfolio or sector information); combine with web search tools
Lack of Budget	Lazy Graph expansion over full KG build, consider NetworkX over Neo4J for graphs storage; reduce scope; prioritise local LLMs
Hallucination Risk	Main objective of adopting a Graph-RAG with traceability and human-verification in workflow
Token bloat / Over-reliance on external LLMs	Retrieval plan with budgets: per-hop top_k caps, MMR de-dup, recency & confidence filters; hierarchical extractive summaries.
Conflicting views	Don't auto-resolve; show both - side-by-side narratives
Regime drift	Temporal metadata, rolling windows in retrieval

7. Possible / Future Extensions

- Using hallucinated graph structure to assist in building up the graphs (HyDE inspired)
- Full Graph-RAG
- Chat with conversational memory (context retention)
- Expanded data coverage (structured fundamentals, alt-data, internal trade logs)
- Multi-agent framework