Graphs + Knowledge Tables + LLMs
Unlocking Claims Intelligence at Suncorp

A focused presentation on GraphRAG & Structured RAG

1. Motivation

• Suncorp has massive structured data (SQL: policies, claims, payments)

• And massive unstructured data (emails, PDFs, reports)

• Today's RAG: retrieves chunks but misses relationships

• GraphRAG: entity + relationship reasoning + LLM = explainable insights

2. What is GraphRAG?

• Retrieval Augmented Generation with a Knowledge Graph backbone

• Entities become nodes, relationships become edges

• Query → extract entities → retrieve subgraph → combine with vector retrieval

• Example: 'Show all claims linked to Provider X before Policy Y started'

3. Structured RAG + Graph

• Structured DB → Graph nodes (Policy, Claim, Customer, Provider)

• Unstructured docs → Graph edges & attributes (emails, PDFs, notes)

• Potential: Faster retrieval, SQL joins, provenance tracking, auditability

4. Math from HybridRAG (Sarmah et al., 2024)

Vector retrieval:

$V_q = top_k(sim(f_{emb}(q), f_{emb}(c)))$

Graph retrieval:

$G_q = hop(E_q, h) \rightarrow$ subgraph of entities & neighbors

Combined context:

$C_q = V_q \cup Docs(G_q)$

LLM generation:

■ = LLM(q, C_q)

5. Toy Example (Suncorp Claim)

Query: 'Has Provider X submitted claims >$50K before Policy Y started?'

• VectorRAG → finds emails/docs mentioning Provider X, claim amounts

• GraphRAG → Policy Y → Claim → Provider X chain with dates/amounts

• Combined: LLM answers with provenance: structured DB + email snippet

6. Architecture Sketch

Flow: Structured DB (SQL) + Unstructured Docs → Entity Extraction → Graph Builder → Hybrid Retrieval → LLM with provenance

7. Takeaways for Suncorp

• Explainability: trace answers to SQL + docs

• Compliance: relationships are explicit

• Efficiency: hybrid retrieval is faster and more precise

• Scalability: applicable across 120+ AI initiatives