

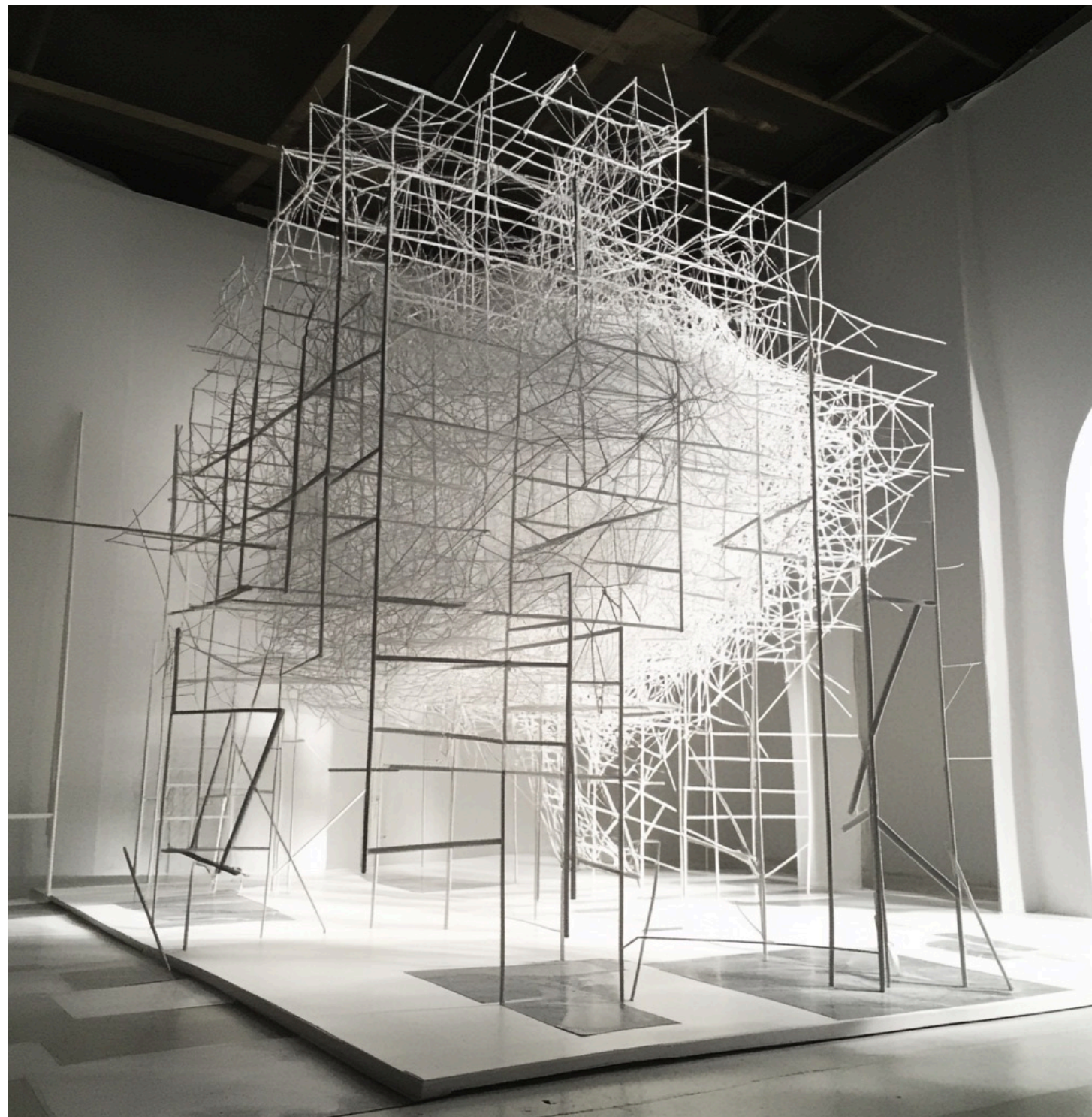


Reasoning in Clarity: The Power of GraphRAG

When documents Aren't Enough...

Dec 24, Özgür Güler

Overcoming Challenges with BaseLine RAG

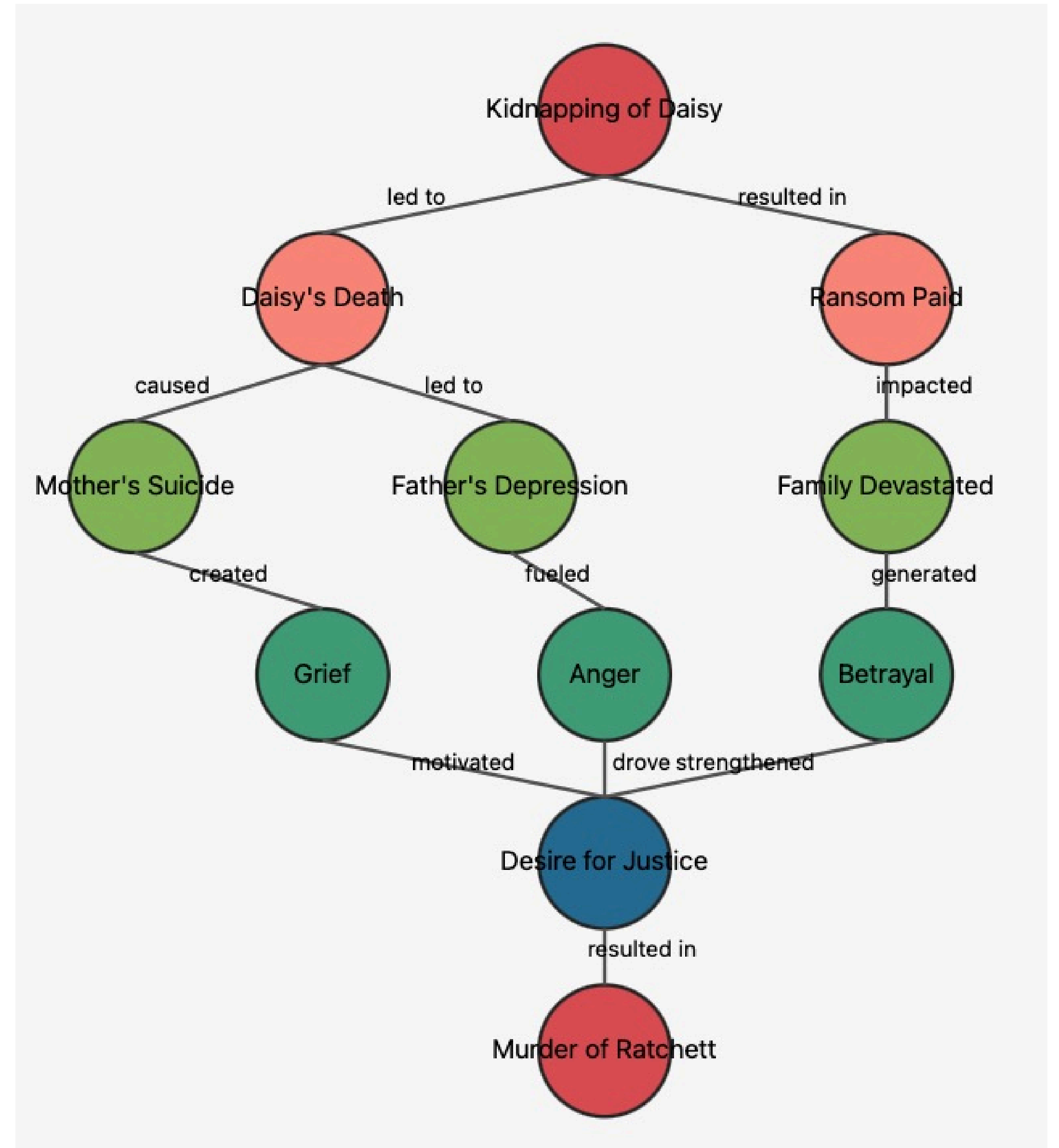


Sparse & Complex Relationships in data

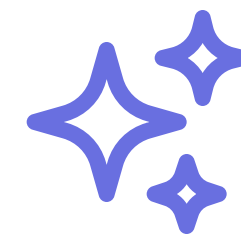
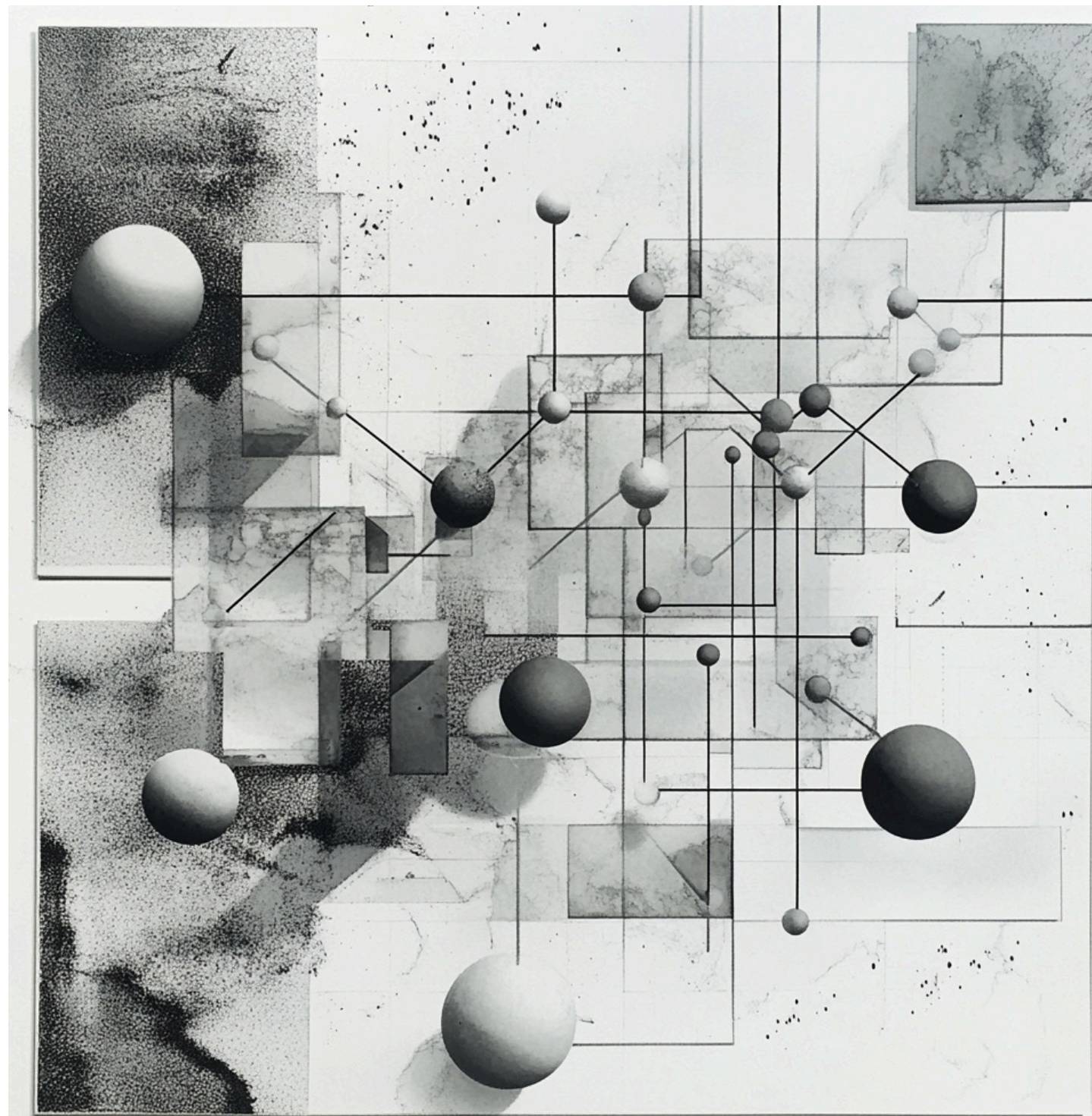
- Knowledge scattered across documents
- Hidden connections missed by similarity
- Global Search e.g. “What are the top 5 themes” in the data...

Knowledge Graph Murder on the Orient Express

Encode information in a graph and feed the context generated by graph retrieval to the LLM



Overcoming Challenges with BaseLine RAG



Advanced Reasoning

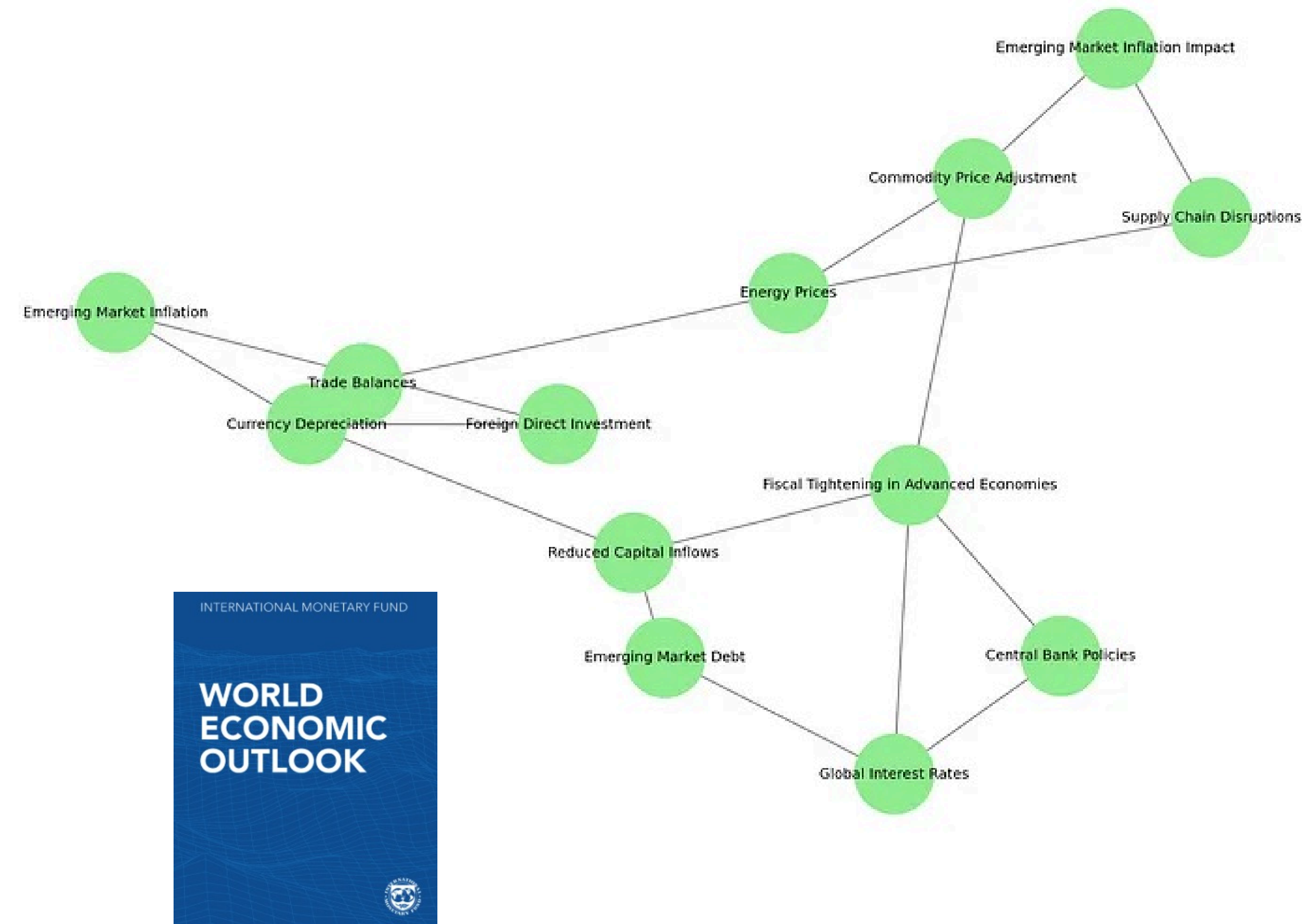
- multi-hop reasoning
- Requires connecting multiple facts
- Exploring deeper “causal pathways”

Exploring Deep “Causal Pathways” in a KG

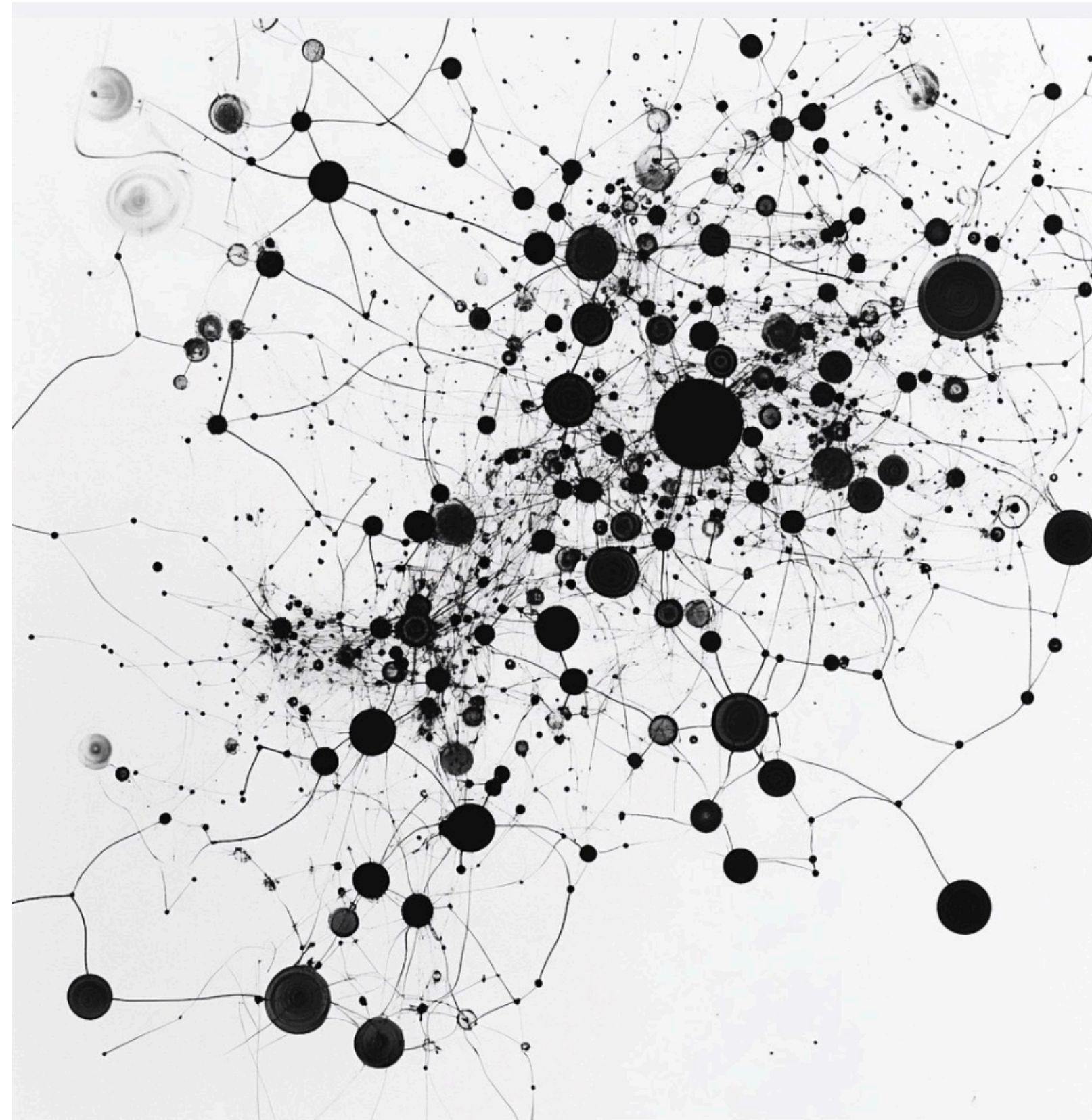
How does fiscal tightening in Advanced Economies affect Emerging Market economies?

- Path A: “Fiscal Tightening” → “Reduced Capital Inflows” → “Currency Depreciation” → “Emerging Market Inflation.”
- Path B: “Fiscal Tightening” → “Commodity Price Adjustment” → “Energy Prices” → “Trade Imbalances” → “Emerging Market Inflation Impact.”

Extended Knowledge Graph: Fiscal Tightening and Global Economic Impact



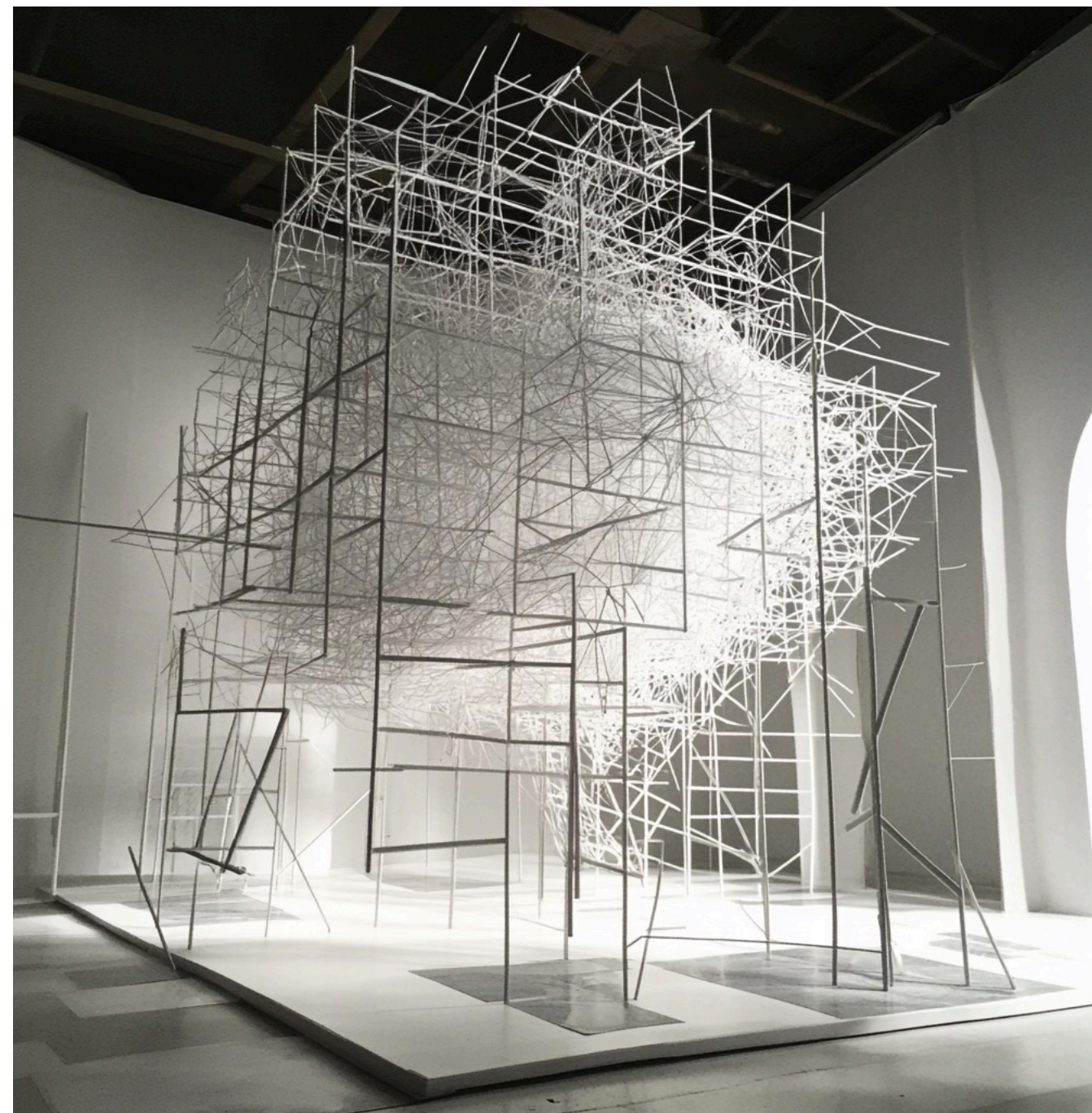
Overcoming Challenges with BaseLine RAG



Structural Information

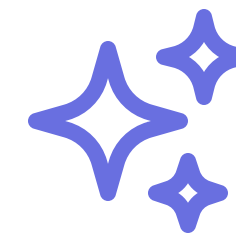
- Hierarchical relationships that are not explicit
- Organizational dependencies
- Network of interconnected facts

Overcoming Challenges with BaseLine RAG



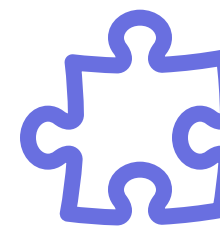
Sparse & Complex Relationships in data

- Knowledge scattered across documents
- Hidden connections missed by similarity
- Global Search “What are the top 5 themes in the data”



Multi-hop Reasoning

- Requires connecting multiple facts
- Complex logical chains a.k.a “causal pathways”



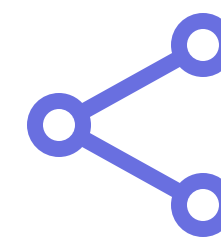
Contextual Ambiguity

- Same terms, different meanings
- Context dependent relationships



Structural Information

- Hierarchical relationships required with “Global Search”
- Organizational dependencies
- inferring structure from scattered incomplete information e.g. build your family tree by reading the diaries of your ancestors...



Transitive Properties

- If $A \rightarrow B$ and $B \rightarrow C$, then $A \rightarrow C$ connections
- Indirect relationships & influences e.g. protein synthesis or medical side effects detection

Knowledge Graphs - Building the scaffolding for your information base



Information soup to KG's

- Entities become nodes
- Relationships become edges
- Properties become attributes
- Context becomes subgraphs



Benefits

- Explicit relationship modelling
- Structured knowledge representation
- Navigable information space
- Preserves semantic connections

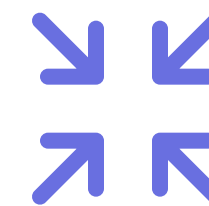


When to use Graph RAG?



Grounding Generations

Where accuracy is more important, data is sparse, multi-domain or difficult to extract e.g. clinical decision support, rare disease diagnosis, cost of hallucinations is extremely high...AdaHealth...SLM's?

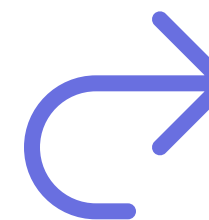


Advanced Reasoning

Causal ML

Where advanced reasoning requires exploring deeper cognitive pathways in the graph...

e.g. Scientific research, legal reasoning, supply chain optimisation, cybersecurity analysis, drug discovery...



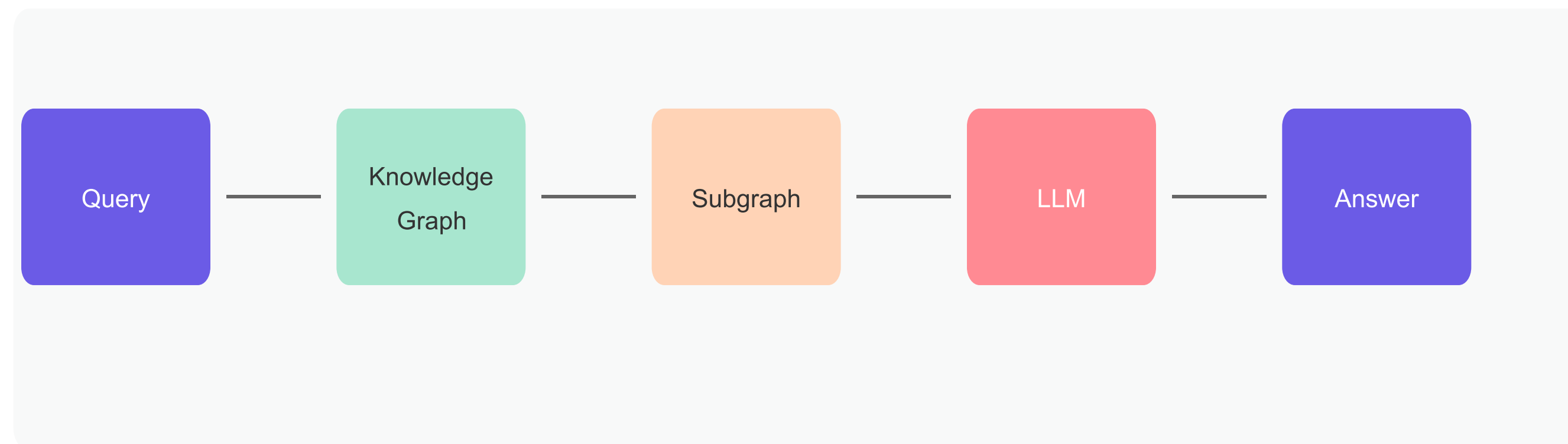
Dynamic knowledge Evolution

- Evolving relationships between entities
- Temporal dependencies and changes
- Emerging patterns and connections
- Knowledge graph updates and versioning

What is GraphRAG?

GraphRAG is LLM generations with Graph context...

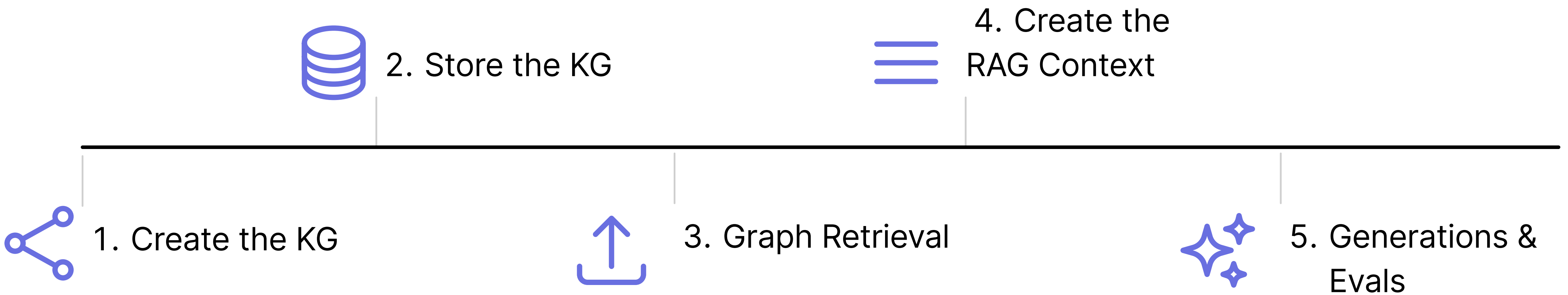
- extracting query relevant context from a KG
- feeding it to an LLM as context...



How it works:

1. Takes your query
2. Finds relevant subgraphs in the knowledge graph
3. Converts these relationships into rich context
4. Uses an LLM to generate an informed answer

Implementing Graph RAG...



1. Create your KG

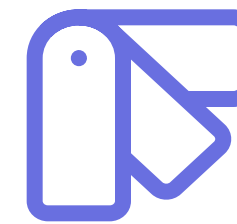
Tools/methods to create your KG



MS Research GraphRAG accelerator library



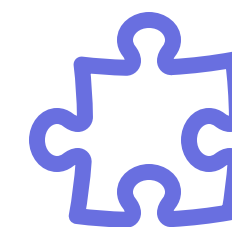
Neo-4j LLM Graph Builder Tool



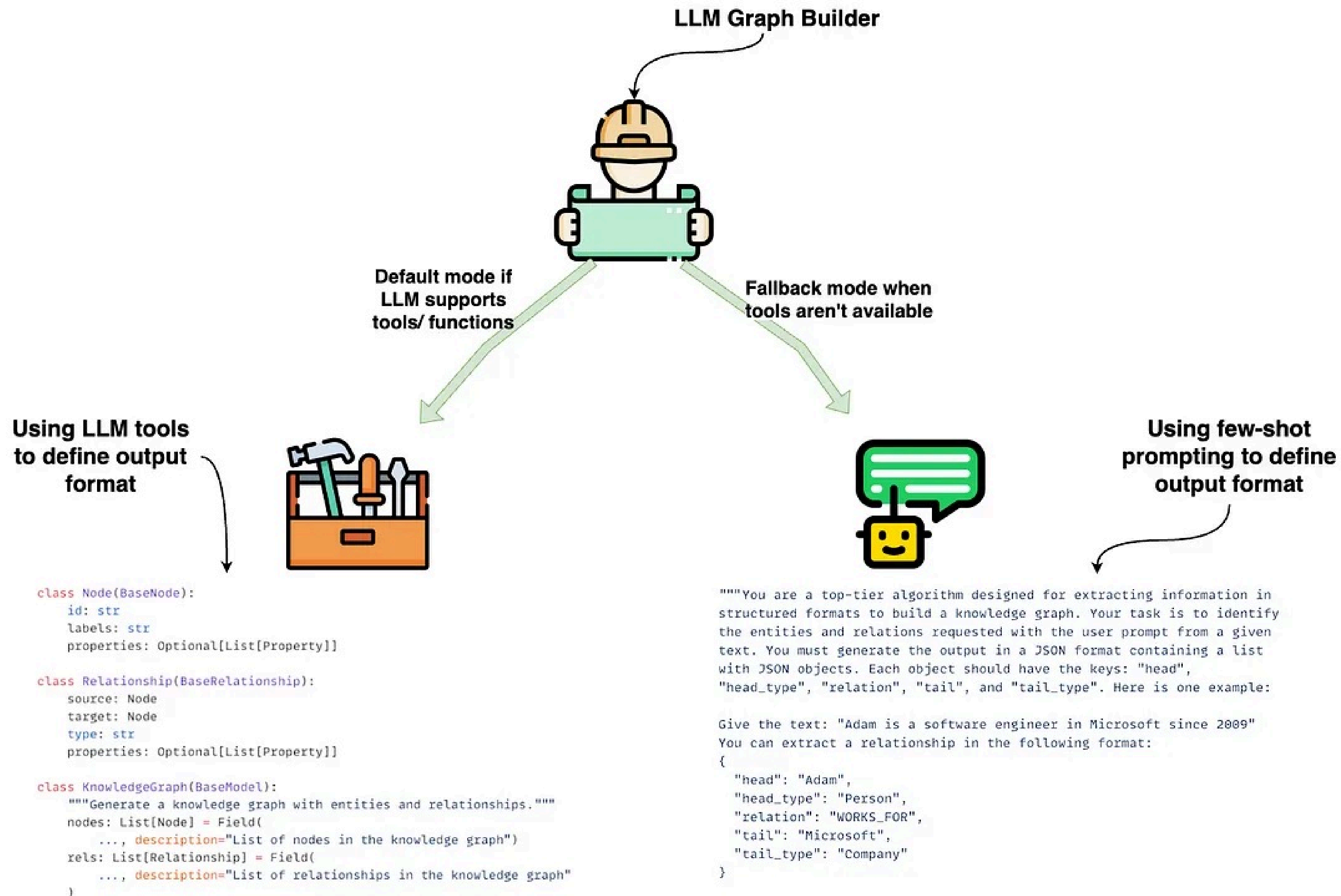
Langchain LLM Graph Transformer



LLama Index LLM Graph Transformer



Prompt Engineer your way



2. Graph DB vs Graph Index

2. Store your KG

Choice depends on...

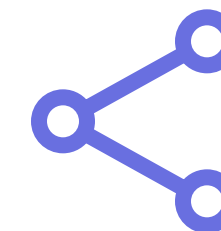
- Scale of your knowledge graph
- Query patterns
- Performance requirements
- Development complexity tolerance
- Infrastructure budget



Document based runtime graphs

Save Graph Index as a document RDF / XML, Turtle, JSON-LD files & build at runtime with rdflib (Cosmos AI Graph) or with Apache Spark GraphFrames ...

- No native indexing
- Memory constraints
- Query Performance



Native Graph DB's or RDF stores

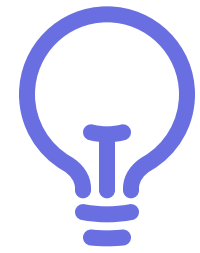
- Optimised for graph operations
- Built-in indices
- Built-in reasoning with RDF stores



Unified Indexing

- Pinecone, weaviate
- Vectors + edge information
- Limited Graph Operations
- Evolving category

3. Graph Retrieval



MS Graph RAG

Uses hierarchical community detection (Leiden algorithm) to create structured knowledge summaries at different levels.

- emphasizes holistic understanding and theme detection
- focuses on hierarchical understanding and community structure
- 3 modes Global Search (community summaries), Local Search (entity neighborhoods), DRIFT Search (combines both)

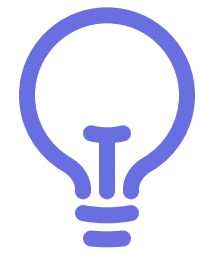


PyG GRetriever

Uses Prize-Collecting Steiner Tree (PCST) optimization to find relevant subgraphs, focusing on maintaining connectivity while retrieving information

- Focuses on question-answering and conversation capabilities
- Integrates Graph Neural Networks (GNNs) with LLMs more directly, using trainable graph encoders and projection layers...

Query: "How does the spice melange connect the Fremen, the Bene Gesserit, and House Atreides?"



MS Graph RAG

1. Hierarchical Communities:

- Power Structures (Houses, Bene Gesserit, Spacing Guild)
- Arrakis Ecosystem (Sandworms, Spice Production, Fremen)
- Prescience Groups (Navigators, Kwisatz Haderach, Paul's Visions)

Search Modes:

- Global Search: The spice's role in shaping galactic civilization
- Local Search: Direct spice interactions of each group
- DRIFT Search: How spice connects political and religious powers

Example Community Summary: "Spice connects these groups through prophecy (Bene Gesserit), survival (Fremen), and destiny (Atreides), forming the central axis of power on Arrakis."



PyG GRetriever

1. PCST Relevant Subgraph:

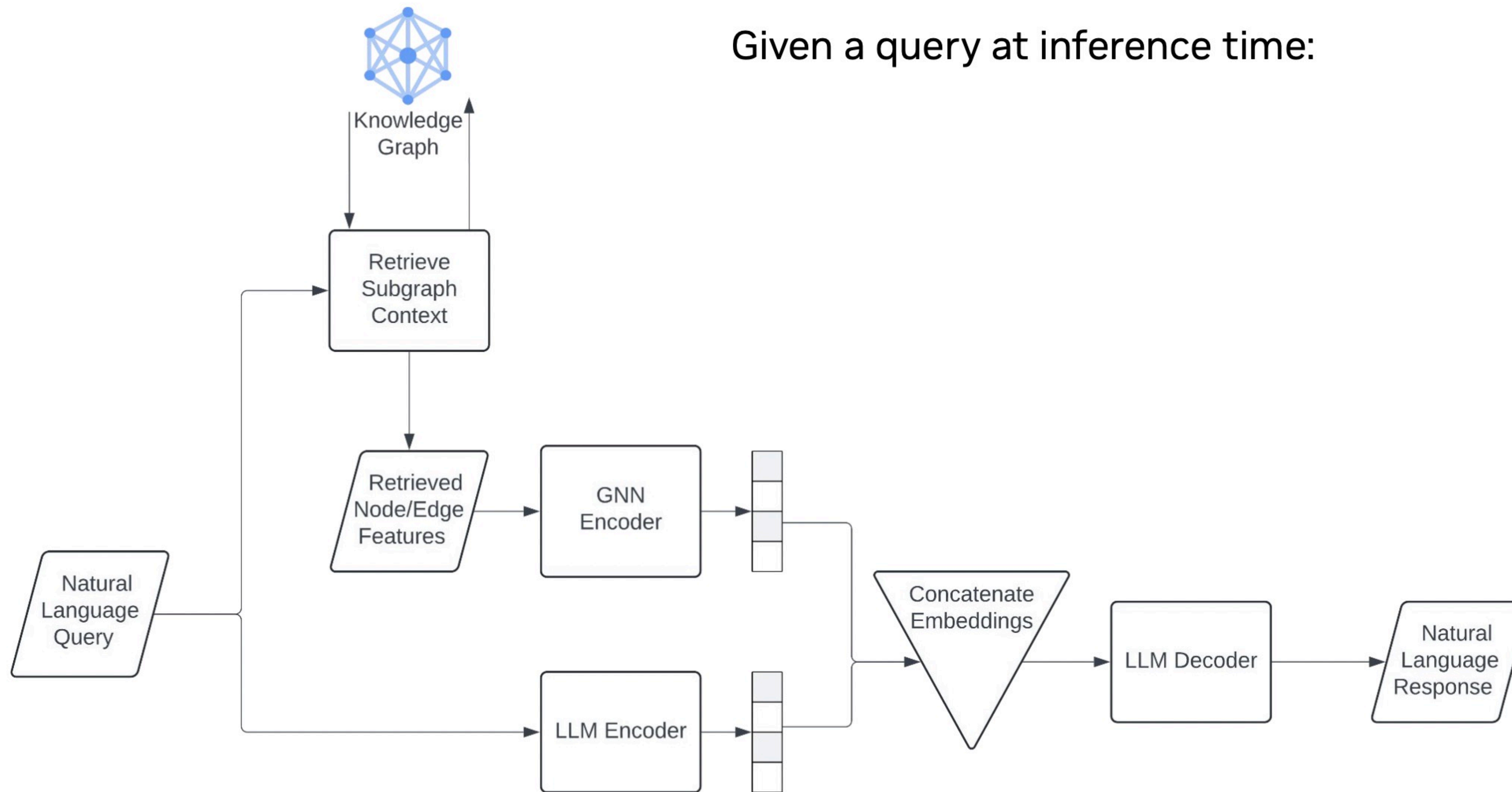
- Core nodes:
 - Spice melange
 - Fremen spice rituals
 - Bene Gesserit prophecies
 - Paul Atreides' visions
- Key connections:
 - Water of Life ceremonies
 - Desert survival techniques
 - Guild navigation requirements
 - Prescient abilities

Direct Evidence Path: Shows precise connections like: Fremen → spice collection → water of life → Bene Gesserit testing → Paul's transformation → fulfillment of prophecy

"The spice melange forms a direct chain of influence: Fremen harvest and use it in their water of life ceremonies, which the Bene Gesserit adapted for their own testing rituals. Paul Atreides, product of the Bene Gesserit breeding program, gains prescient abilities through these spice rituals, ultimately uniting all three groups under his leadership as the Kwisatz Haderach."

General Graph Based RAG Workflow

Given a query at inference time:



Resources

- Microsoft Research GraphRAG portal [\[link\]](#)
- Microsoft GraphRAG global query sample notebook [\[link\]](#)
- Microsoft Research GraphRAG paper [\[link\]](#)
- Neo4j Graph Builder [\[link\]](#)
- G-Retriever paper [\[link\]](#)
- PyG GReTriver examples [\[link\]](#)
- CosmosAI Graph [\[link\]](#) [\[MSLearn\]](#)
- Hybrid RAG paper [\[link\]](#)
- Graph Attention Networks paper [\[link\]](#)
- The Missing Piece in Graph RAG: Graph Attention Networks [\[link\]](#)



Thank you!