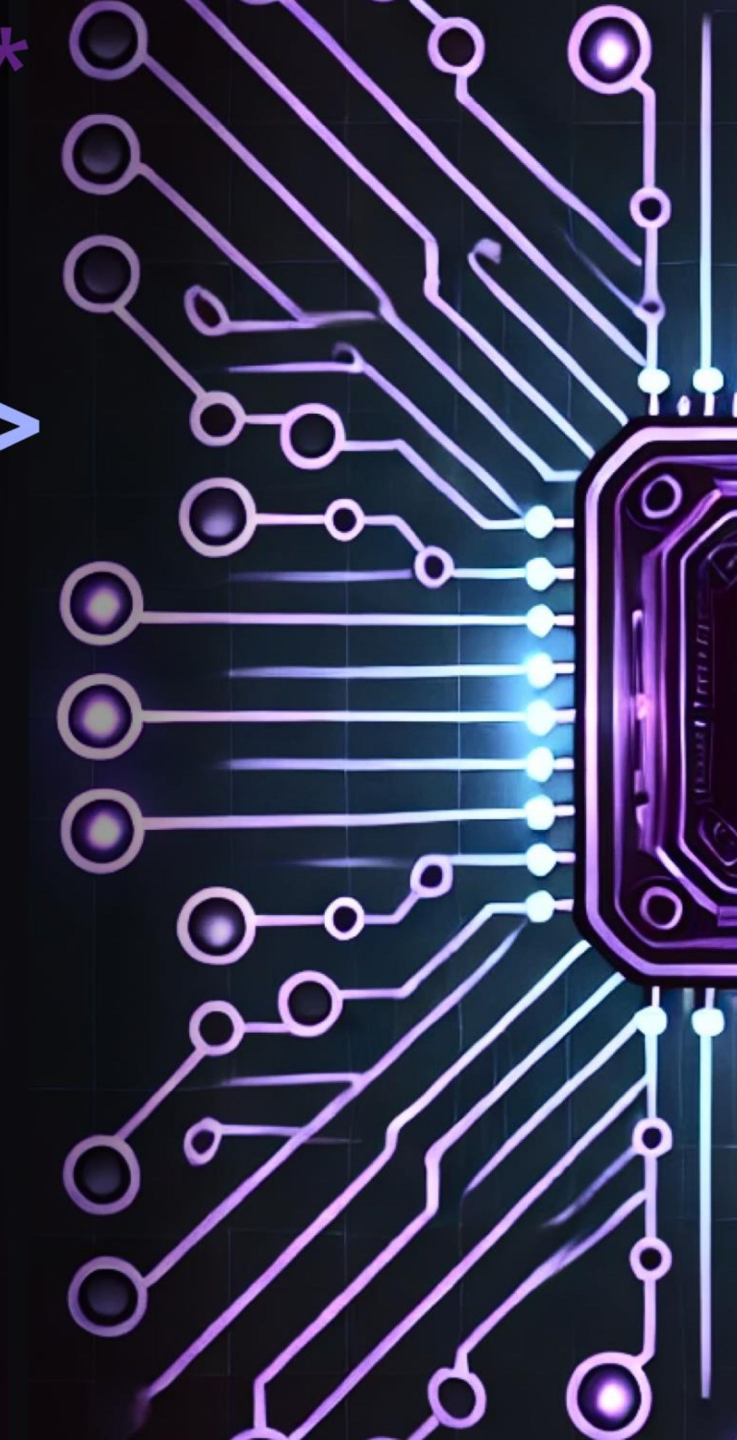
September 3rd - 16th

Hack Together:
RAG Hack

Build, innovate, and #HackTogether
aka.ms/RAGhack



Hack Together: RAG Hack

aka.ms/RAGhack

Week 1: September 3 - 8

9 MONDAY	3 TUESDAY	4 WEDNESDAY	5 THURSDAY	6 FRIDAY	7 SATURDAY	8 SUNDAY
	RAG 101 **.NET** **Azure AI Studio** **Python**	**Langchain4J** **LangchainJS**	**Responsible AI** **MongoDB** **Azure AI Search** **PostgreSQL** **Azure SQL**	**GraphRAG** **Multi-channels**	**Hack!** **Hack!** **Hack!**	**Hack!** **Hack!** **Hack!**

Week 2: September 9 - 14

9 MONDAY	10 TUESDAY	11 WEDNESDAY	12 THURSDAY	13 FRIDAY	14 SATURDAY	15 SUNDAY
Semantic Kernel **Spring AI** **Vision models** **Internationalization**	**VSCode Agent** **Agentic RAG** **Code Interpreter**	**AI Studio Advanced** **AutoGen** **Data Access Control** **Fine Tuning**	**Model catalog** **Evaluations**	**Hack!** **Hack!** **Hack!**	**Hack!** **Hack!** **Hack!**	**Hack!** **Hack!** **Hack!**

Submissions due September 16, 11:59 PM PT!

Microsoft

Intro to GraphRAG

John Alexander
Sr. Content Developer, Azure AI Apps
Microsoft

Apurva Mody
Principal Program Manager, Office of CTO RD
Microsoft

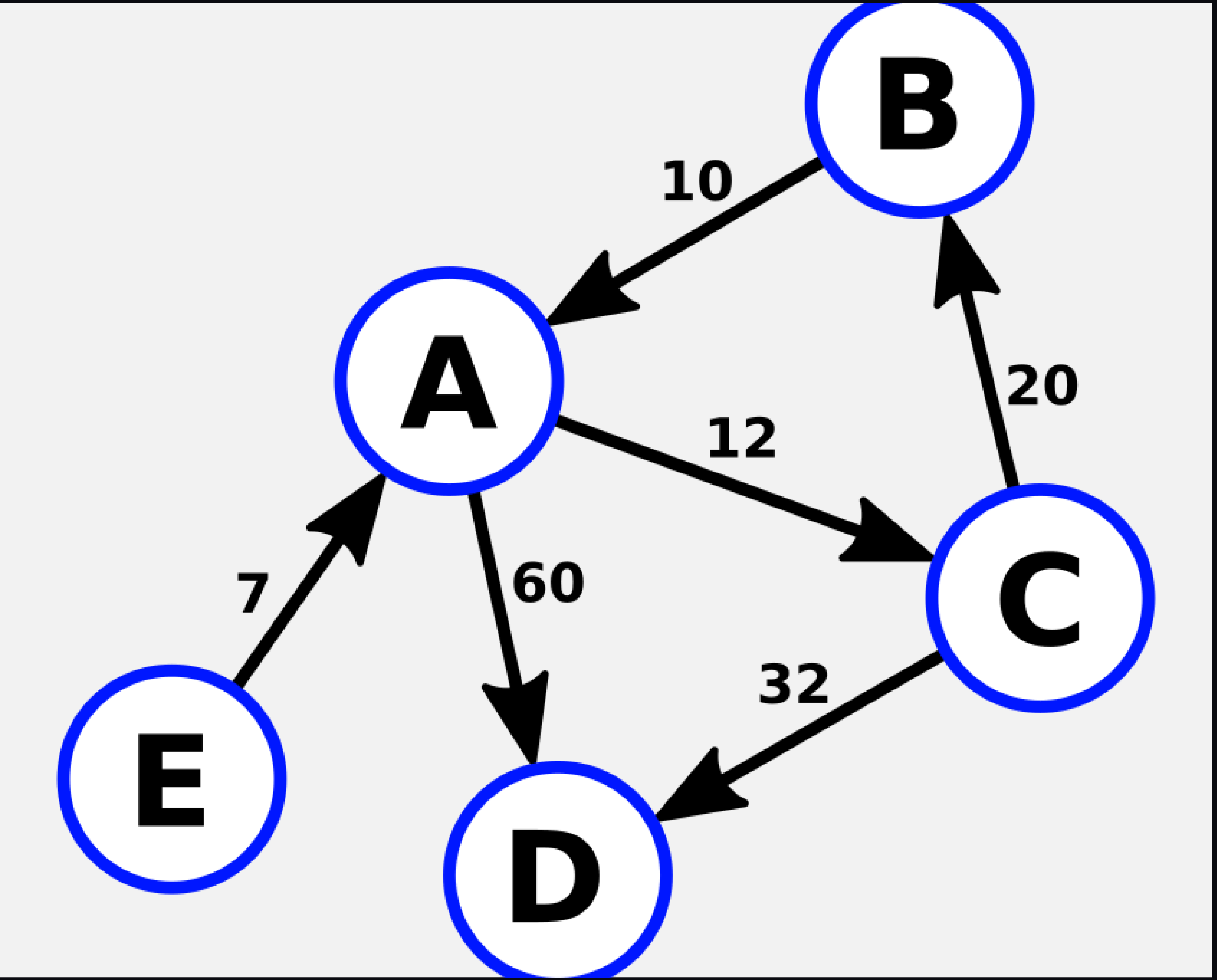
Agenda

- Graph Fundamentals
- Querying the Graph
- RAG With Neo4j and LangChain
- MSFT GraphRAG

Graph

Fundamentals

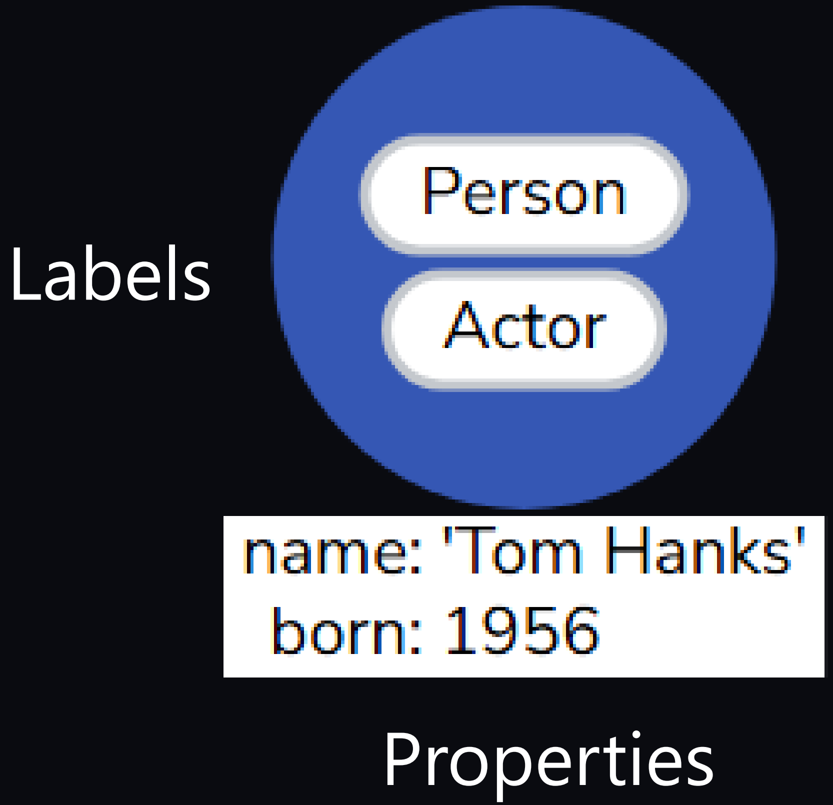
Graph Theory



- Graph is a set of Nodes (Vertices) and Relationships (Edges)

Nodes

- Describe entities (discrete objects) of a domain.
- Have zero or more labels to define (classify) what kind of nodes they are.
- Can have properties (key-value pairs), which further describe them.

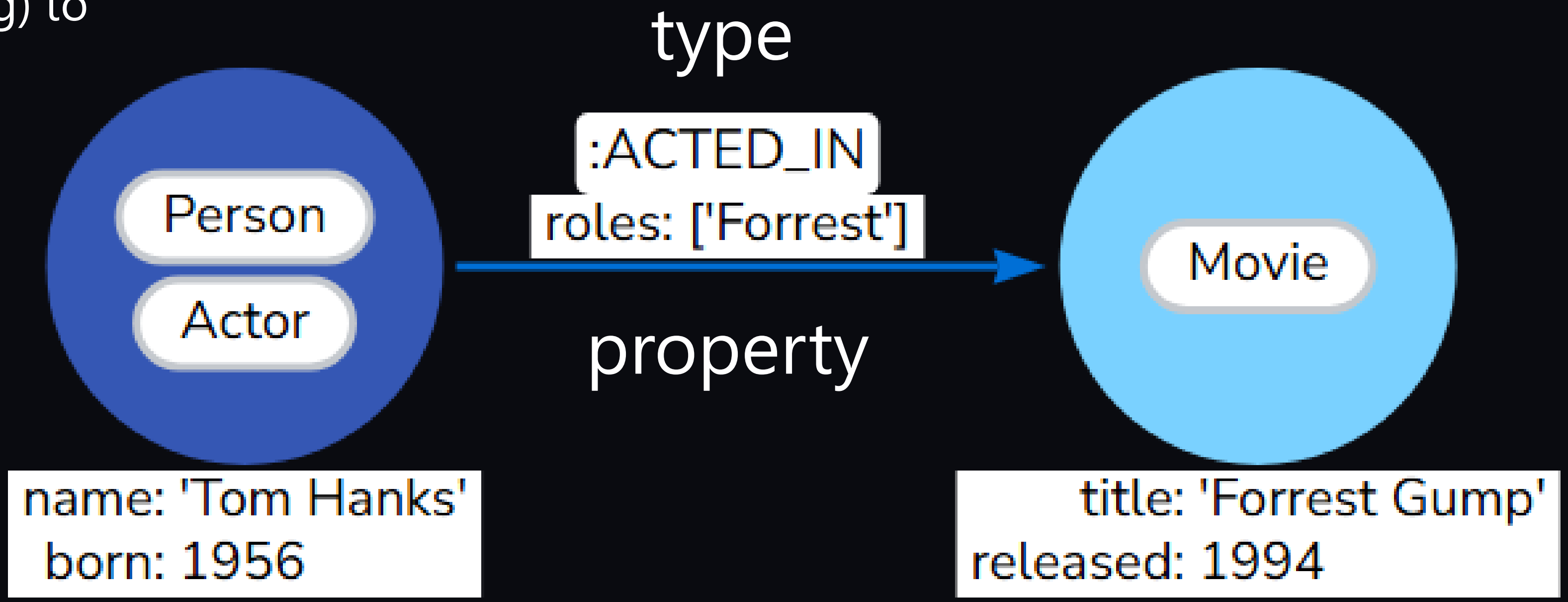


name: 'Tom Hanks'
born: 1956

Image source: neo4j.com/docs/getting-started/_images/graph_single_node-arr.svg

Relationships

- Connect a source node and a target node.
- Has a **direction** (one direction).
- Must have one **type** (ingoing, outgoing) to define the relationship type.
- Can have **properties** (key-value pairs)



type

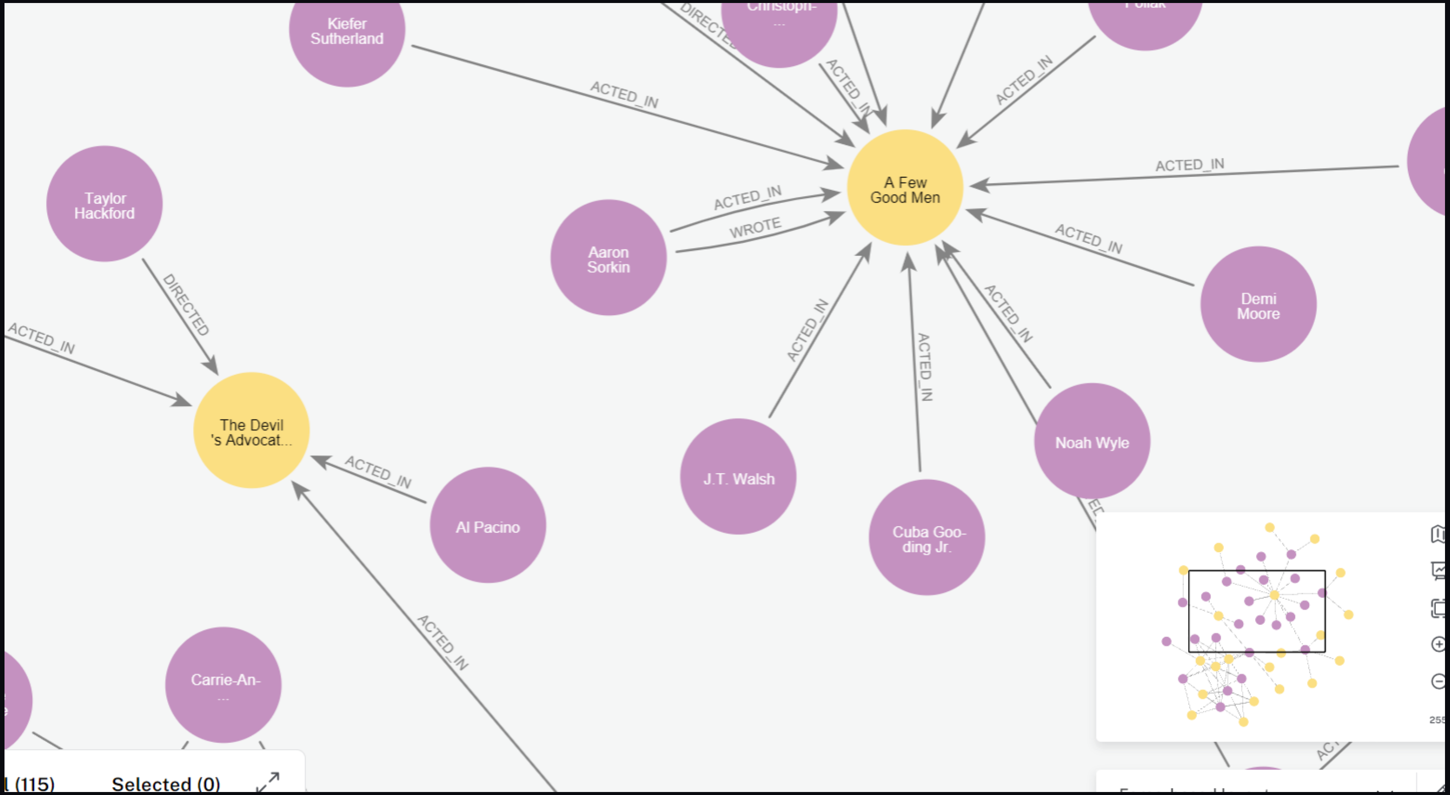
property

Image source: neo4j.com/docs/getting-started/_images/graph_simple-arr.svg

Graph Database

uses graph structures with nodes, edges, and properties to represent and store data.

designed for more efficient data handling with complex relationships and connections than traditional relational databases.

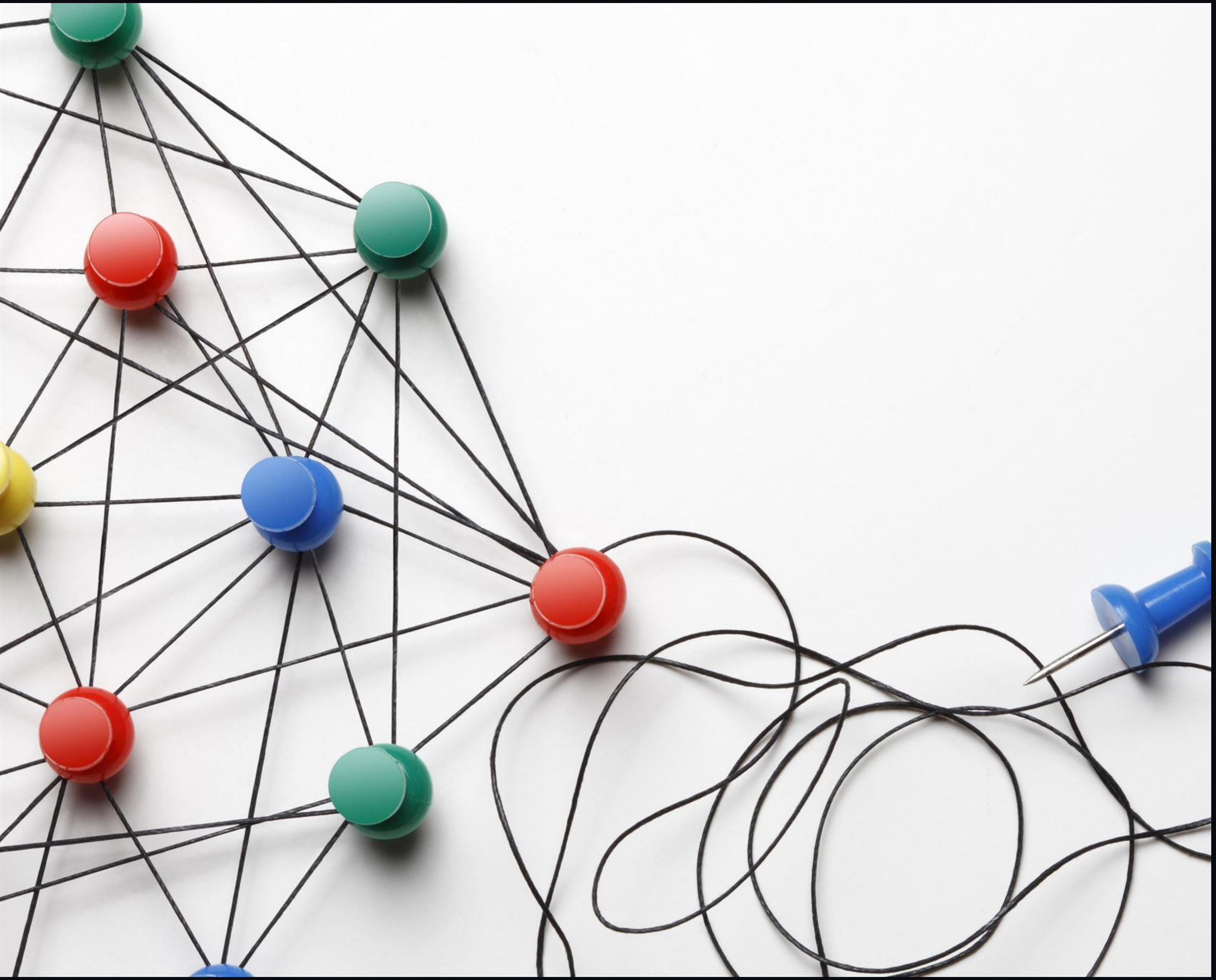


Knowledge Graph is a knowledge base

that uses a graph-structured data model or topology to represent and operate on data

Or

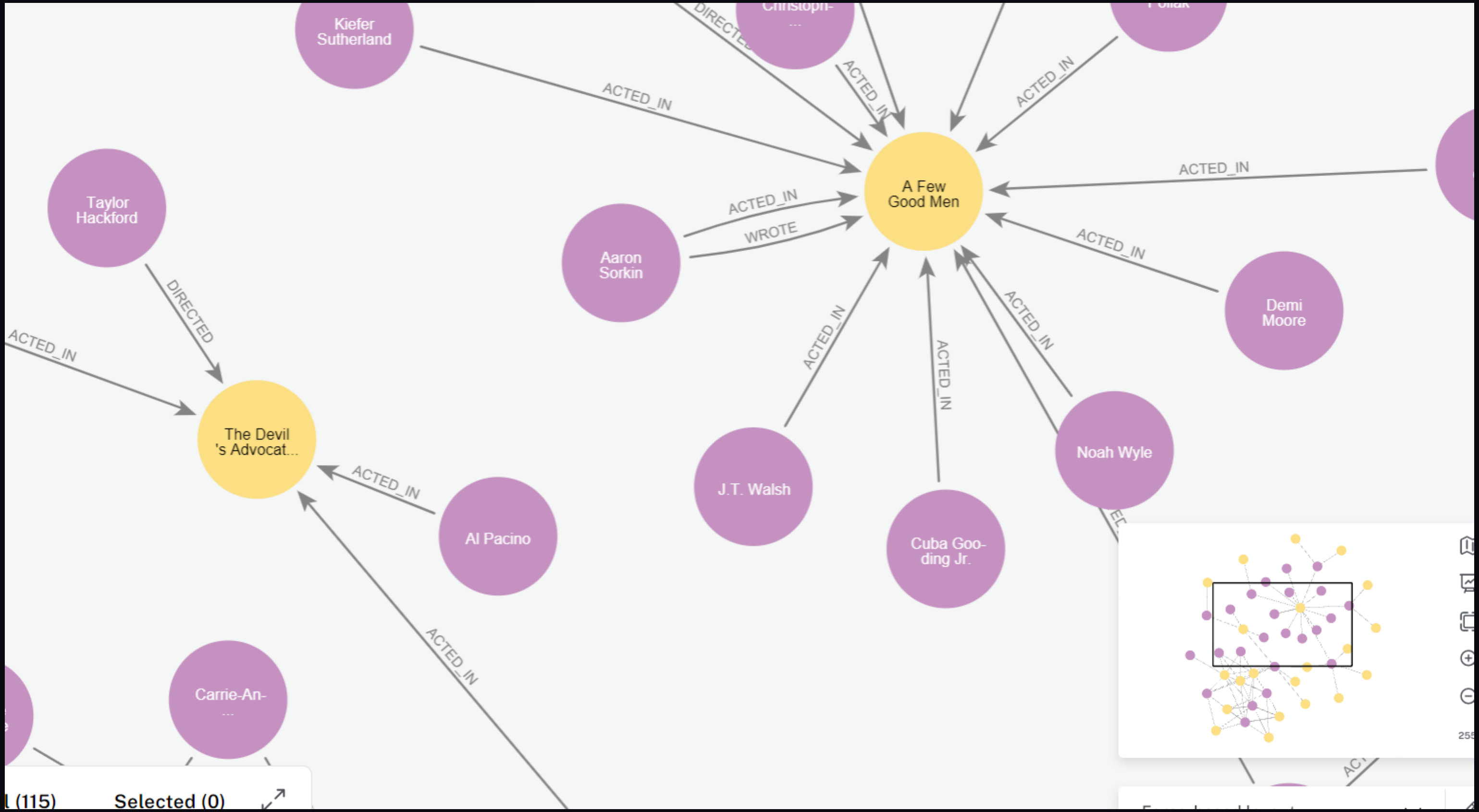
of interconnected entities defined by a schema



Knowledge Graph

why...

- Rich Semantics: Capture complex relationships and concepts
- Flexibility: Schema-less -> makes it easy (ish) to combine multiple data sources
- Enhanced Queries: More powerful and expressive query language for working with densely connected data



Ontology: provides the foundational schema or blueprint that guides how entities (like concepts, objects, or events) and their relationships are represented in the knowledge graph.

Microsoft

Querying the Graph with Cypher - Demo

Ontology + Data = Knowledge Graph

1. **Classes (or Concepts)**: Categories or types of entities within the domain. For example, in a knowledge graph about healthcare, classes might include "Disease," "Symptom," "Medication," and "Patient.“

2. **Instances (or Individuals)**: Specific entities that belong to a class. For example, "Diabetes" might be an instance of the class "Disease," and "Aspirin" an instance of the class "Medication.“

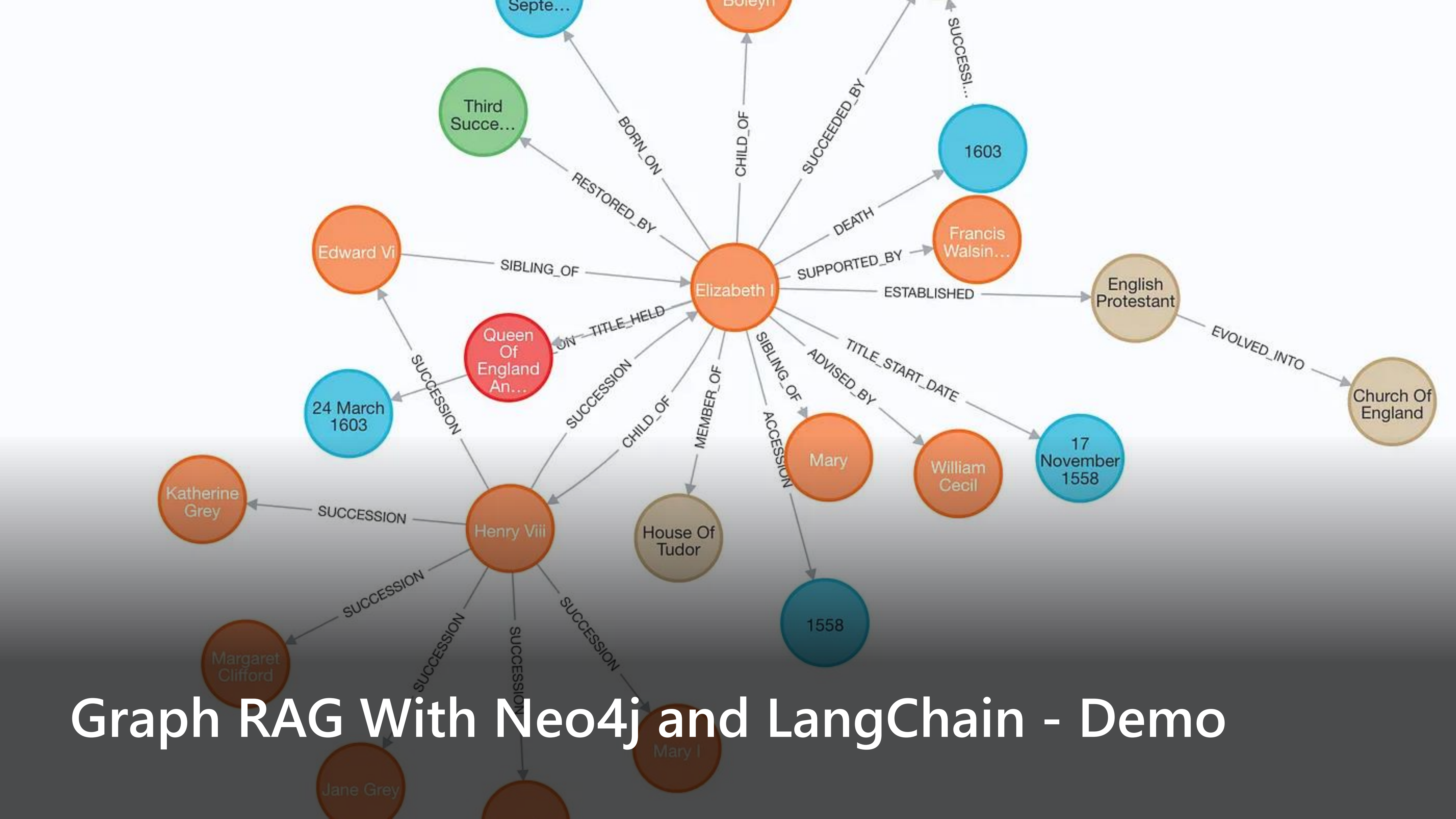
3. **Properties (or Attributes)**: Characteristics or attributes of the classes or instances. For example, the "hasDosage" property might describe the dosage of a medication.

4. **Relations**: Defined connections or associations between classes or instances. These might include hierarchical relationships (e.g., "is a" relationships like "Aspirin is a Medication") or more specific associations (e.g., "treats" relationships like "Aspirin treats Headache").

5. **Rules/Constraints**: Logical rules or constraints that govern the relationships and properties within the ontology. For example, a rule might state that "Every medication must have a dosage."

LangChain and Neo4j

- **Neo4jGraph** - database connector
- **GraphDocument** -Represents a graph document consisting of nodes and relationships.
- **LLMGraphTransformer** – Transform documents into graph-based documents using an LLM.
- **Neo4jVector –** Neo4j vector store (requires neo4j python package install)



Graph RAG With Neo4j and LangChain - Demo

Microsoft

GraphRAG

Baseline RAG

- Uses vector similarity (e.g. ada) for information retrieval on snippets of text (chunks) to add to the context.
- Struggles with complex queries that involve connecting multiple pieces of information
- Often fails to provide relevant context for complex questions that may span the dataset

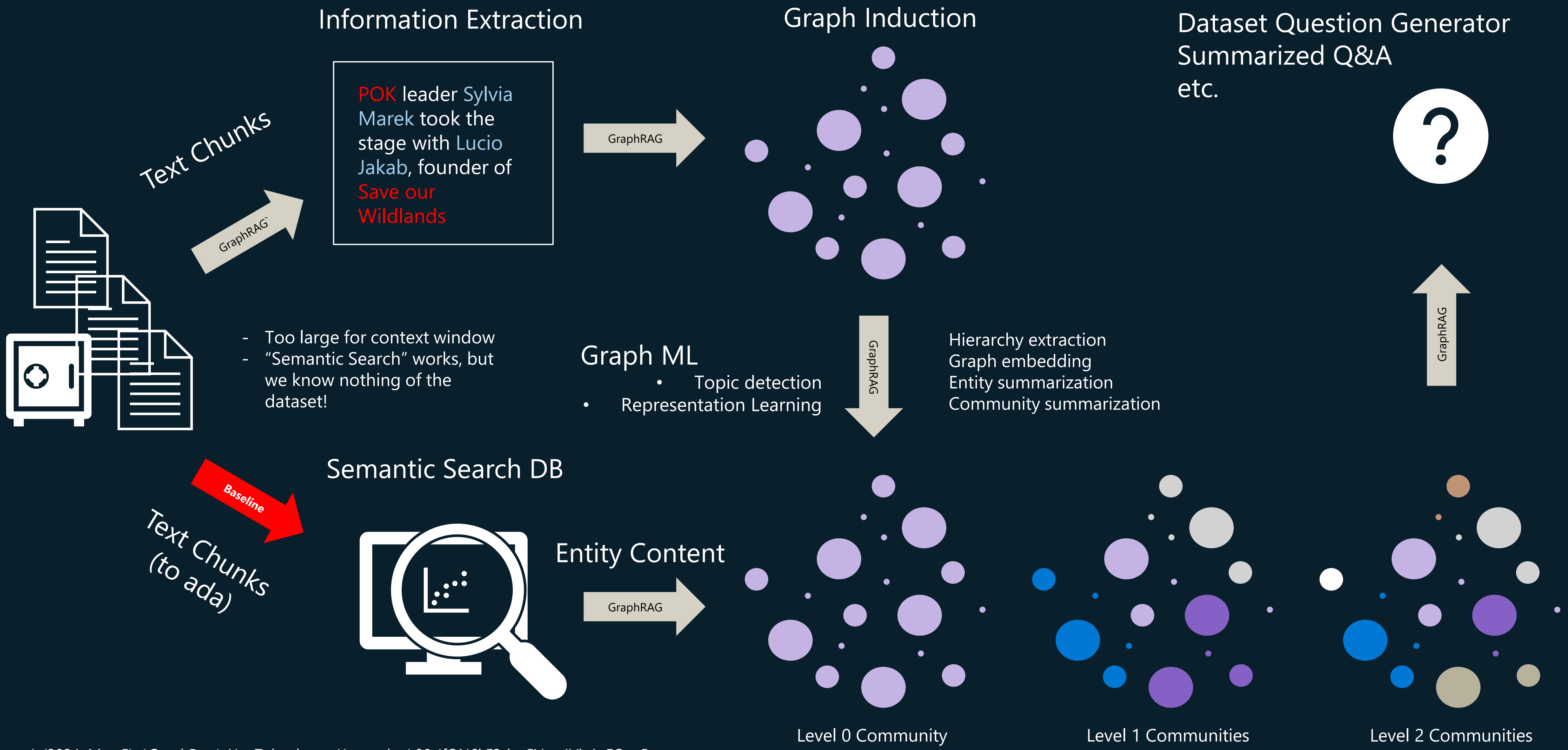
Knowledge Graph + RAG

- Performs a vector similarity search on entities
- Uses a knowledge graph to pull in relevant connected data
- Injects the collected data into the context
- Often fails to provide relevant context for complex questions that may span the dataset

MSFT GraphRAG

- Utilizes LLM-generated knowledge graphs for improved retrieval
- Excels at connecting entities and understanding relationships
- Splits graph into communities and creates summaries to assist search
- Handles whole dataset reasoning, summarizing themes and aggregating information effectively

Comparing Baseline RAG to GraphRAG



Information Extraction

POK leader Sylvia Marek took the stage with Lucio Jakab, founder of Save our Wildlands

Graph Induction

Dataset Question Generator
Summarized Q&A
etc.

- Too large for context window
- "Semantic Search" works, but we know nothing of the dataset!

Graph ML

- Topic detection
- Representation Learning

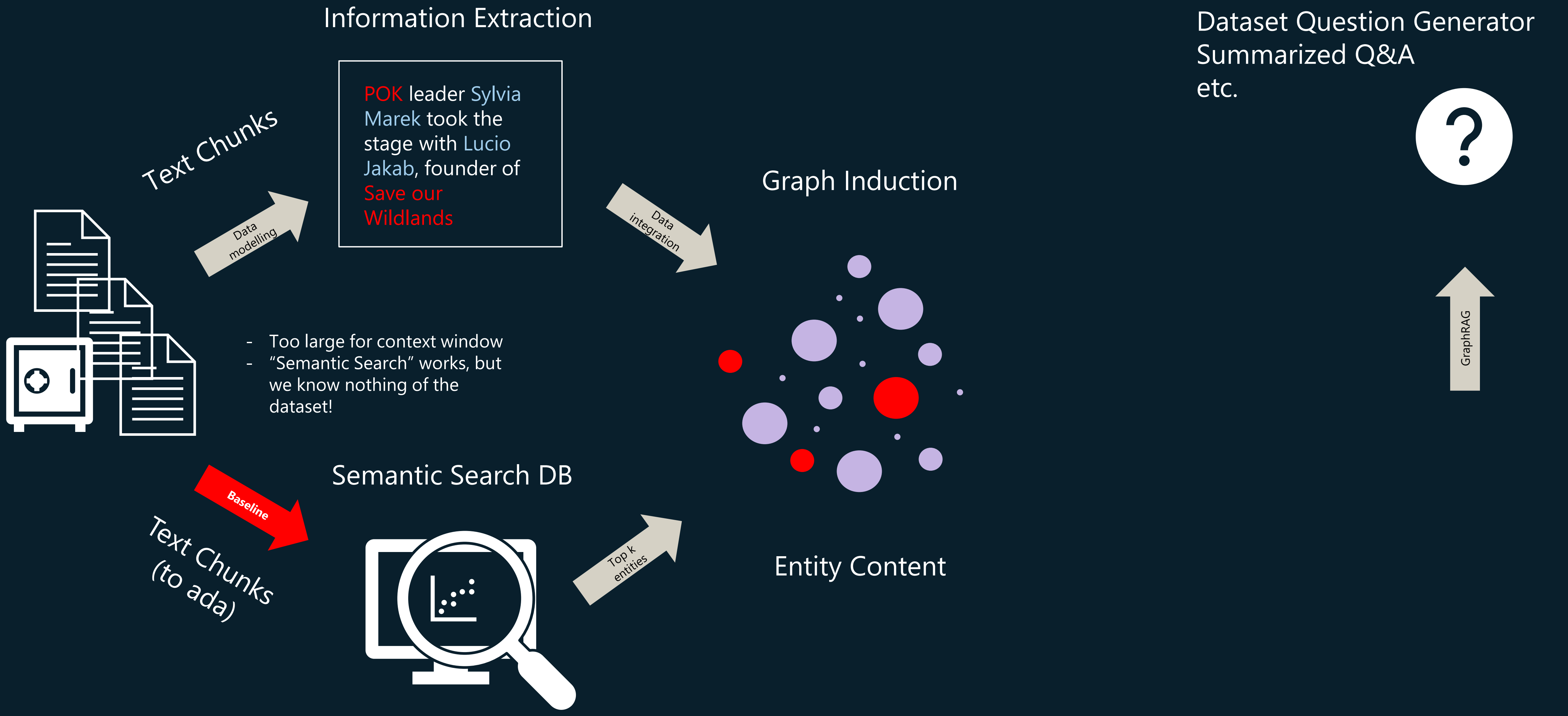
Hierarchy extraction
Graph embedding
Entity summarization
Community summarization

Semantic Search DB

Entity Content

Larsson, J. (2024, May 5). *GraphRag*. YouTube. https://youtu.be/r09tJfON6kE?si=r5V-w4VlxAzFQcu5

Comparing Baseline RAG to GraphRAG



POK leader Sylvia Marek took the stage with Lucio Jakab, founder of Save our Wildlands

- Too large for context window
- "Semantic Search" works, but we know nothing of the dataset!

Larsson, J. (2024, May 5). *GraphRag*. YouTube. https://youtu.be/r09tJfON6kE?si=r5V-w4VIxAzFQcu5

Microsoft

GraphRAG - Demo

Microsoft

Questions?

Resources

- https://github.com/ms-johnalex/intro-to-graphrag
- neo4j-graph-examples (github.com)
- Enhancing the Accuracy of RAG Applications With Knowledge Graphs | by Tomaz Bratanic
- https://github.com/tomasonjo/blogs/
- Project GraphRAG Web Site
- GraphRAG: Unlocking LLM discovery on narrative private data - Microsoft Research
- microsoft/graphrag: A modular graph-based Retrieval-Augmented Generation (RAG) system (github.com)
- Azure-Samples/graphrag-accelerator: One-click deploy of a Knowledge Graph powered RAG (GraphRAG) in Azure (github.com)
- GraphRAG python implementation

RAGHack: Next steps

Register for the hackathon @ aka.ms/raghack/register

Hack, hack, hack! Try our code samples or start from scratch! 🧑🏿‍💻 🧑🏻‍💻 👩🏼‍💻

For any questions, post in GitHub forum or attend Discord Office Hours. 🙋🏼 🙋🏾‍♀️

Join more live RAG sessions! aka.ms/raghack/streams

Submit your project before September 16th 11:59 PM PT to win prizes! 💵