Amazon Neptune

DISA TEM Presentation

June 2025

Nicole Moldovan
Principal NoSQL GTM Specialist
Amazon Web Services

Meet Your AWS Account Team

Account Manager

Enables and facilitates your business account

Account Manager

Solutions Architect

Provides technical guidance

Customer Solutions Manager

Helps you drive your cloud journey

Most complete set of relational & purpose-built databases

Amazon RDS

Amazon Aurora

KEY-VALUE

Amazon DynamoDB

DOCUMENT

Amazon DocumentDB

CACHING

Amazon ElastiCache

GRAPH

Amazon Neptune

TIME-SERIES

Amazon Timestream

MEMORY

Amazon MemoryDB

WIDE COLUMN

Amazon Keyspaces

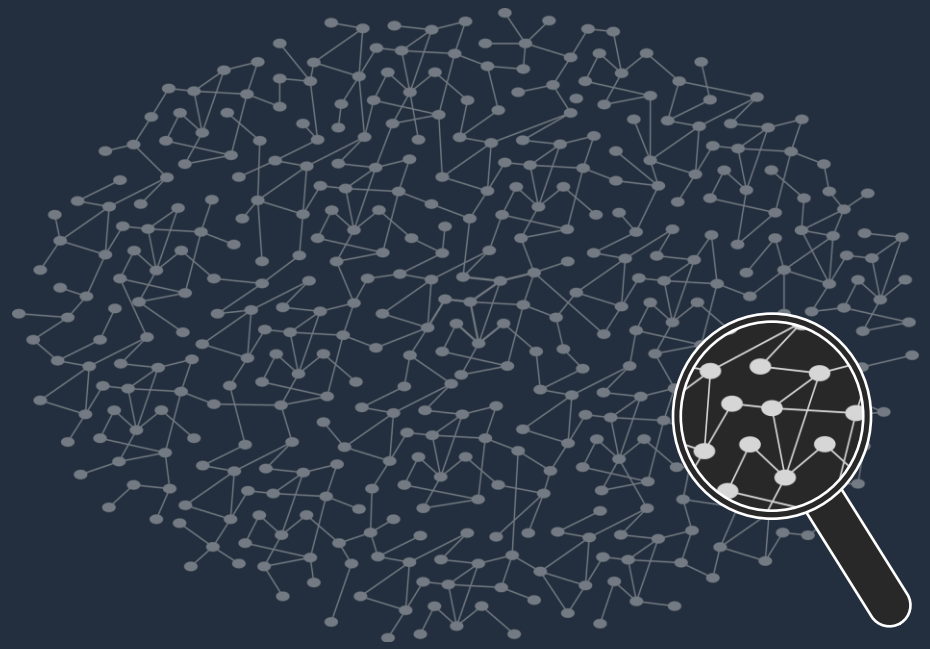
Reference customers

Amazon Neptune

Customers across different verticals and use cases use Amazon Neptune in production today

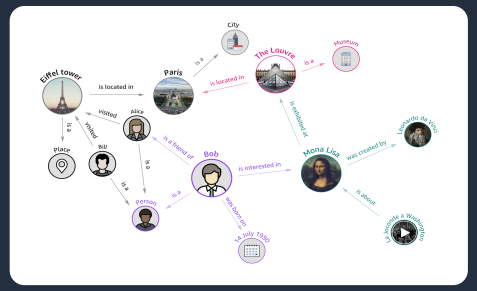
THOMSON REUTERS
finra
SIEMENS
AstraZeneca
LIFEOMIC
freshworks
amazon alexa
intuit
FACTSET
SAMSUNG
Blackfynn
Rappi
Uber ATG
HUUUGE
NBCUniversal
NETFLIX
asurion
COX AUTOMOTIVE
PaySense
zeta
Pearson
MARINUS ANALYTICS
noonum
audible FOR BUSINESS

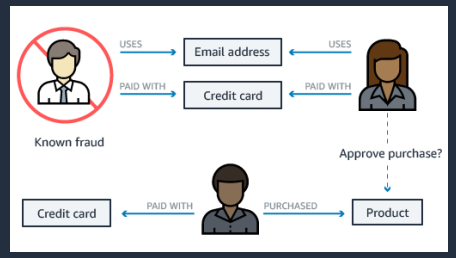
Graphs enable customers to innovate



1. Model data based on relationships
2. Applications explore connections and patterns in connected data
3. Processing graphs is hard due to random data access
4. Generalized graph operations require purpose-built processing

Amazon Neptune Use Cases





Knowledge Graphs

- Link disparate and heterogeneous data sources together to discover hidden connections
- Make data easily accessible through improved search results

Identity Graphs

- Persistent identifier to link all related devices and ids, enabling unified profile creation, targeting, and personalization
- Create audiences based on interests, preferences, and purchases

Fraud Graphs

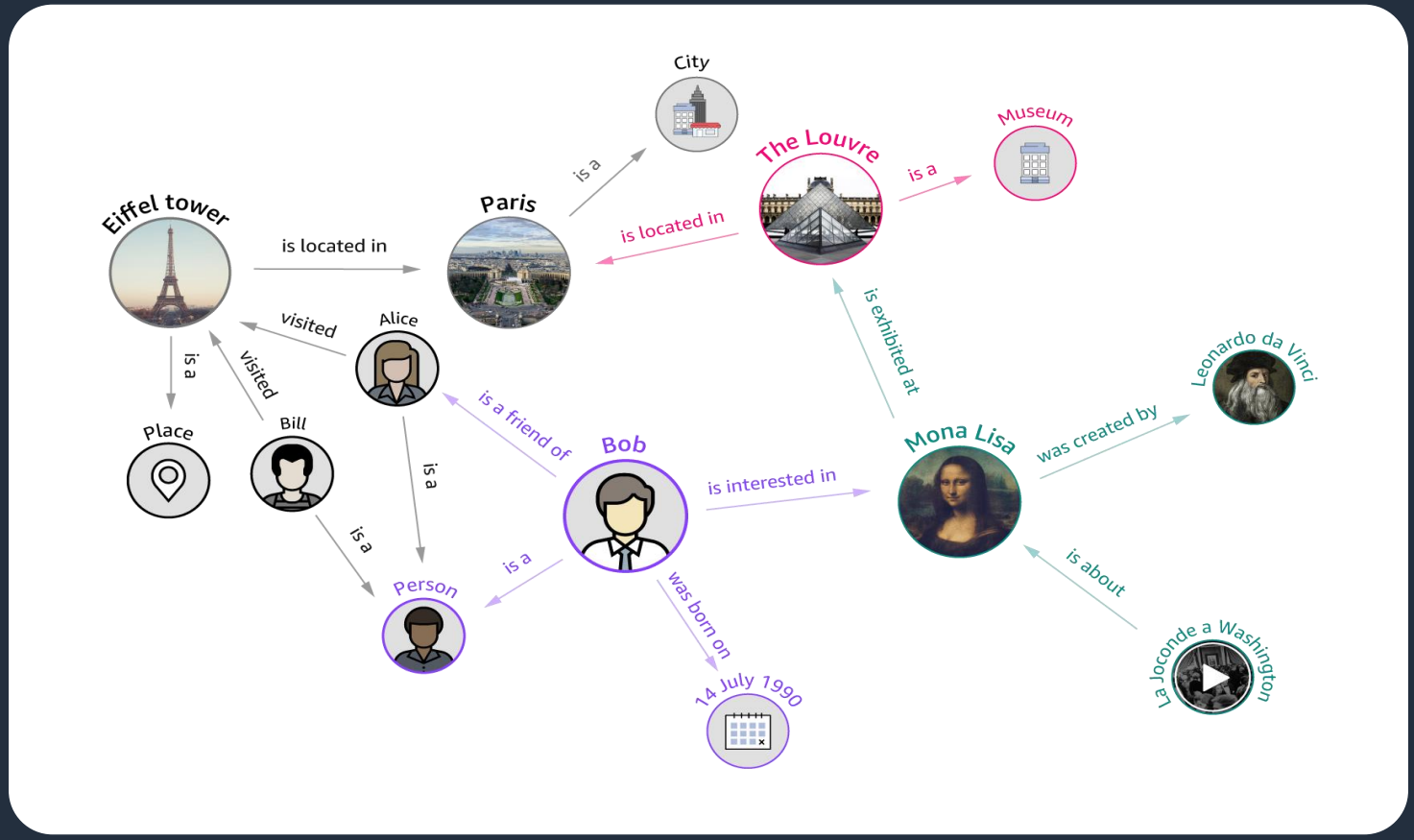
- Locate and prevent fraudulent patterns of transactions as they are occurring
- Derive unique insights commonly used to identify fraud

Security Graphs

- Use graphs to detect threats to your environment such as unwanted user access to applications, or exposed resources

Knowledge graphs

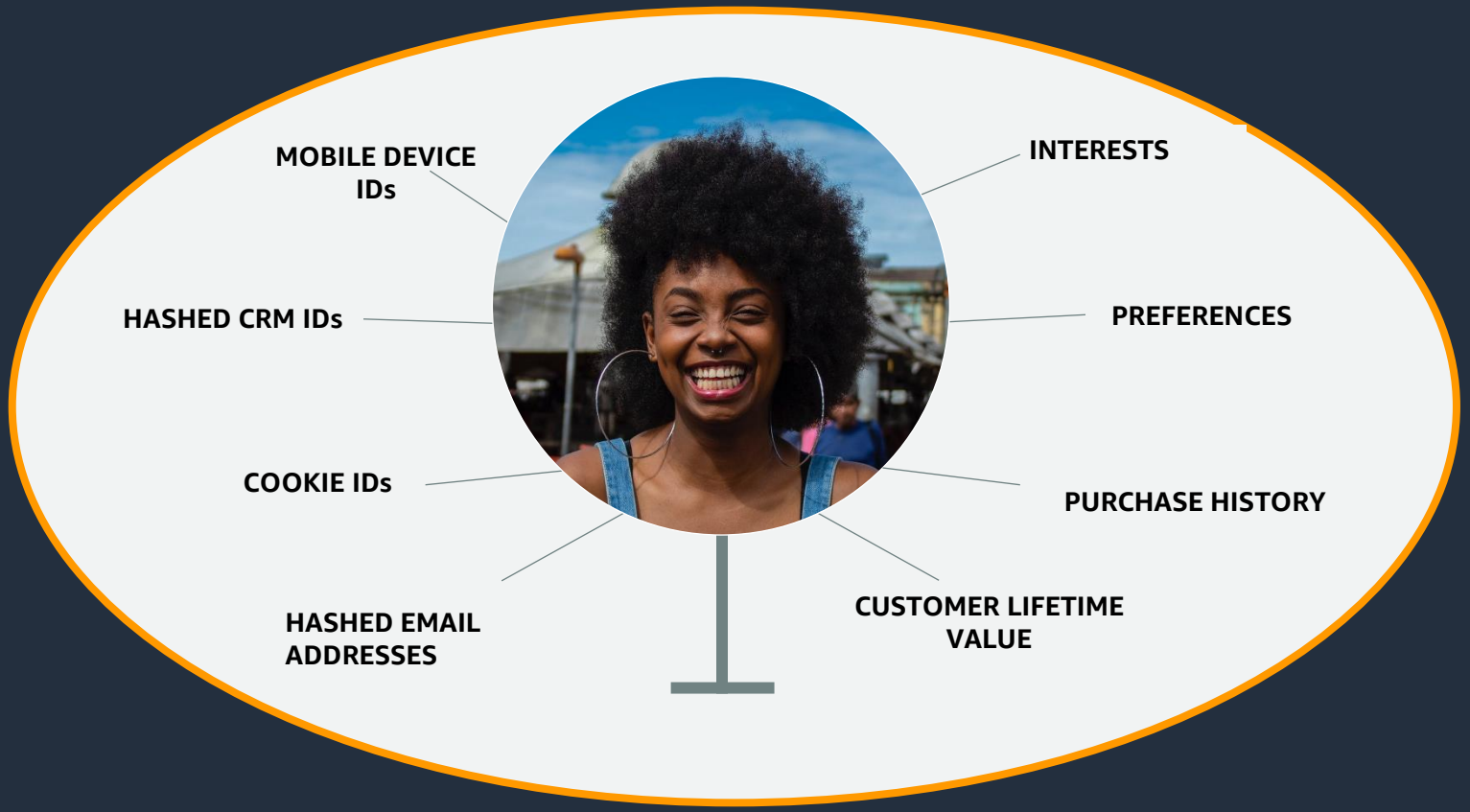
UNDERSTANDING THE WHO, WHAT, WHEN, AND WHERE



https://aws.amazon.com/neptune/knowledge-graphs-on-aws/

Identity graphs

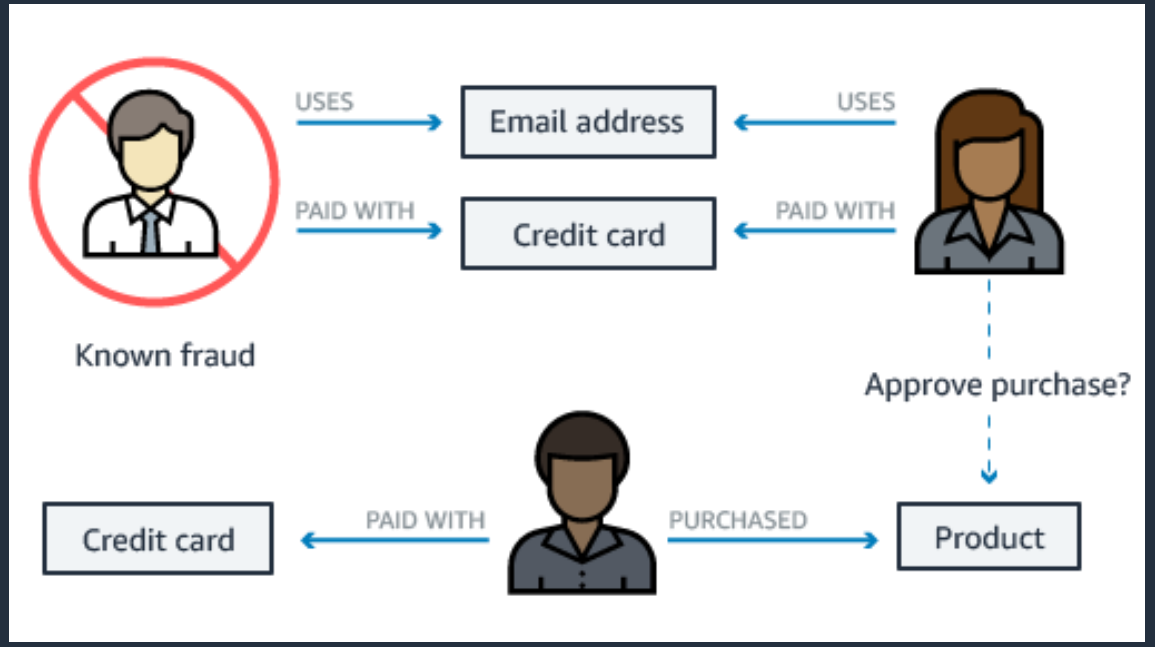
UNIFIED 360° VIEW OF THE CUSTOMER



https://aws.amazon.com/neptune/identity-graphs-on-aws/

Fraud graphs

DETECTING FRAUD AS IT HAPPENS USING RELATIONSHIPS



https://aws.amazon.com/neptune/fraud-graphs-on-aws/

As India’s leading gaming company, Games24x7 is known for its flagship products like RummyCircle, which offers online rummy, and My11Circle, which offers fantasy sports.

Challenge:

As the game of Rummy involves real money, Games24x7 has to stay vigilant to prevent fraud and collusion during tournaments.

Solution:

It uses the Amazon Neptune graph database to detect if two players in a game are colluding to beat the other four players. This is accomplished by assigning a table in the database to each player when they log in.

Security graphs

UNDERSTAND SECURITY VULNERABILITIES ACROSS LAYERS

1. Cloud Security Posture Management
2. Data Flow/Exfiltration
3. Identity and Access Management

https://aws.amazon.com/neptune/security-graphs-on-aws/

Wiz: Redefining cloud security

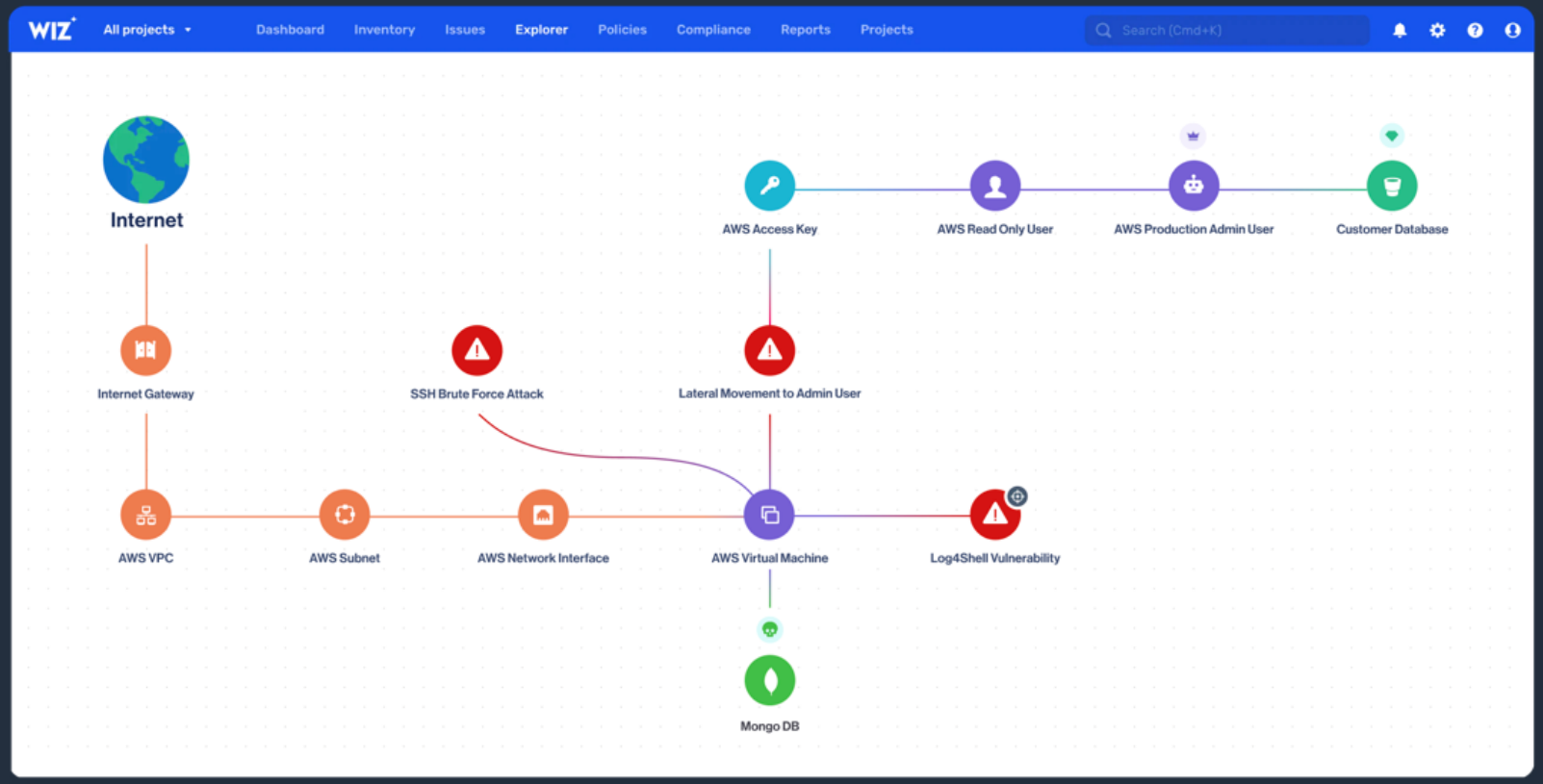
WIZ CONTINUES TO GROW AT AN ASTONISHING RATE*

5M workloads

Scanned daily

Over 30%

of the Fortune 100



*Forbes article by Peter Cohan: https://bit.ly/3EZGJtq

Neptune's Family

Neptune Databases

- **Store and manage graph data sets**
- A serverless graph database designed for superior scalability and availability.
- For graph database workloads that need to scale to 100,000 queries per second, multi-AZ high availability, and multi-region deployments.
- Example use cases: Social Networking Applications, Fraud Alerting, and Customer 360

Neptune Machine Learning

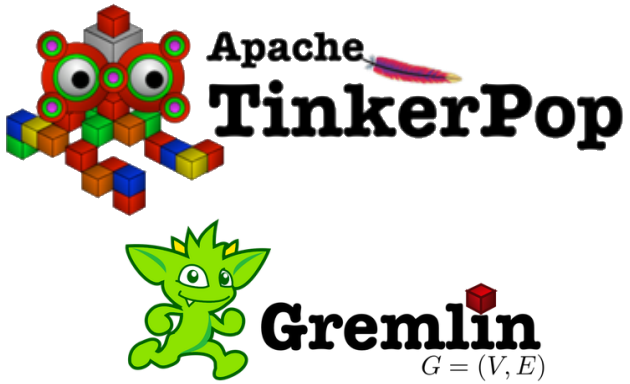
- **Creates, trains, and applies ML models on your graph data**
- An integration with SageMaker that uses the Deep Graph Library (DGL) and graph neural networks (GNNs) to make easy, fast, and more accurate predictions using graph data
- For automating the selection and training of ML models on graph data and for using Neptune Database APIs to make predictions on live data.
- Example use cases: Link prediction for recommendations, fraud detection, or identity resolution.

Neptune Analytics

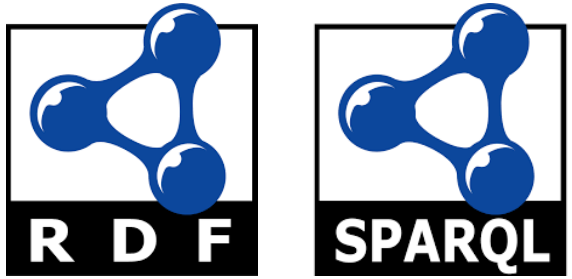
- **Analyze graph data sets in-memory**
- A graph analytics database engine to quickly analyze large amounts of graph data to get insights and find trends.
- For quickly analyzing existing graph databases or graph data sets stored in a data lake using popular graph analytic algorithms and low-latency analytic queries.
- Store vector embeddings as properties of nodes. Add explainability to semantic search results
- Example use cases: Targeted content recommendations, Fraud investigation, and Network threat detection

Leading graph models and frameworks

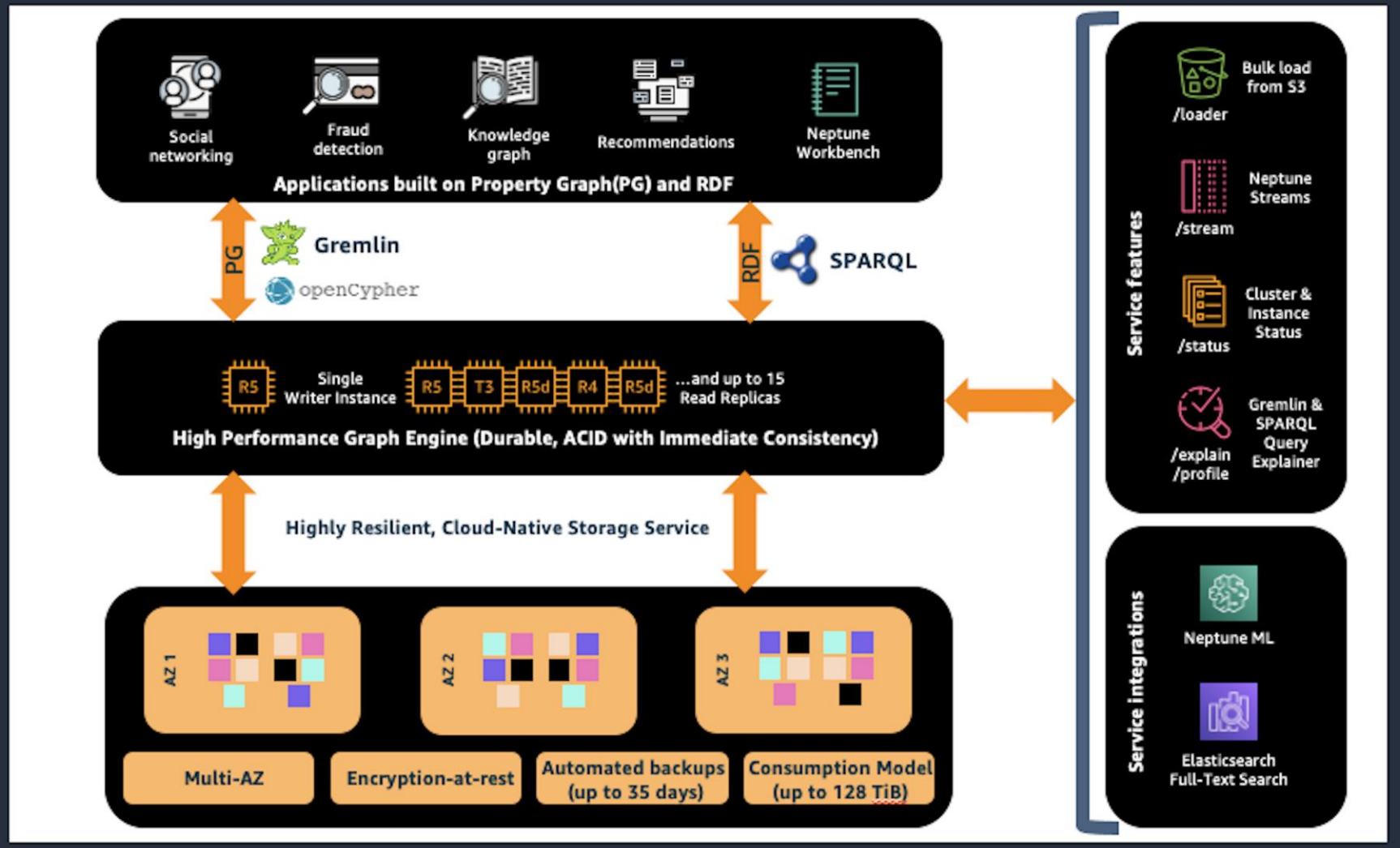
PROPERTY GRAPH



RESOURCE DESCRIPTION FRAMEWORK (RDF)



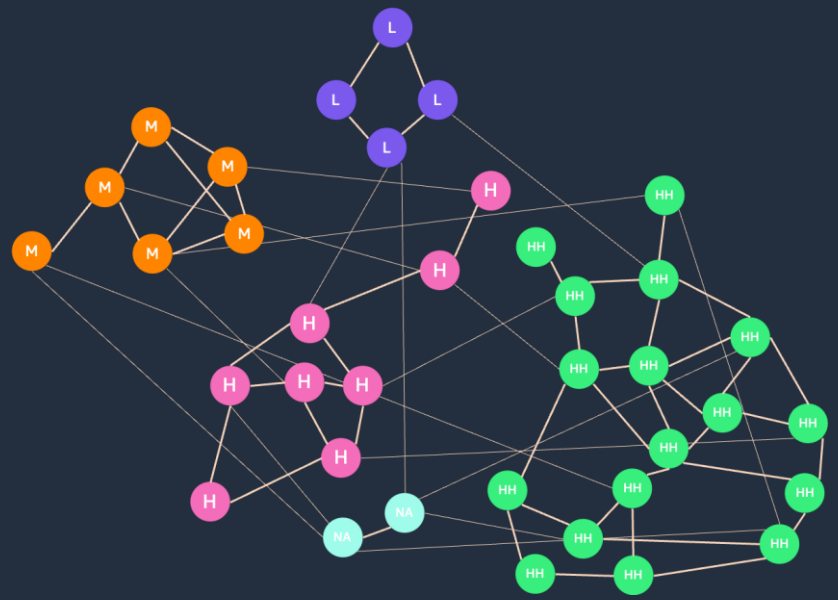
Amazon Neptune Database High Level Architecture



Amazon Neptune Analytics

Amazon Neptune Analytics

New analytics database engine for Amazon Neptune helps data scientists and application developers make data discoveries faster by analyzing graph data with tens of billions of connections in seconds



One Endpoint to orchestrate your graph workloads

- Use one API endpoint to create a graph, load data, invoke queries, and perform vector similarity search
- Popular OSS query language: openCypher

High-performance graph analytic queries and graph algorithms

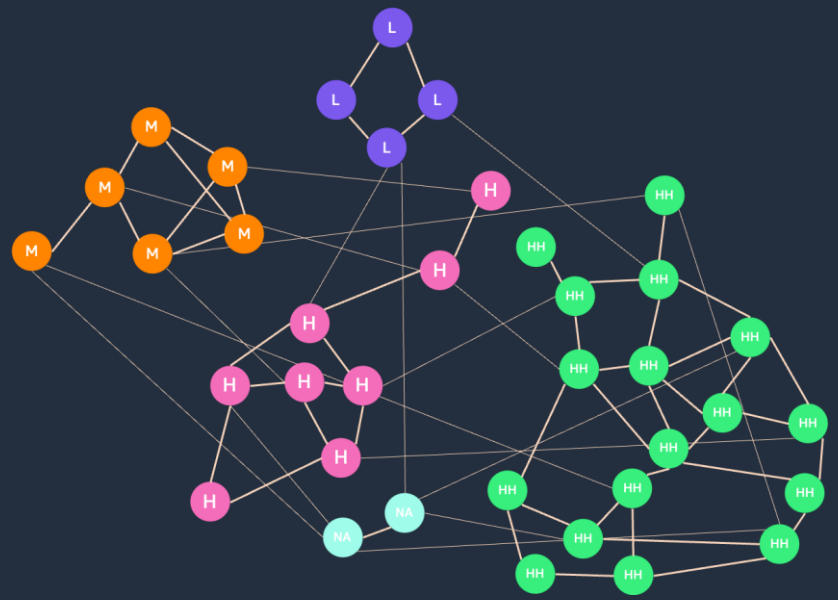
- 80X faster than previous AWS solutions (10M/edges/sec)
- 5 families of algorithms with 30 variants
- Via openCypher, integrated query syntax with mutations

Store and Search Vectors for Generative AI Applications

- Store and search embeddings trained in graph queries
- Combine graph algorithms, pattern matching traversals, and vector search into one query to reduce services required for generative AI workloads.

Amazon Neptune Analytics

Example Use Cases



Ephemeral Analytics

Customers load and analyze large datasets quickly for hourly/daily metric computations, near real-time analytics, risk modeling, and situations requiring quick insights.

Low Latency Analytical Queries

With low-latency graph queries, you can extract graph features from connected data to fuel predictive models in real-time to fuel personalized AI models.

Vector Search with Graph Data

Best of both worlds: build AI applications using techniques like retrieval augmented generation (RAG) that combine graph traversals and vector search for context augmentation.

GenAI & GraphRAG with Amazon Neptune

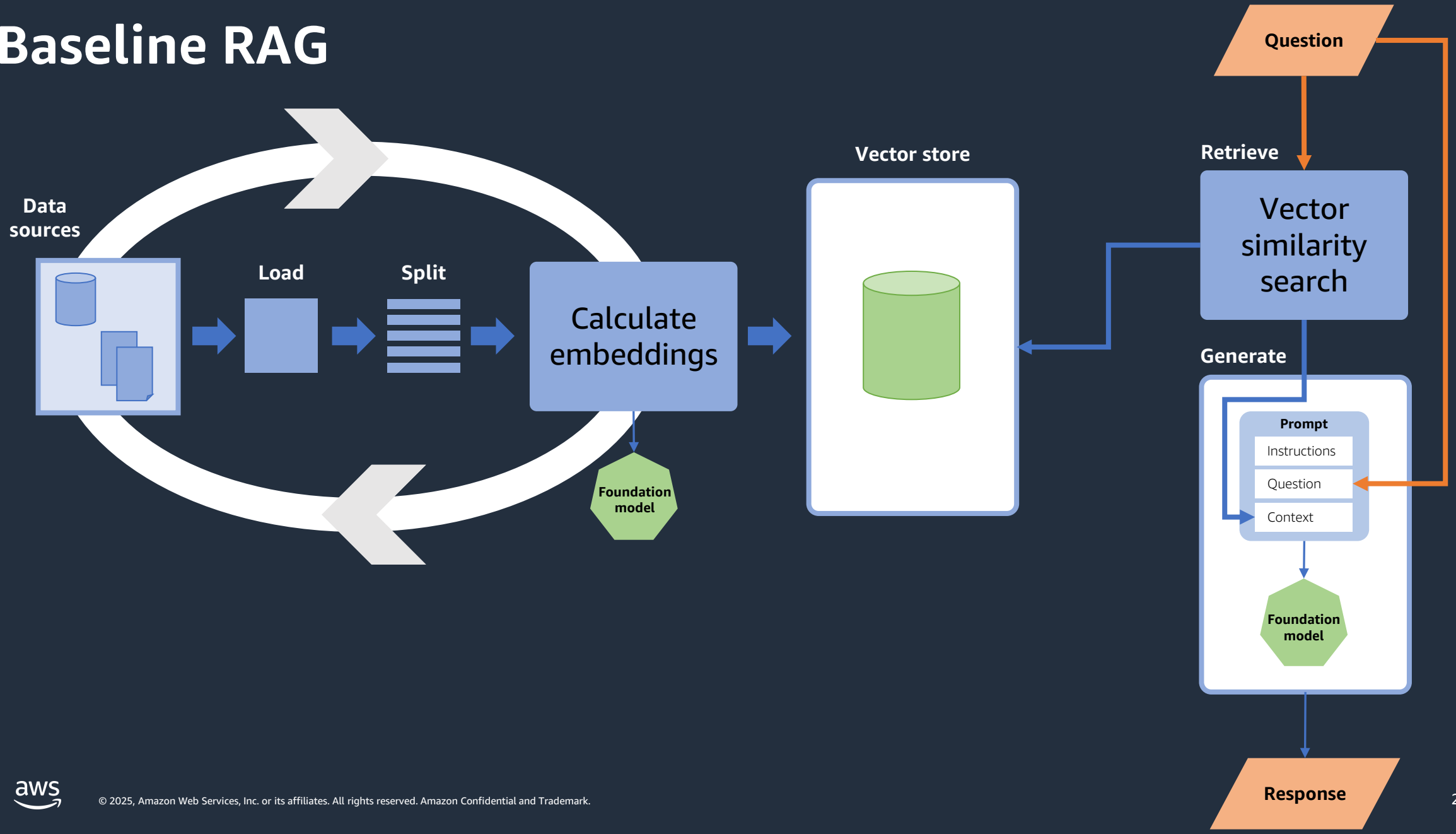
RAG is a powerful architecture pattern but has complex data challenges

CONNECTEDNESS
data spread across multiple disparate documents is hard to retrieve

SPECIFICITY
embeddings are sparse representations of data which may lack crucial details

EXPLAINABILITY
explaining the relevance of data retrieve is demanding

Baseline RAG



Sometimes the most relevant information to answering a question is not the closest in meaning.

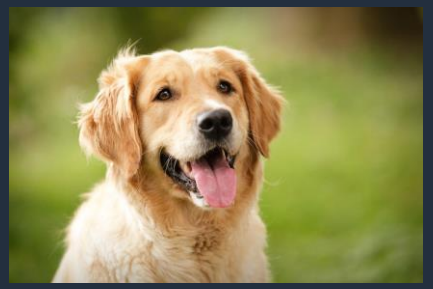
Searching in Vector space vs. Graph space?

Similarity in vector space compares mathematical closeness

Zucchini is similar to *summer squash* and *courgette*

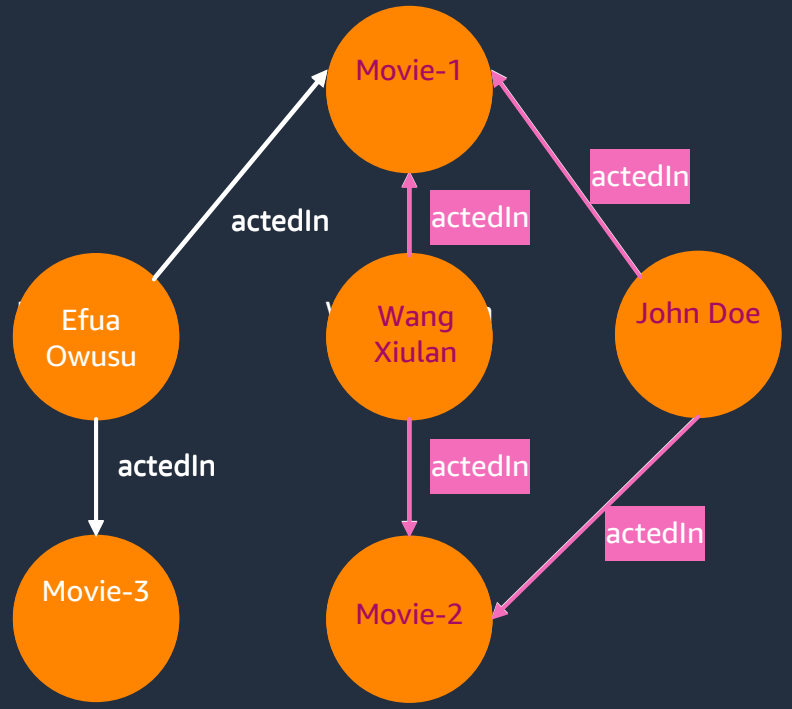
Vectors can represent ...

... A text embedding model
... An image embedding model

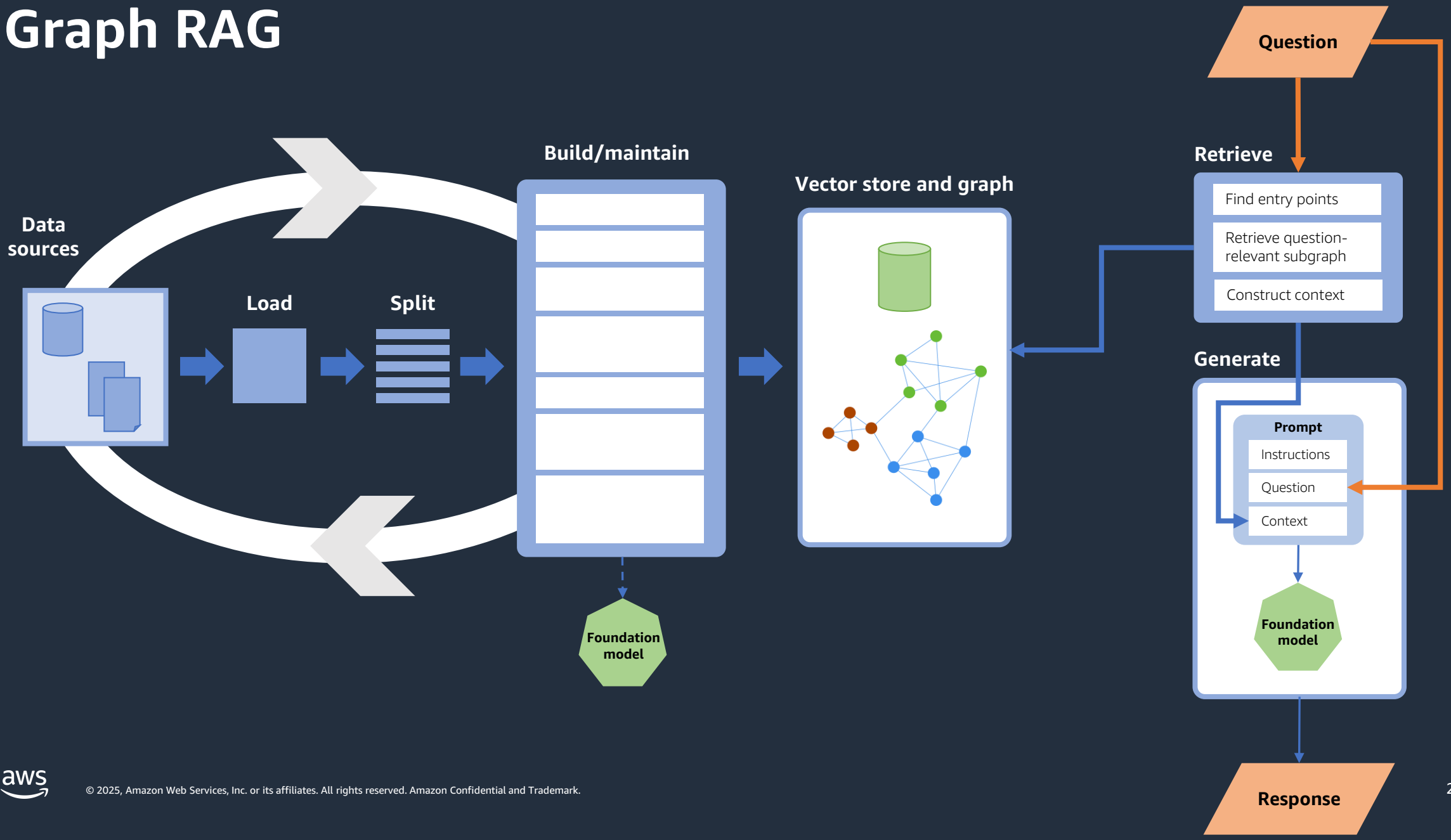


Relatedness in graph space compares shared connections

Wang Xiulan and *John Doe* are related because they've starred in multiple movies together



Graph RAG



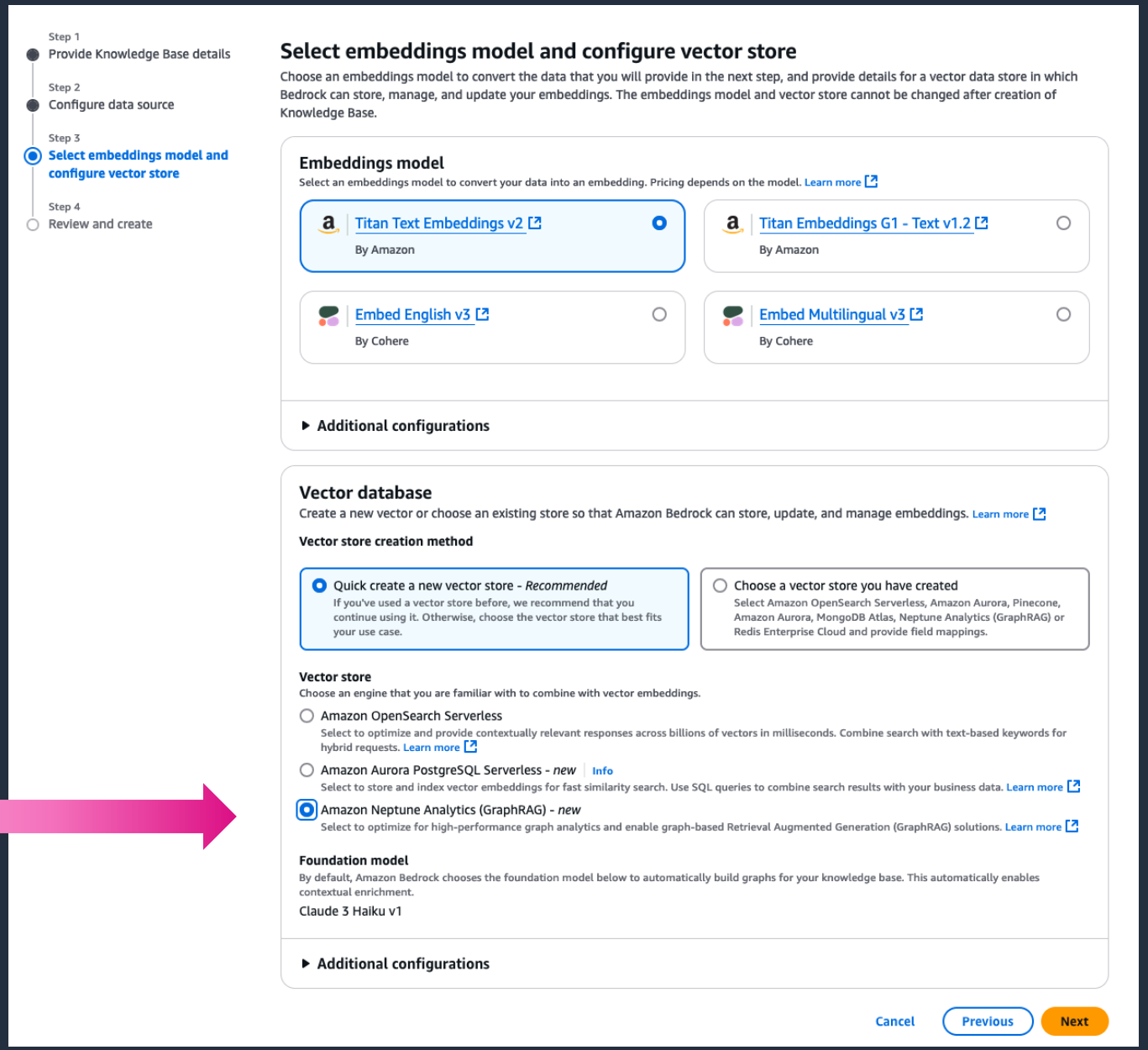
When to use GraphRAG

Multiple documents and data sources: When relevant information is dispersed across multiple sources or documents and the relationships between entities are not clear or hidden

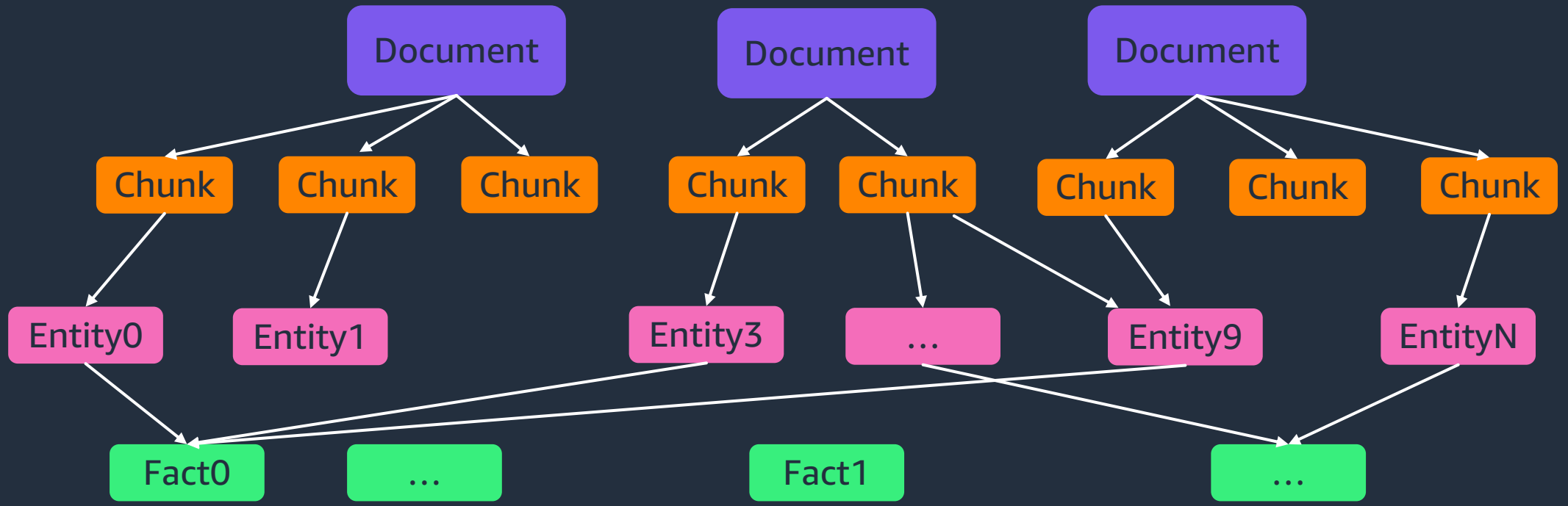
Summarization: When a summarization question targeting a specific domain across multiple documents

GraphRAG with Amazon Neptune in Amazon Bedrock Knowledge Bases

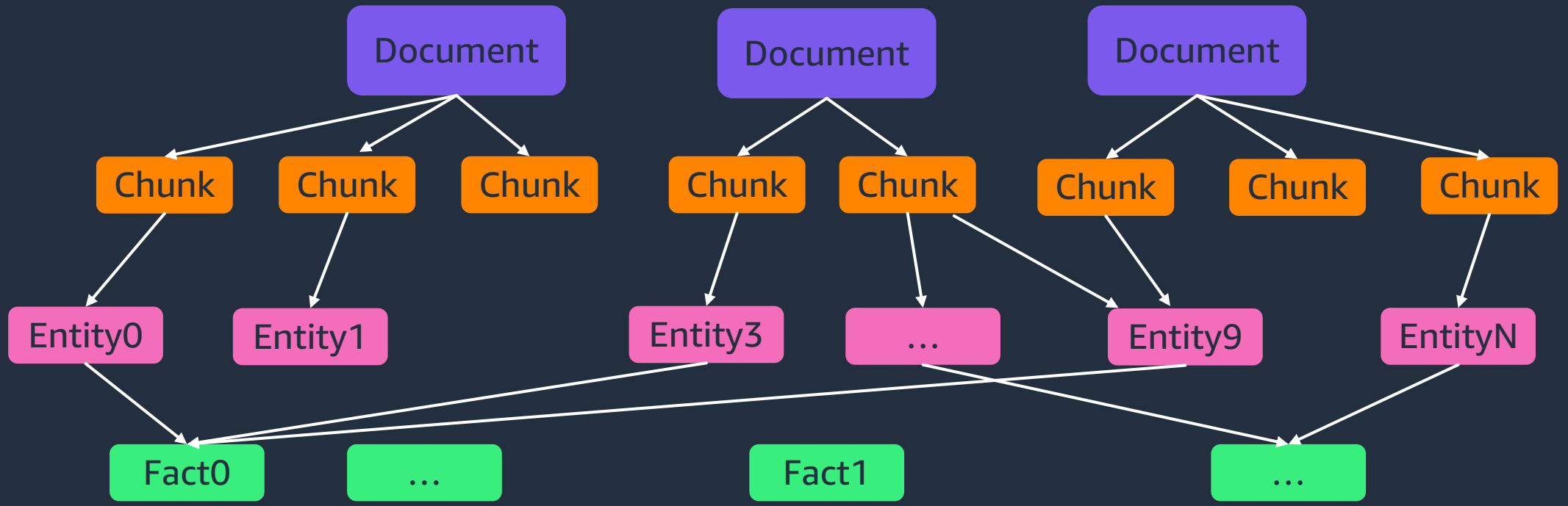
This will create a graph and embeddings and store them in provided Neptune Analytics



Automatically generates a graph from your data that is updated as you add new content

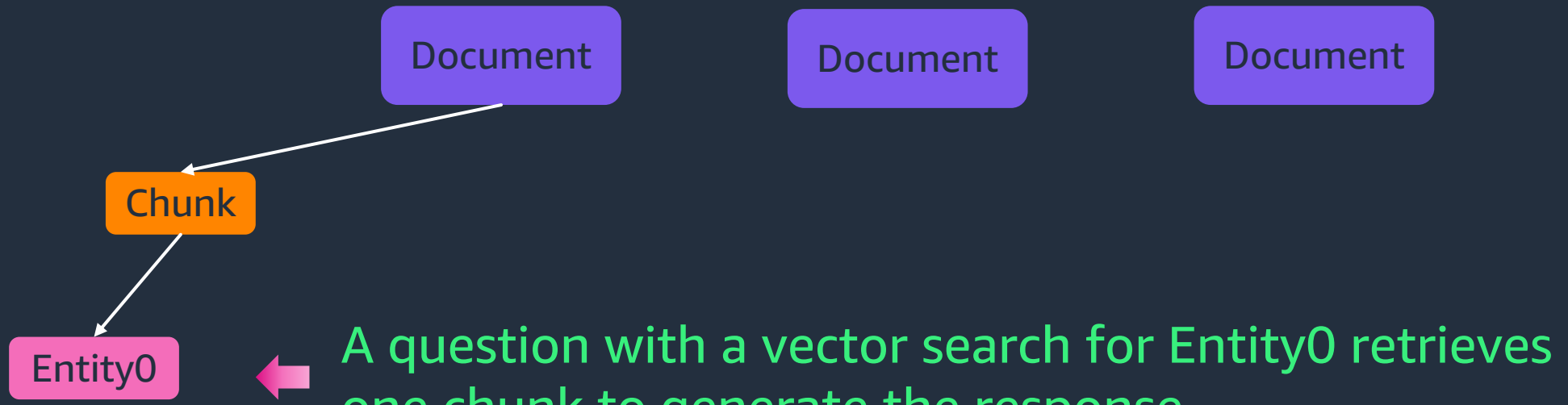


Uses the graph within the retrieval process to improve the responses



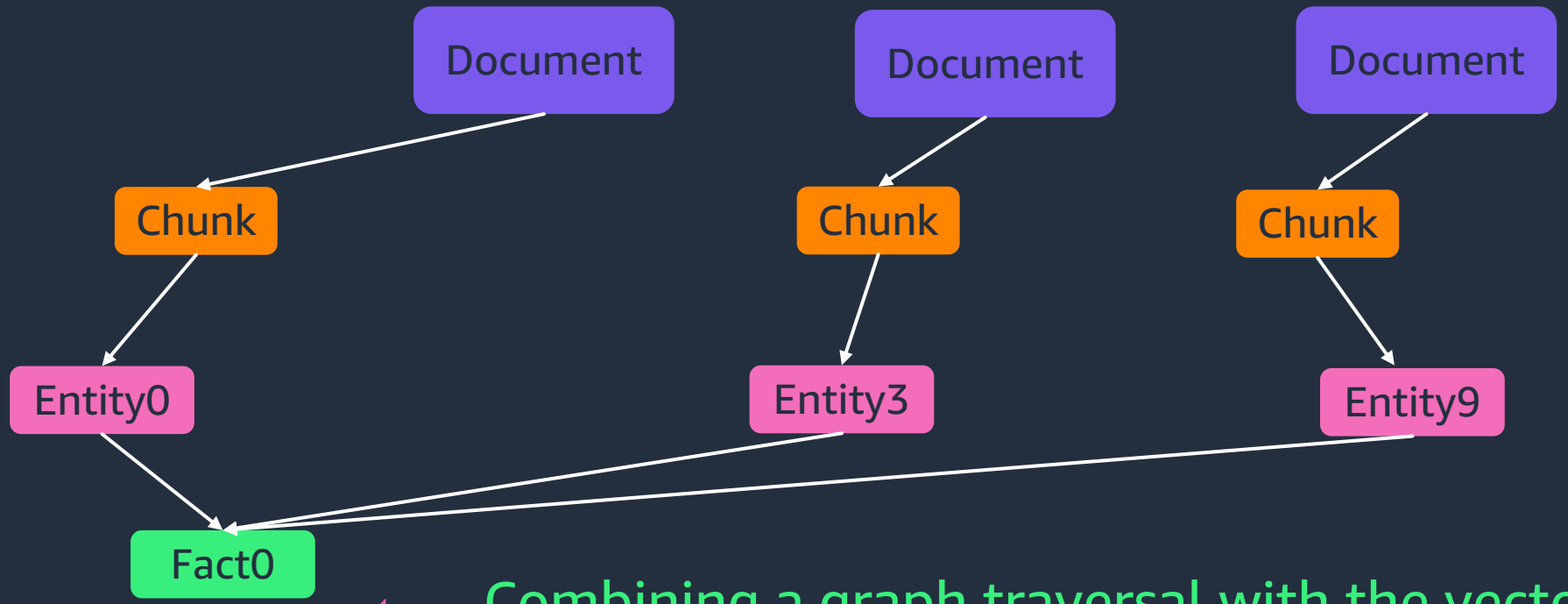
Responses can be more accurate and comprehensive, especially for multi-document retrieval, because the graph can be used to identify relevant chunks that are not top-K vector search results.

Uses the graph within the retrieval process to improve the responses



A question with a vector search for Entity0 retrieves one chunk to generate the response.

Uses the graph within the retrieval process to improve the responses



Combining a graph traversal with the vector search for Entity0, retrieves three chunks to generate the response

Scenario: Example Corp's Quarterly Financial Statement

Example Corp's quarterly report had four paragraphs:

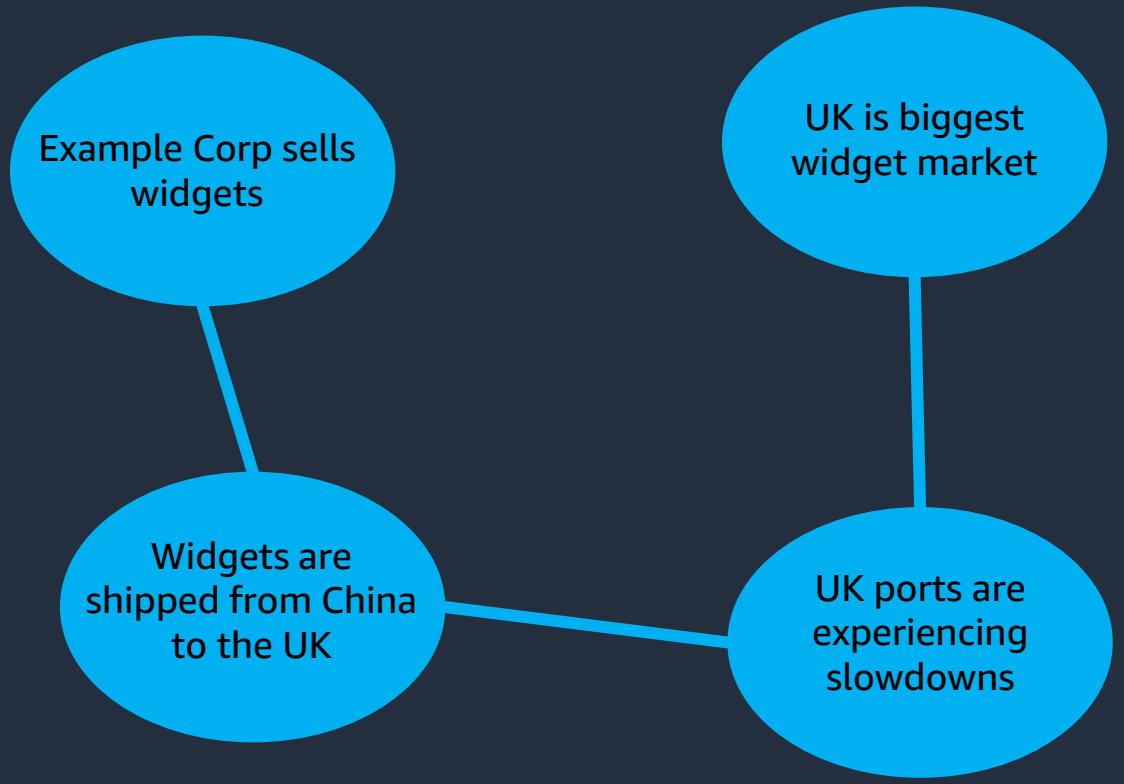
- Paragraph 1 discussed how "Example Corp sells widget"
- Paragraph 2 discussed that the "UK is the biggest widget market"
- Paragraph 3 discussed that the "Widgets are shipped from China to the UK"
- Paragraph 4 discussed that "UK ports are experiencing slowdowns"

Each of paragraph has the text and associated embedding stored

Question: What is the outlook for widget sales in the UK?

STEP 1: AN EMBEDDING IS CREATED OF THE QUESTION BEING ASKED

question embedding

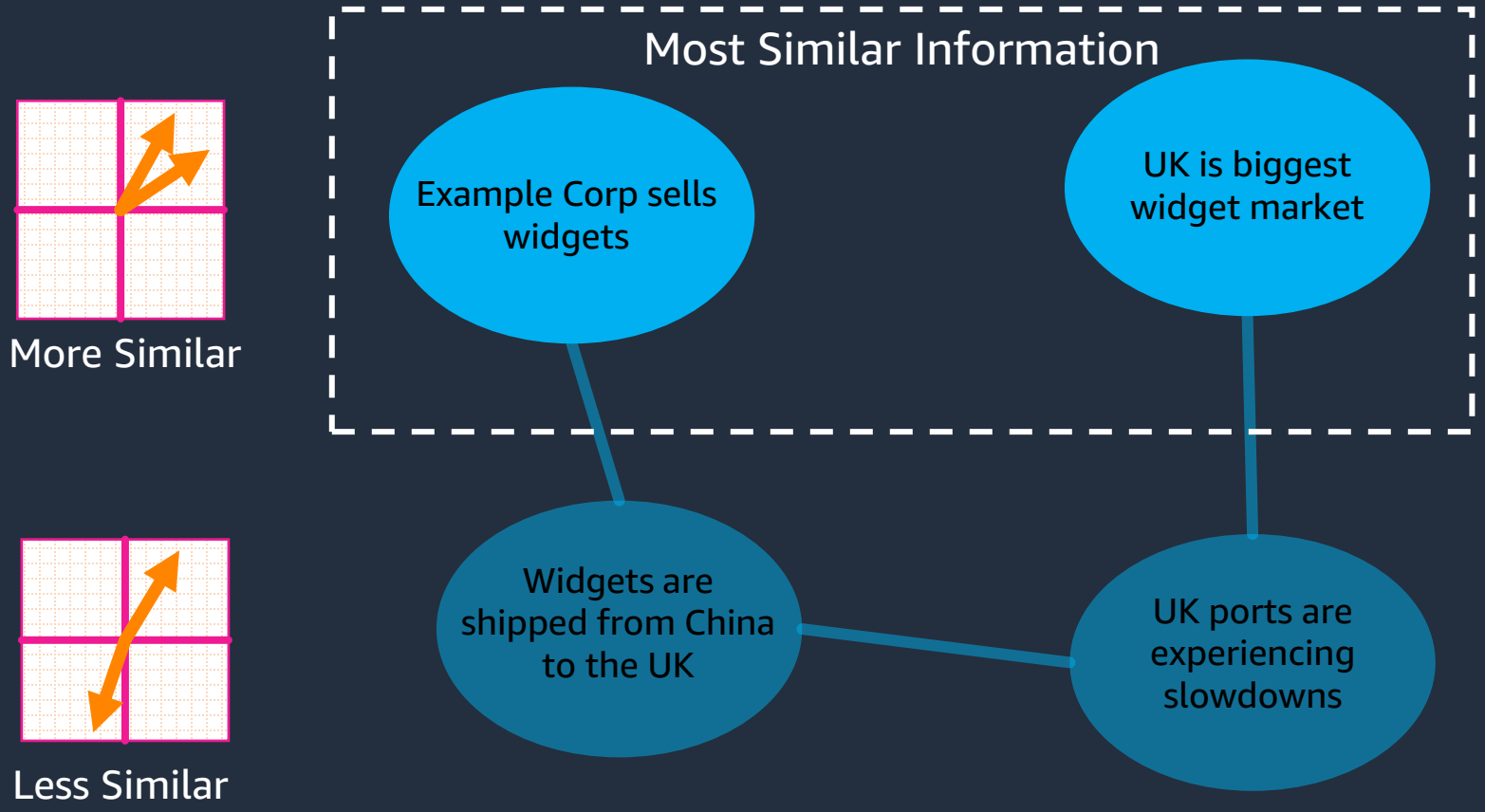


LLM Response

Question: What is the outlook for widget sales in the UK?

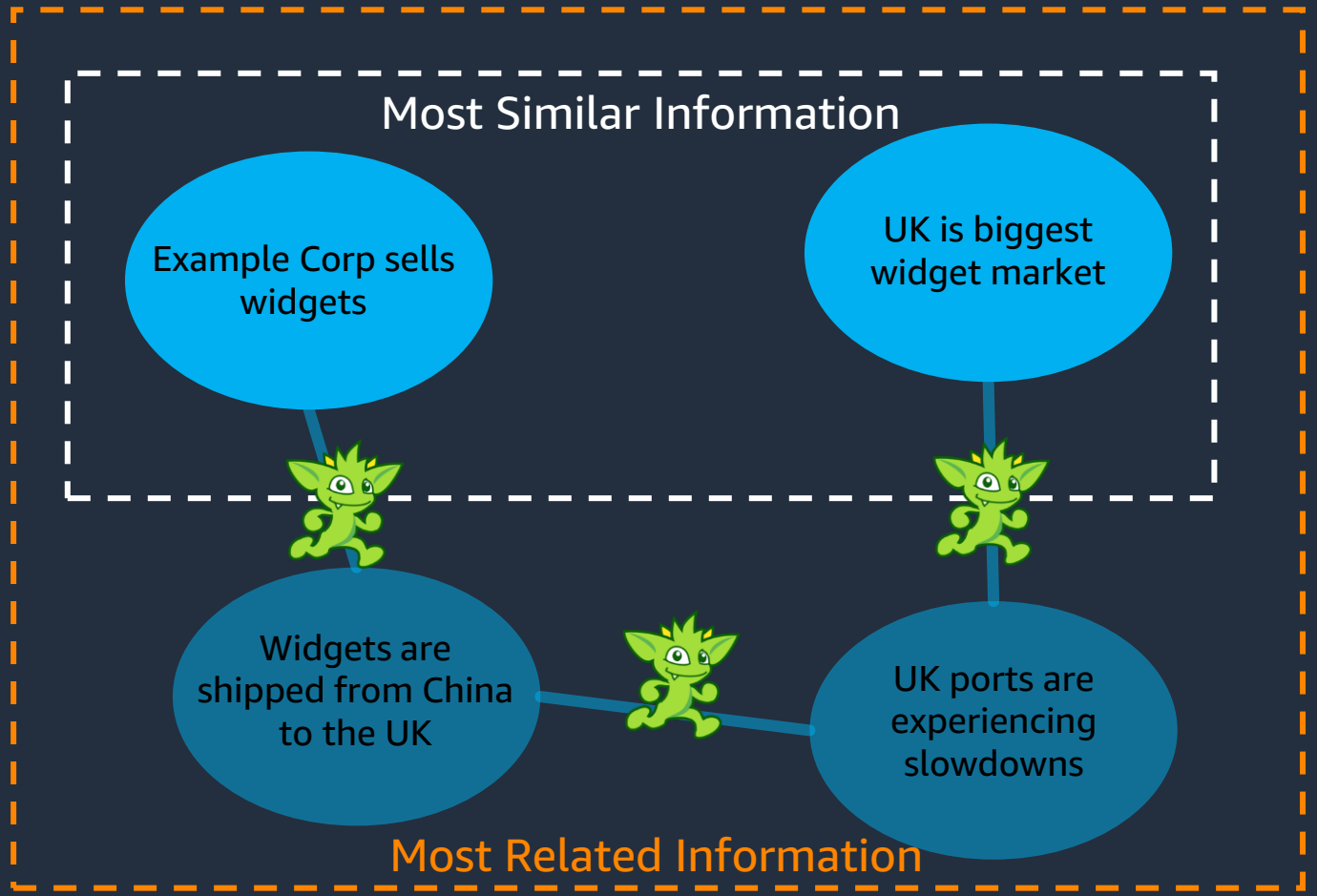
STEP 2: SIMILARITY SEARCH IS RUN TO FIND THE STARTING NODES

question embedding



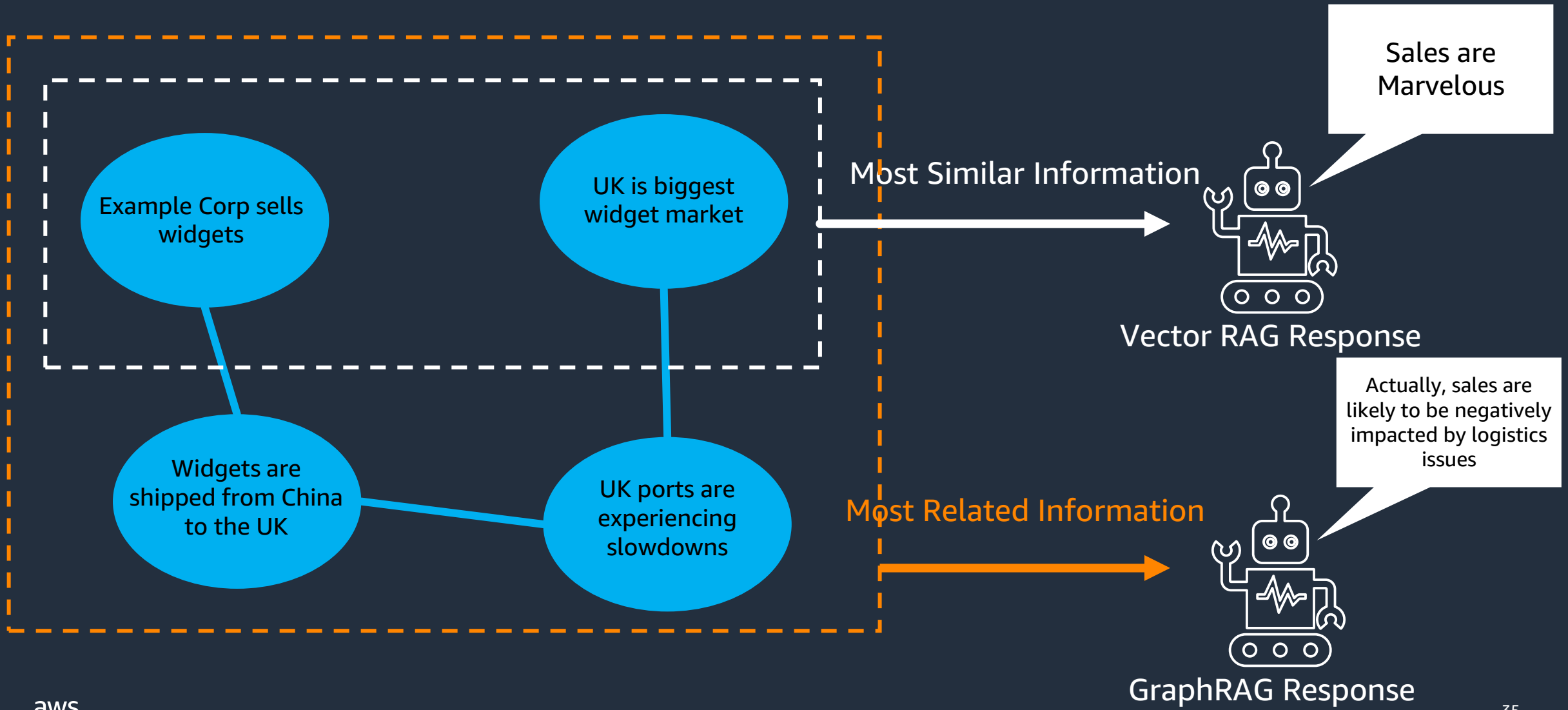
Question: What is the outlook for widget sales in the UK?

STEP 3: IN GRAPHRAG THE GRAPH IS ALSO USED TO LOCATE RELATED IDEAS



Question: What is the outlook for widget sales in the UK?

STEP 4: MOST RELEVANT INFORMATION PASSED TO GENERATE A RESPONSE



The combination of graph and vector search produces responses that can be more accurate and comprehensive, especially for multi-document, multi-hop questions

Leading players in many industries are already building GraphRAG applications

Finance	Cyber Security	Automotive	Pharmaceuticals
Various documents such as research reports, projections, analyst projections are leveraged to provide recommendation for better financial investments	Analyzing logs to interpret alerts, summarize security reports, do threat intelligence, and make recommendations	Increasing efficiency for internal processes – quickly pull the right datasets/databases for further analysis such as spare parts analysis	Improving Drug trials processes with better analysis, and performing publication analysis

Amazon Web Services announces support for GraphRAG with Amazon Neptune in Amazon Bedrock Knowledge Bases

Amazon Neptune Analytics

Amazon Bedrock Knowledge Bases

Fully managed AWS solution
Seamless integration within the AWS ecosystem for high security and performance

Now Generally Available

https://aws.amazon.com/neptune/graphrag/

AWS Labs Open Source GraphRAG toolkit

Amazon Neptune

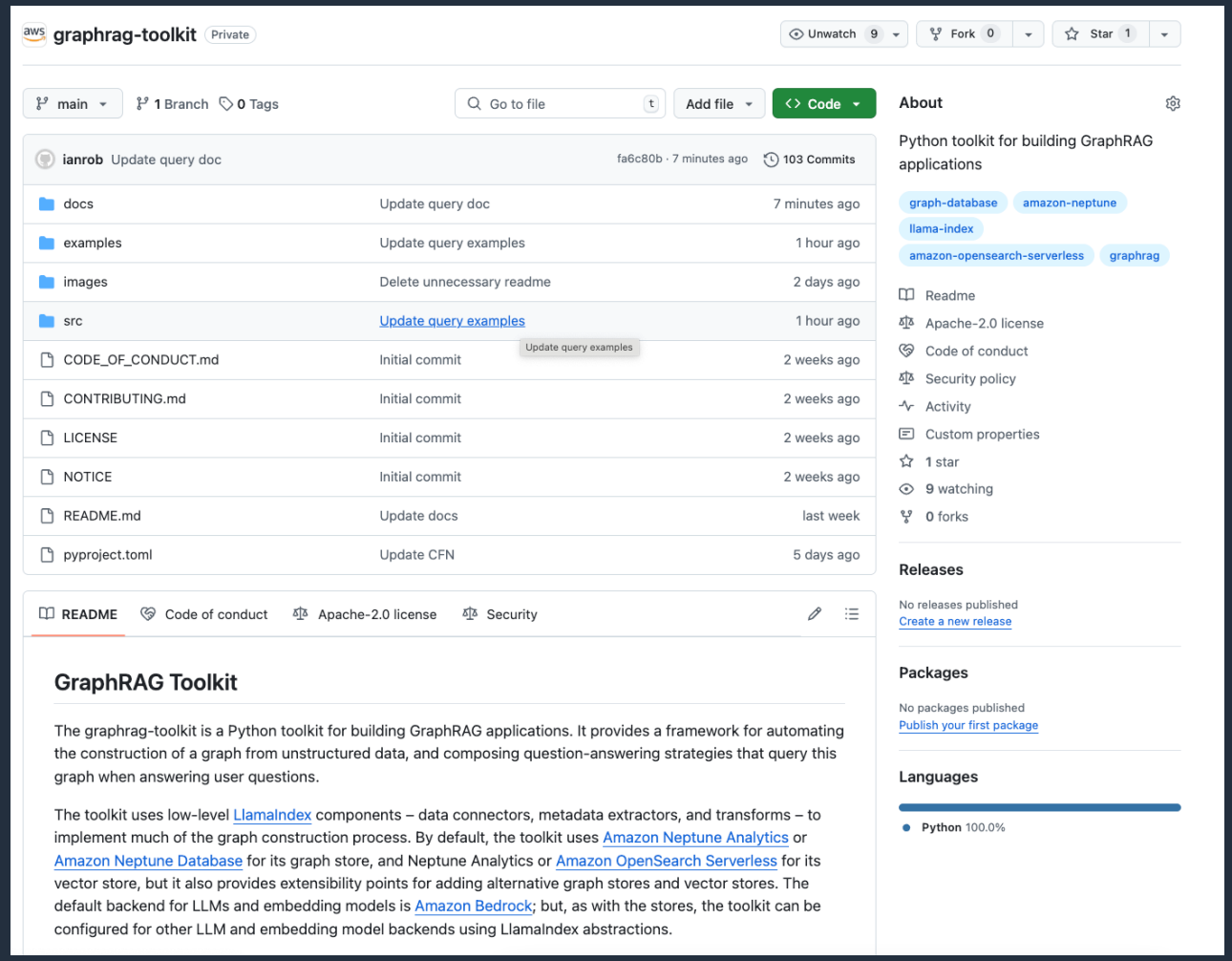
Amazon Bedrock

Flexible, adaptable solution for organizations that prefer open-source options and customizability

Now on GitHub!

https://github.com/awslabs/graphrag-toolkit

Open-Source GraphRAG Python Toolkit




```python
from graphrag_toolkit import LexicalGraphQueryEngine
from graphrag_toolkit.storage import GraphStoreFactory
from graphrag_toolkit.storage import VectorStoreFactory

import nest_asyncio
nest_asyncio.apply()

graph_store = GraphStoreFactory.for_graph_store(
    'neptune-db://my-graph.cluster-abcdefghijkl.us-east-1.neptune.amazonaws.com'
)

vector_store = VectorStoreFactory.for_vector_store(
    'aoss://https://abcdefghijkl.us-east-1.aoss.amazonaws.com'
)

query_engine = LexicalGraphQueryEngine.for_traversal_based_search(
    graph_store,
    vector_store
)

response = query_engine.query('''What are the differences between Neptune Database
                                 and Neptune Analytics?''')

print(response.response)
```

$ pip install https://github.com/awslabs/graphrag-toolkit/releases/latest/download/graphrag-toolkit.zip

Thank you!

Nicole Moldovan

nictune@amazon.com
linkedin.com/in/nicolemoldovan

Chris Walker

cwalkcw@amazon.com
linkedin.com/in/christopherawalker/

Neptune Team

Neptune-gtm@amazon.com

Product Page

https://aws.amazon.com/neptune/