

**Transforming Pharma Master Data
Management: Integrating Neo4j
Knowledge Graphs with GraphRAG
and Generative AI for Enhanced
Data Querying and Analytics**



Data Driven Decisions

[Solutions](#)

[Platforms](#)

[Services](#)

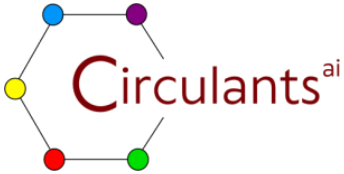
ISO-27001 BSI Certified

Agenda



- 1 About Circulants, **New Age Specialist Technology Company**
- 2 Introduction to Graph Database
- 3 MDM Solution with Neo4J
- 4 Graph RAG and GenAI
- 5 Questions

New Age Specialist Technology Company



**Domain
Expertise**



**Global
Presence**



**140+
Associates**



**155
Projects**



**35+
Clients**

Head-Quarters
Founded in 2018, Bay
Area-based company

IP-Based,
Domain-Specific
Solutions

ISO-27001
Certified



Key Clinical Focus Areas



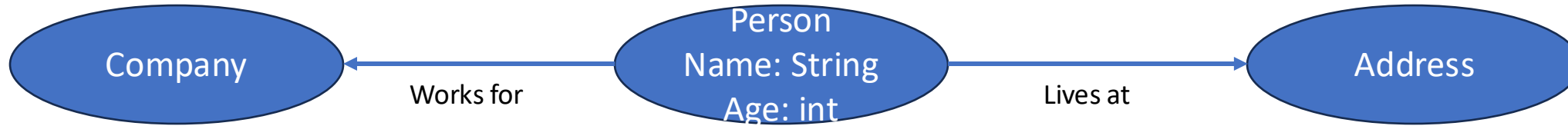
Our Pharma resources stand out from the crowd



Graph Database



A graph database stores data points as 'Nodes' and connections between them as 'Relationships'.



Nodes: Entities or objects in a graph, representing data points (e.g., Person, Company, Address).

Edges: Relationships between nodes, defining how they are related (e.g., Works for, Lives at).

Properties: Key-value pairs attached to nodes or edges, storing additional information (e.g., Name, Age)

Graph Database

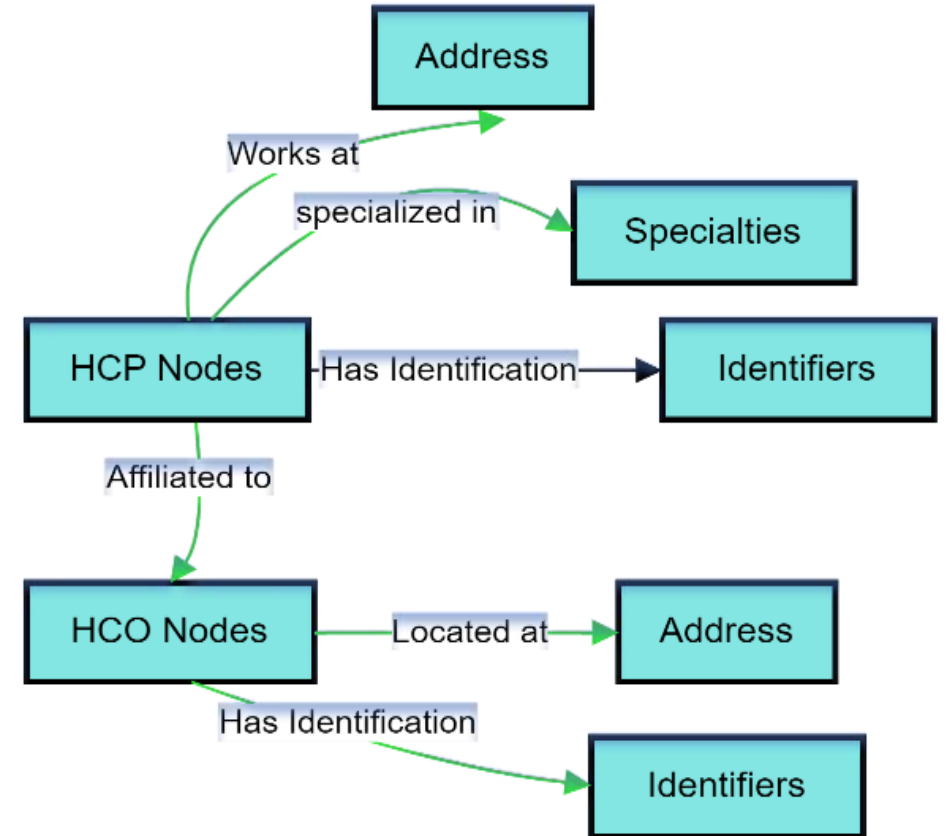


Nodes:

- Nodes are the main data element.
- Nodes are connected to other nodes via relationships/edges.
- Nodes can have one or more properties.

Relationships/Edges:

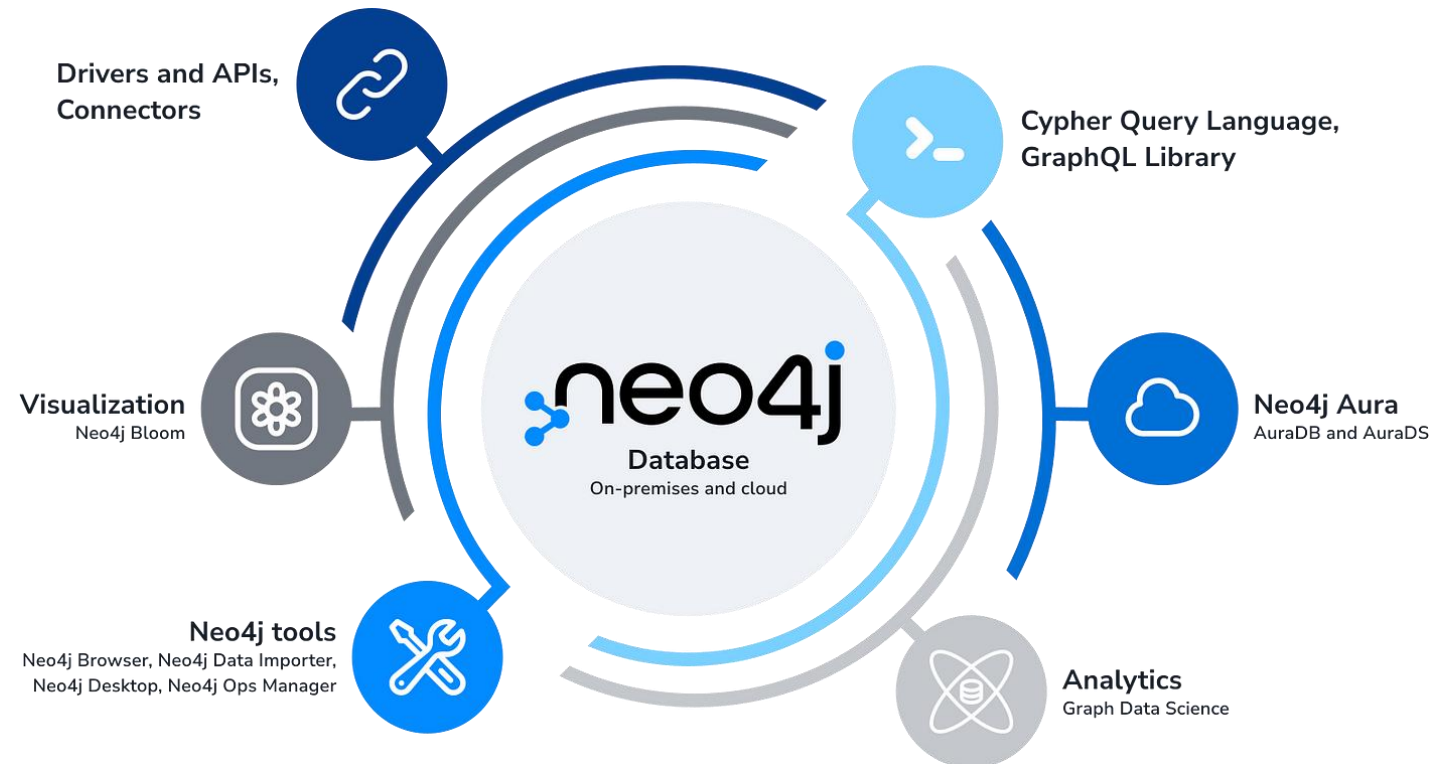
- Relationships connect two nodes.
- Relationships are directional
- Nodes can have multiple relationships
- Relationships can have multiple properties.



Some Leading Graph Databases:- Neo4J, Cosmos DB, Virtuoso, ArangoDB

Introduction to Neo4j:

- Neo4j is a NoSQL (schema optional) graph database management system.
- The data elements Neo4j stores are nodes, edges connecting them, and attributes of nodes and edges.
- It uses the Cypher query language for querying graph data.
- It is ACID compliant and highly scalable.



MDM Solution with Neo4J:

Overview of MDM:

- MDM is about creating a single, accurate view of key business entities (e.g., customer, product, supplier).
- Ensures consistency across various data sources and systems.

Neo4j in MDM:

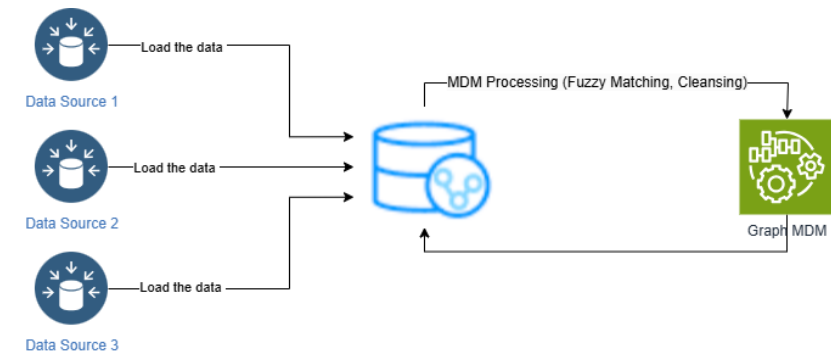
- Graph representation of entities and relationships, allowing us to manage and link diverse data sources efficiently.
- Neo4j helps in data cleaning by using graph relationships to identify inconsistencies, standardizing data through Cypher queries.
- Helps in resolving duplicate records and identifying entities that are the same across systems.

Using Sorensen-Dice Similarity for Matching:

- Sorensen-Dice Similarity is used for comparing the similarity between strings.
- In MDM, it's helpful in matching similar records.
- Neo4j uses the similarity score to group records into the same entity.

Benefits of Using Neo4j for MDM:

- Provides a single view of data through relationships and entity resolution.
- Ensures accurate and consistent master data.
- Efficiently links related entities and handles large datasets using graph traversal and similarity metrics.



Graph RAG & Gen AI:-

Graph RAG:

- RAG (Retrieval-Augmented Generation) combines Knowledge Graphs and Vector Databases to improve information retrieval and content generation.
- Enhances the ability to retrieve context-aware data and generate relevant outputs.

How it Works:

- **Knowledge Graph:** Represents structured data with entities and relationships for deeper insights and connections.
- **Vector Database:** Stores high-dimensional vectors (numeric representations of data) to enable semantic search based on meaning, not just keywords.
- **Integration:** The system retrieves relevant information (via graph and vectors) and uses it to generate more accurate and contextually relevant responses or content.

Benefits:

- **Enhanced retrieval:** Combines semantic search with structured data for better context and relevance.
- **Improved accuracy:** Enables context-aware content generation by using linked data and vector representations.

Graph RAG & Gen AI:-

What is GenAI?

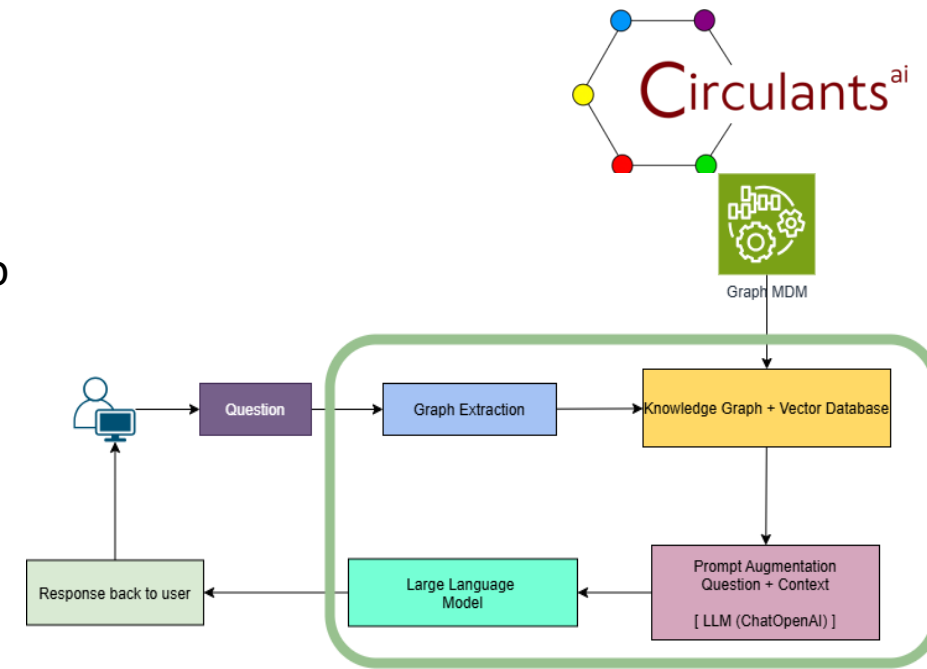
- GenAI (Generative AI) leverages Large Language Models (LLMs) like GPT, to generate human-like text based on prompts.
- LLMs are pre-trained on vast amounts of data to understand and generate diverse content.

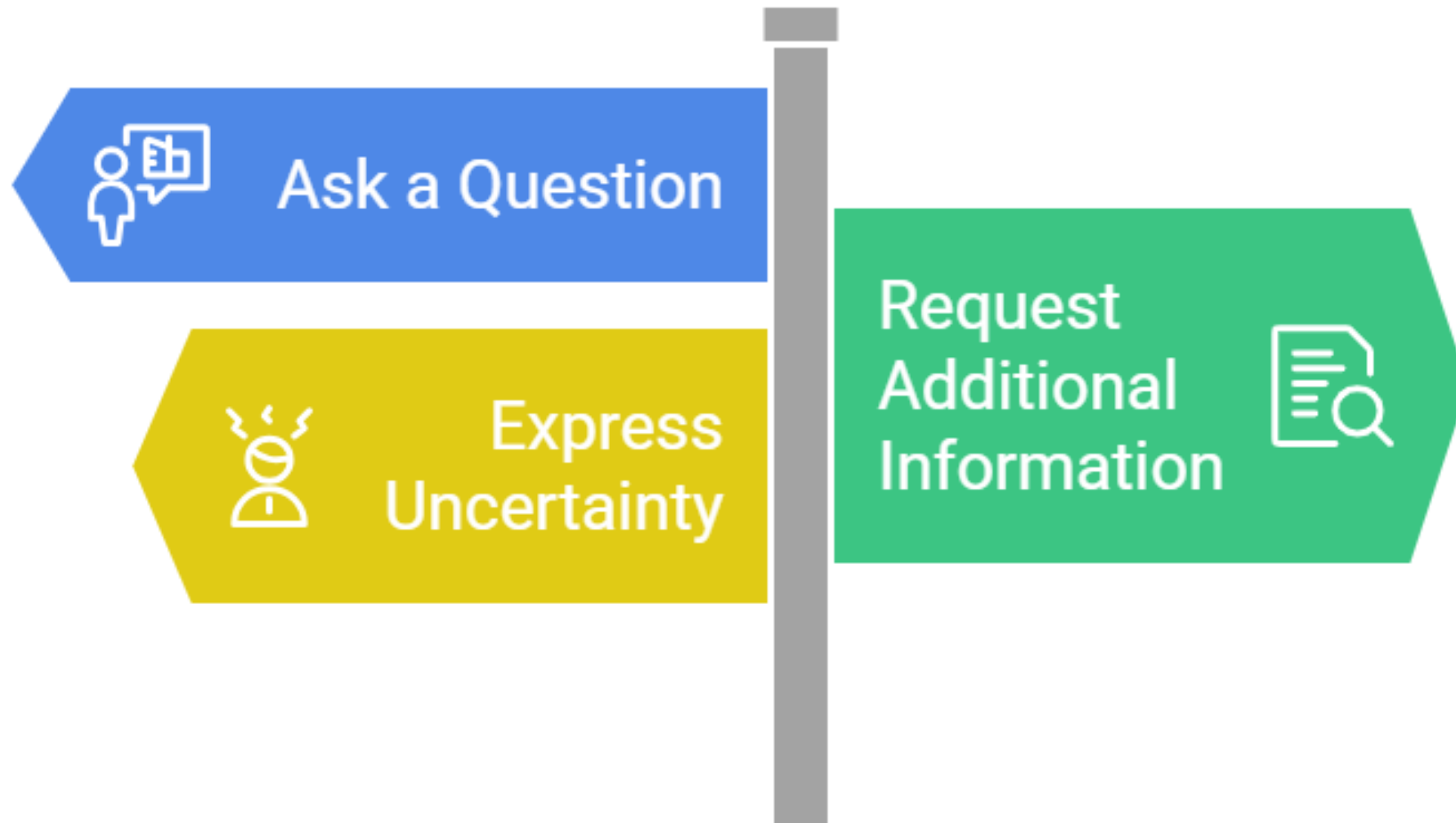
Graph RAG + GenAI Integration:

- **Combining RAG with LLMs:** LLMs generate more relevant and accurate responses by retrieving and processing data from knowledge graphs (RAG).
- The knowledge graph provides structured data and relationships, while vectors provide semantic context, enhancing the LLM's ability to generate context-specific answers.

Benefits:

- **Contextual understanding:** More relevant, data-backed answers due to graph-powered retrieval.
- **Enhanced content generation:** Produces smarter, more relevant outputs by using a mix of structured and unstructured data.
- **Democratizes Data Access:** Enables users at all levels (without advanced technical skills or programming knowledge) to easily access, retrieve, and utilize relevant information when needed, empowering data-driven decision-making across the organization.





Thank You



Pleasanton, CA | Bridgewater, NJ | Jaipur, India | Pune, India