neo4j

GraphRAG Workshop

Workshop Rules

Ask questions straight away, this is an interactive session

Raise your hand if you are stuck

Have fun

Before We Start

1/ Create a **blank-graph-data-science** Neo4j Sandbox at sandbox.neo4j.com

2/ Open the notebook in Colab (needs a **google account**!)

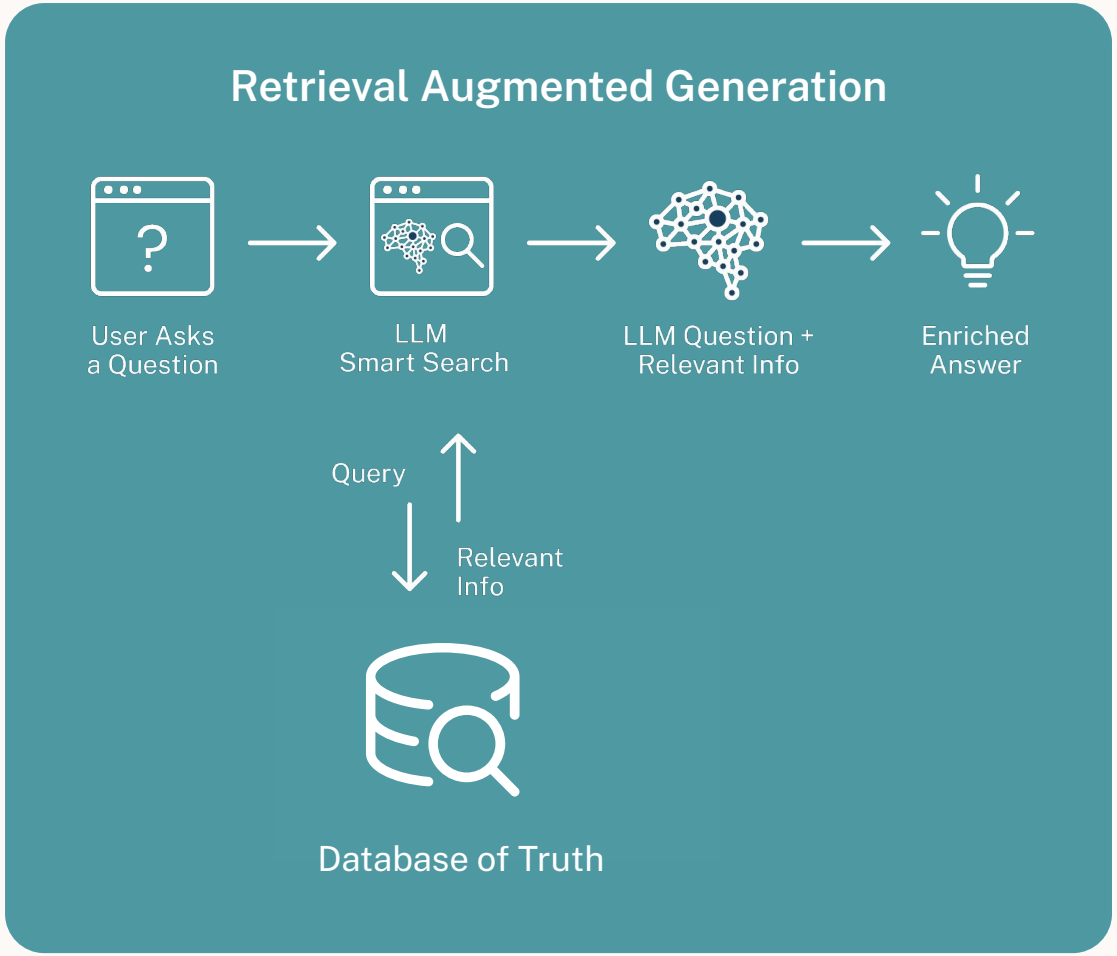
https://github.com/neo4j-product-examples/genai-workshop

open **genai-workshop.ipynb**

Retrieval-Augmented Generation Is Becoming an Industry Standard

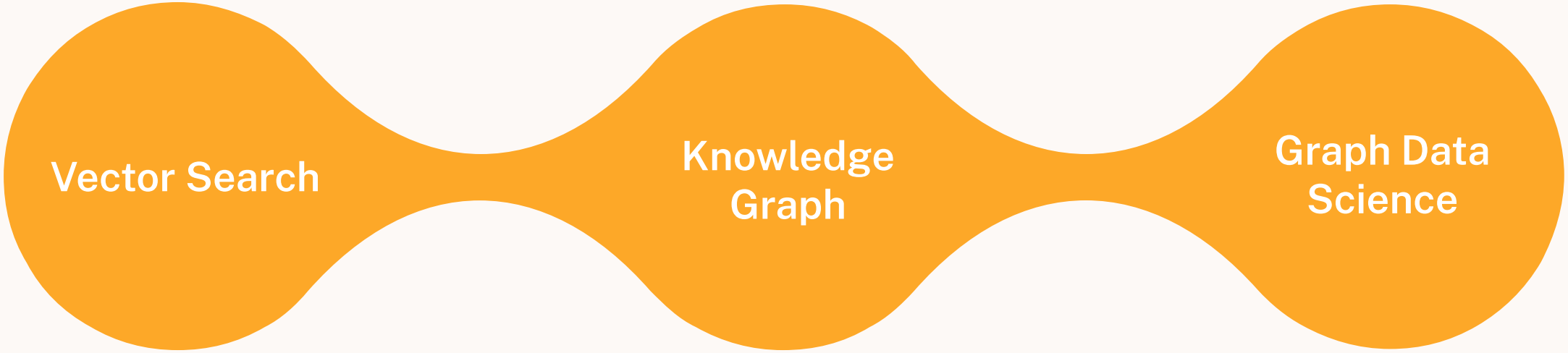
RAG augments LLMs by retrieving up-to-date, contextual external data to inform responses:

- Reduce hallucinations with verified data
- Provide domain-specific, relevant responses
- Enable traceability back to sources



GraphRAG with Neo4j

Unify vector search, knowledge graph and data science capabilities to improve RAG quality and effectiveness



Find similar documents and content

Identify entities associated to content and patterns in connected data

Improve & enrich GenAI insights. Discover new relationships and entities

Today We are Choosing Clothes

This is Eva

For her birthday, Eva wants
a **Halter Neck Top**

How do you choose what to buy?

This is Eva

For her birthday, Eva wants
a **Halter Neck Top**

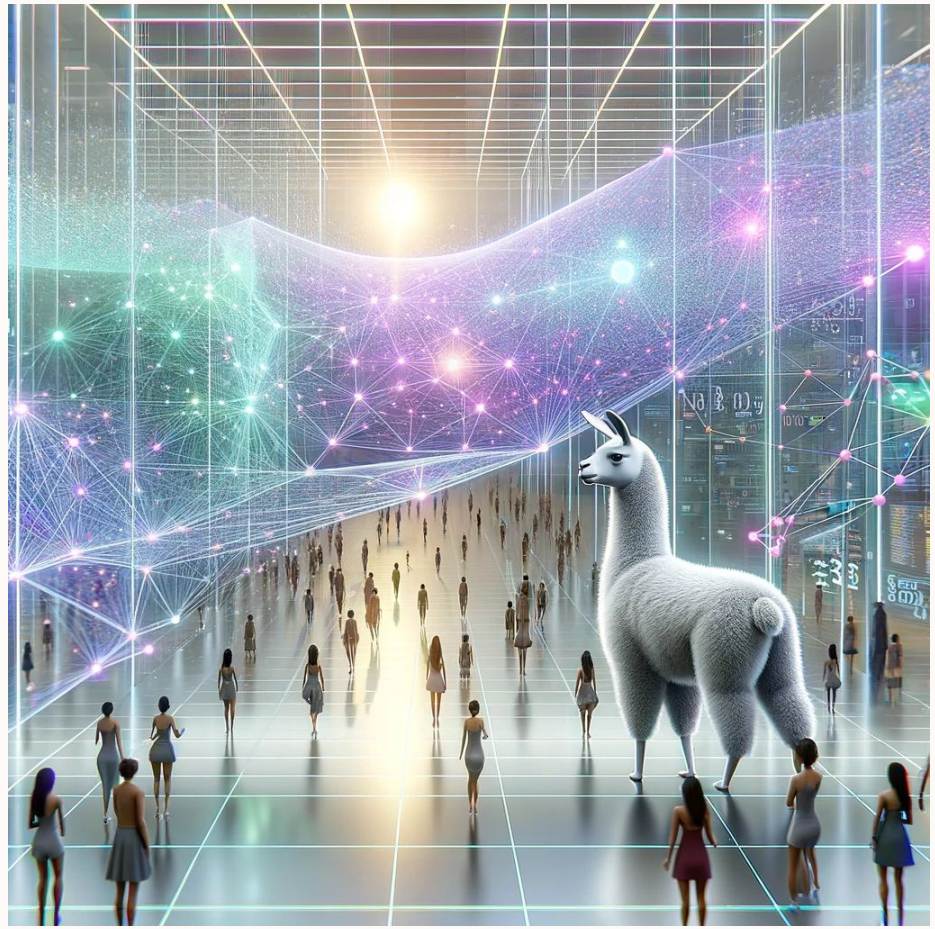
How do you choose what to buy?

1/ Get a list of Halter Neck Top
2/ Match the Halter Neck Top for Eva
3/ Accessorize or add another item

At Scale: Let's Build an AI Fashion Assistant

This Requires:

1. Data: Kaggle H&M
2. Search & Retrieval
3. Context for Personalization
4. Recommendation Engine
5. LLM Powered Content Generator



Solution: LLMs and Knowledge Graphs

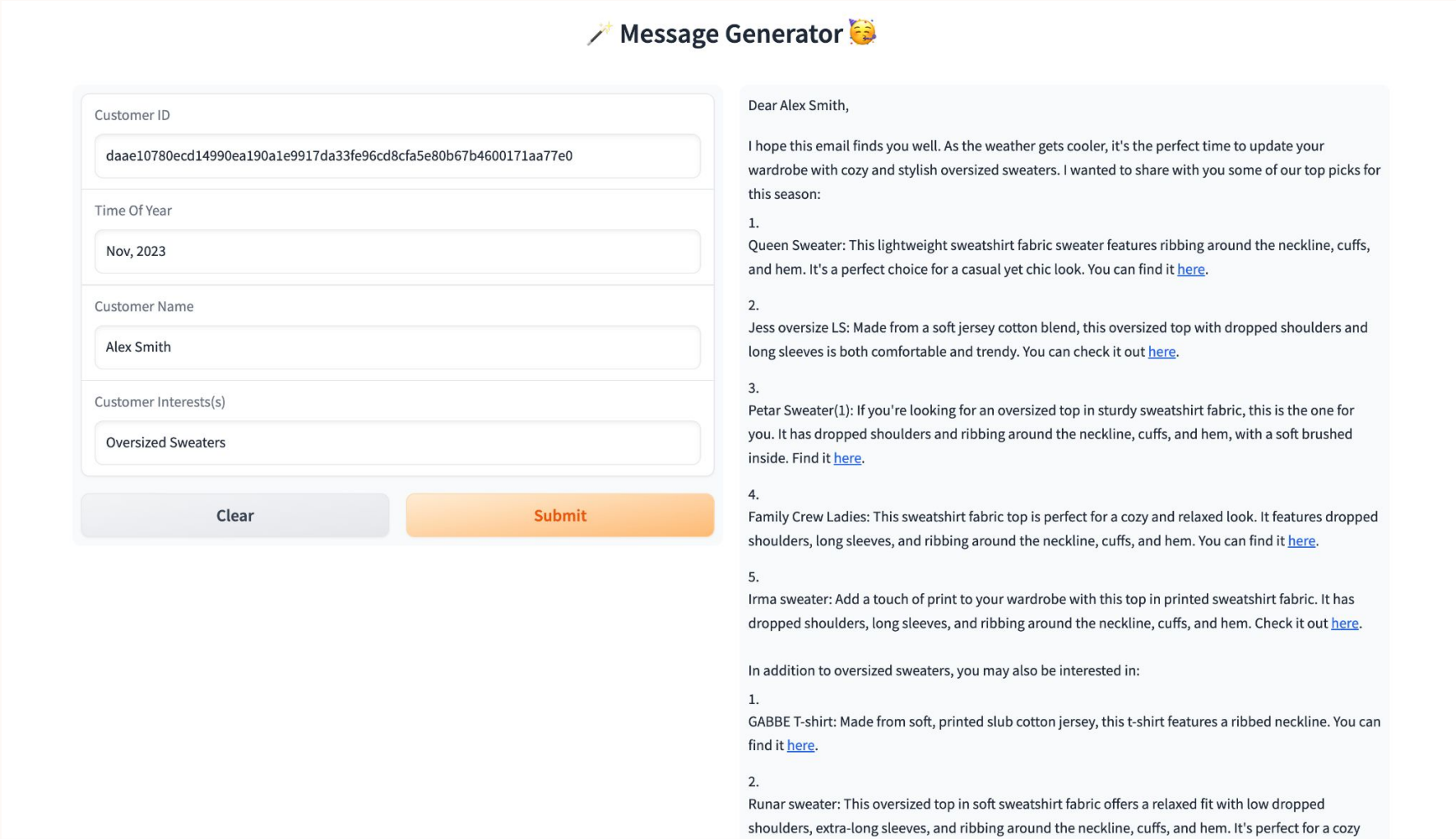
Knowledge
Context
Enrichment

KGs + LLMs

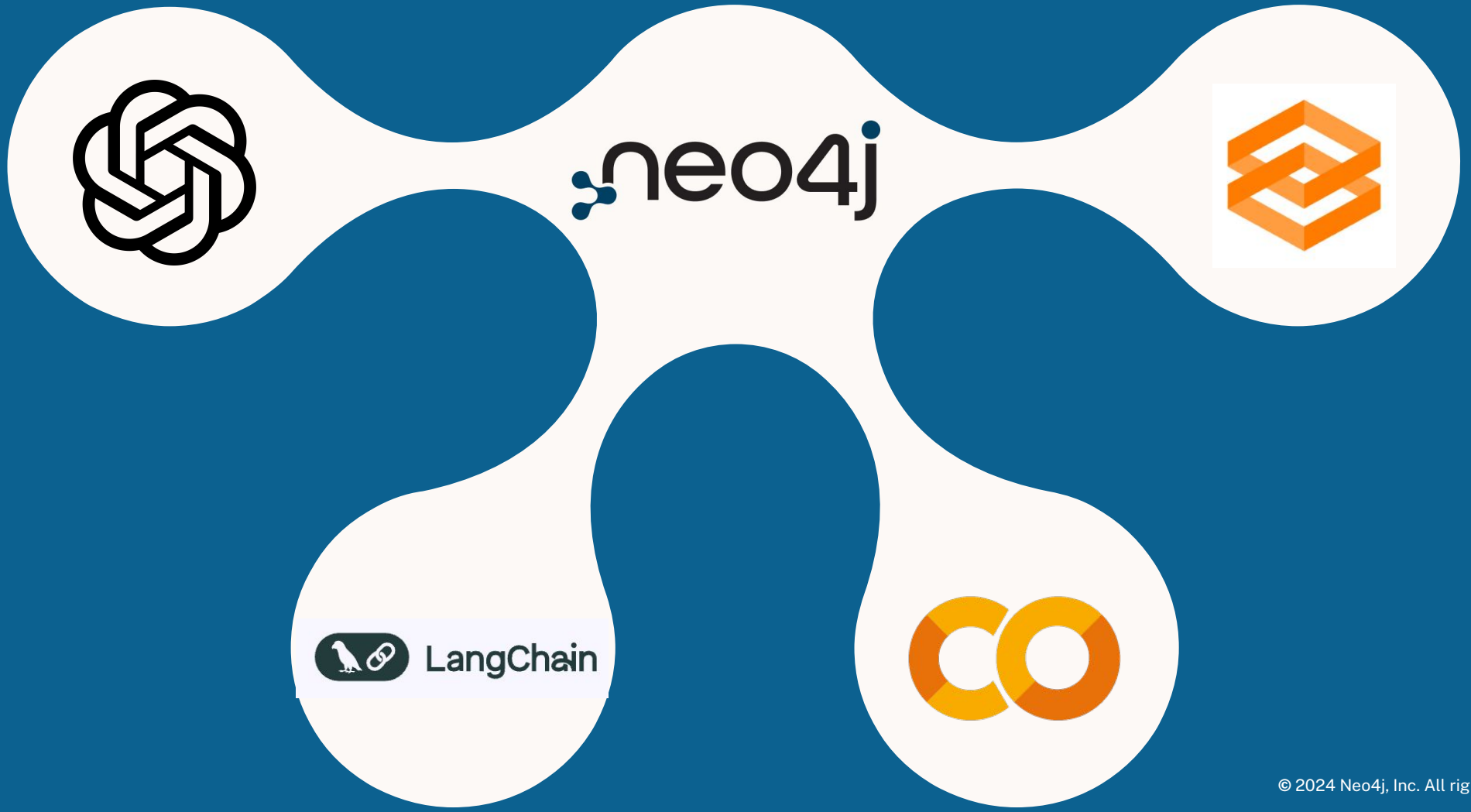
Language
Reasoning
Creativity

BETTER TOGETHER

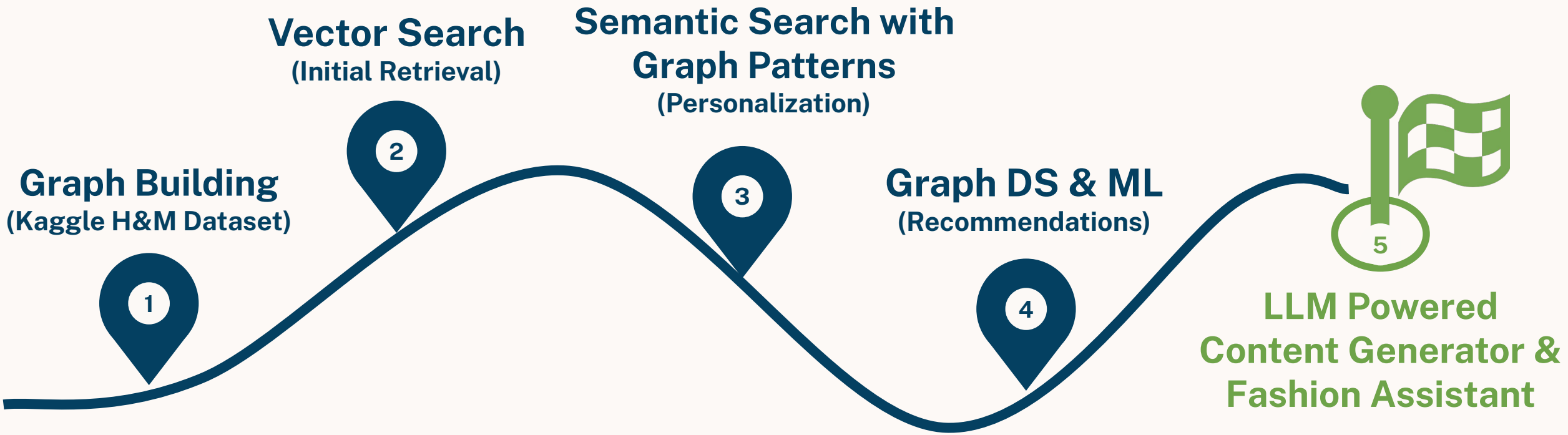
We'll Build a Targeted, Personalized Content Generator

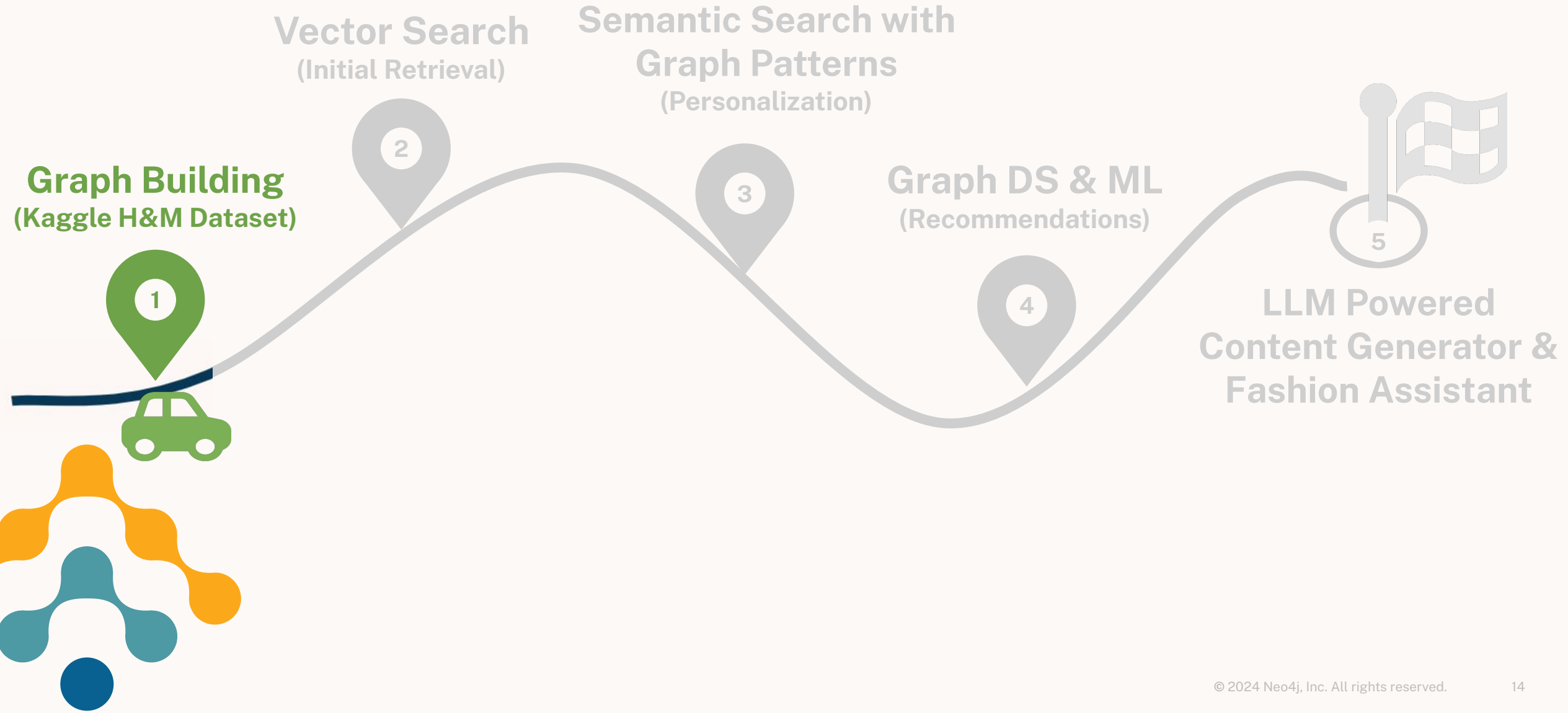


Tools

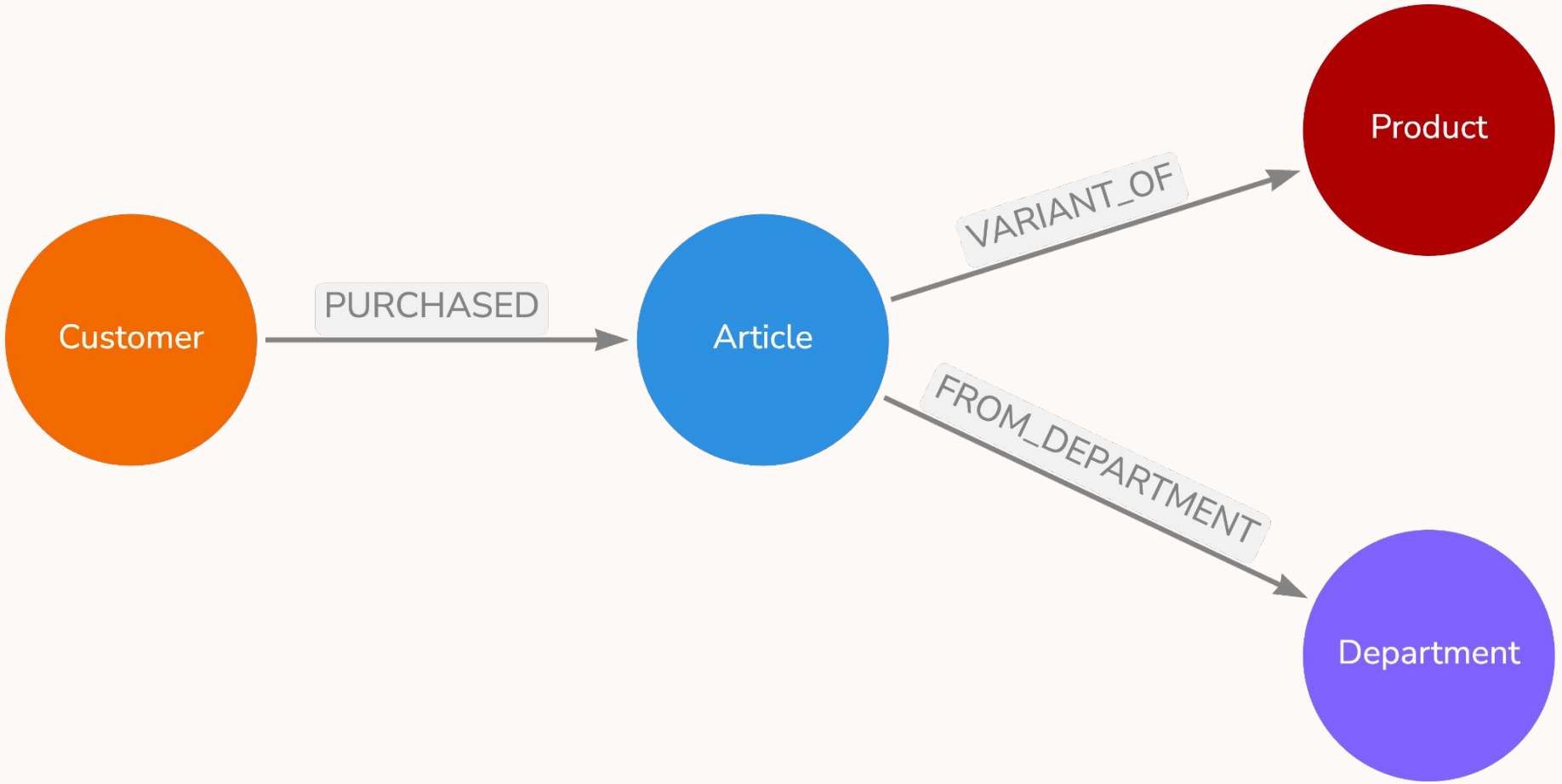


Journey

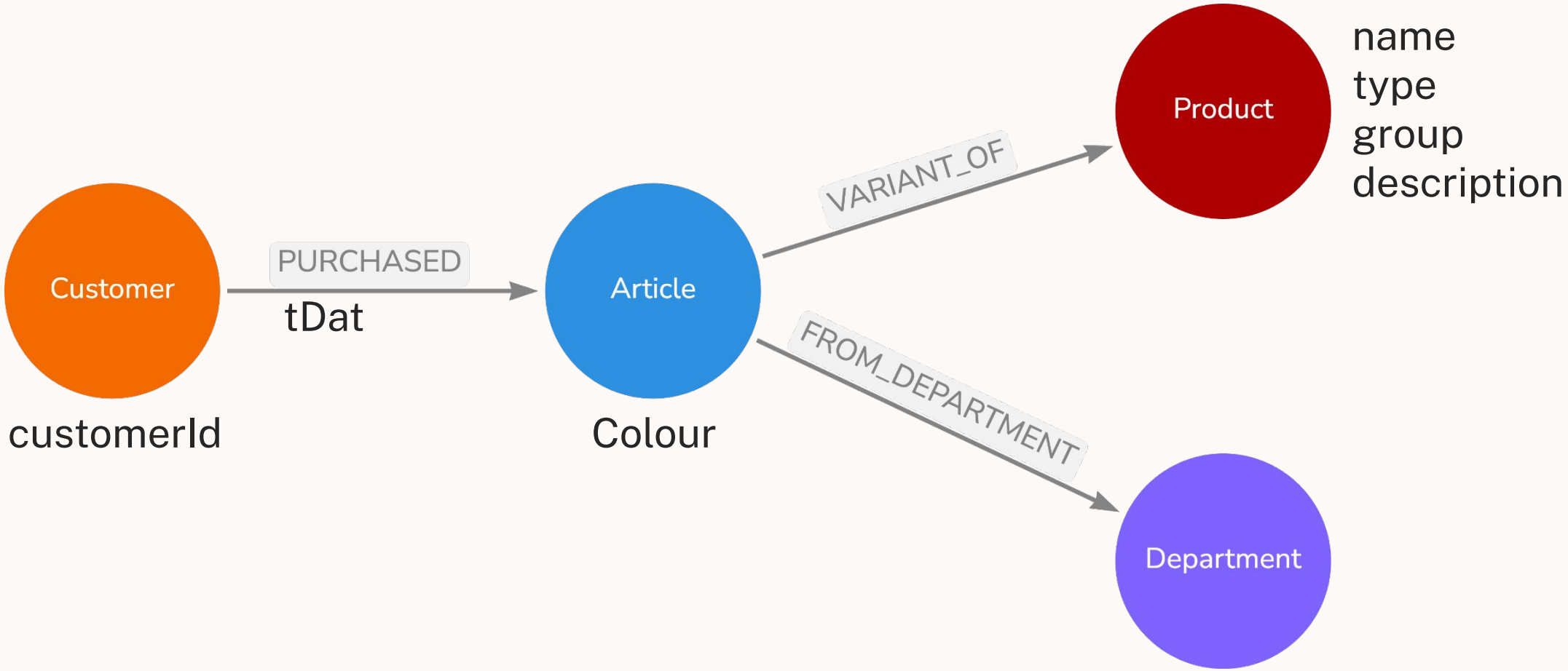




Data Model



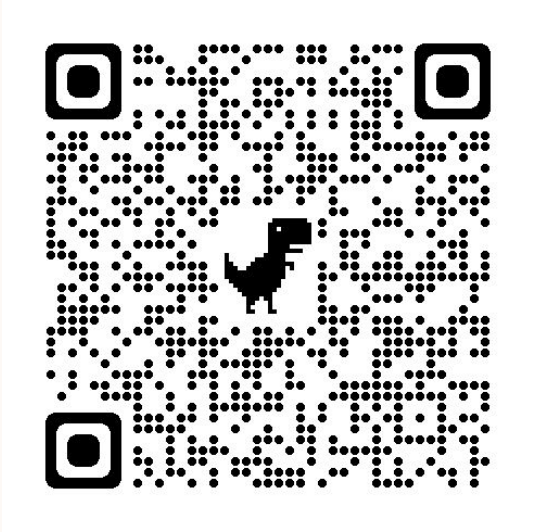
Data Model



Now We Can Start

1/ Create a **blank** Neo4j GDS Sandbox at sandbox.neo4j.com

2/ Open the notebook in Colab (needs a **google account**!)

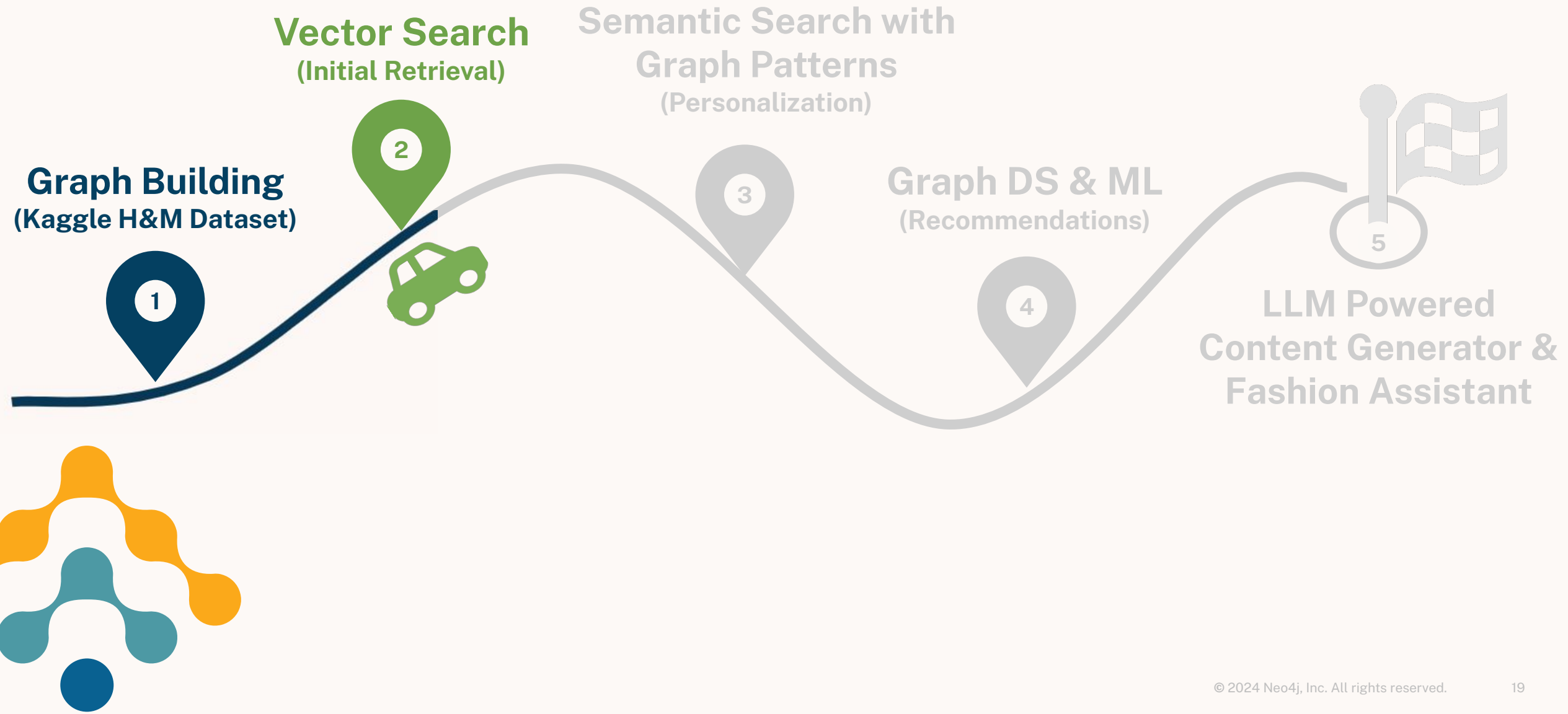


https://github.com/neo4j-product-examples/genai-workshop

open **genai-workshop.ipynb**

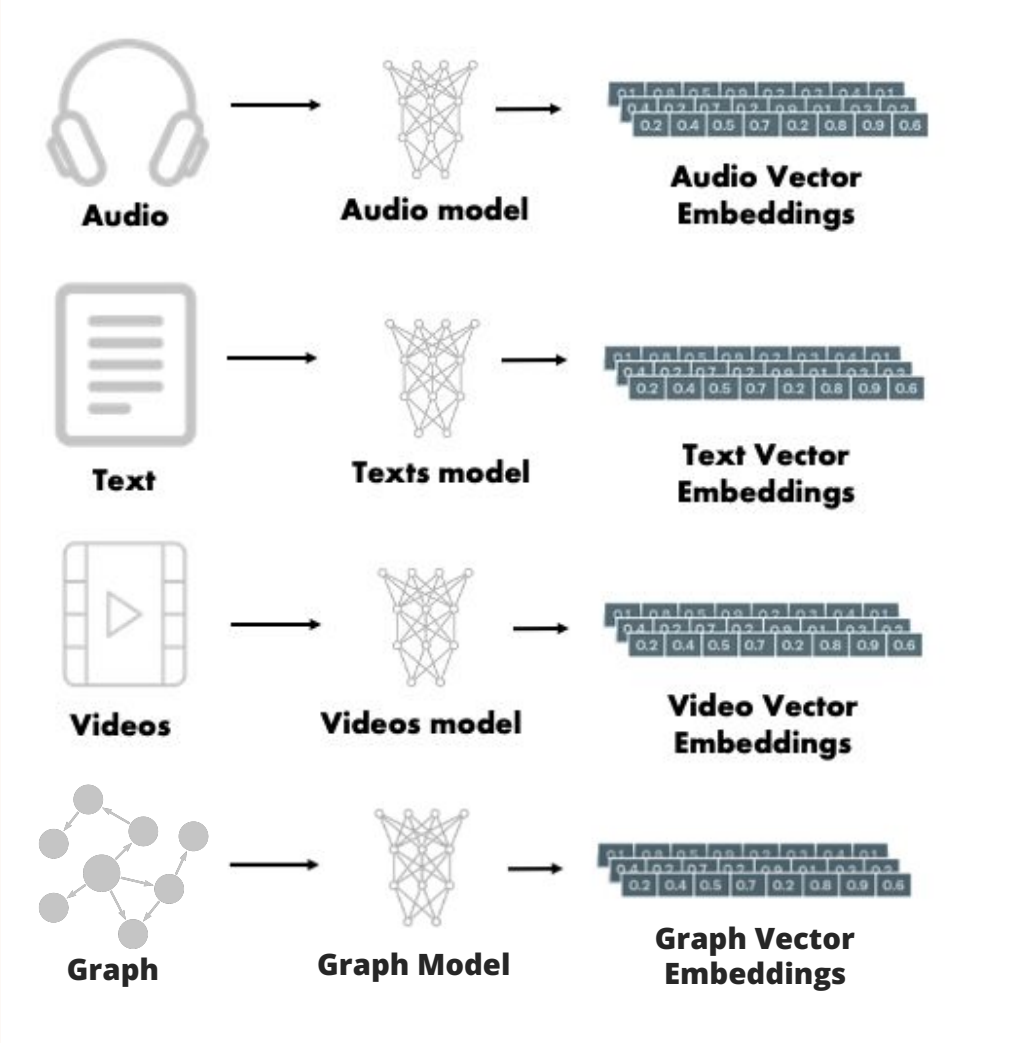
1/ Graph Building

Stop before **Vector Search**



What are Embeddings?

- A type of data compression
- Transform messy data (text, images, audio, etc.) into a compact format for ML algorithms
- Most often numeric **vectors** (a.k.a arrays) with 100s or 1000s of elements
- Preserve information such that “similar” items have proportionally “similar” embedding vectors
- Similarity is measured with vector algorithms (cosine, euclidean, etc.)



What Does “Similarity” Mean?

It Depends:

- Text Embeddings => Semantic Similarity, the meaning behind a text sequence

- Graph Embeddings => similarity in position or structure in a graph - can have semantic meaning too

What will we use embedding for?

1. **(now) Vector search** using text embeddings

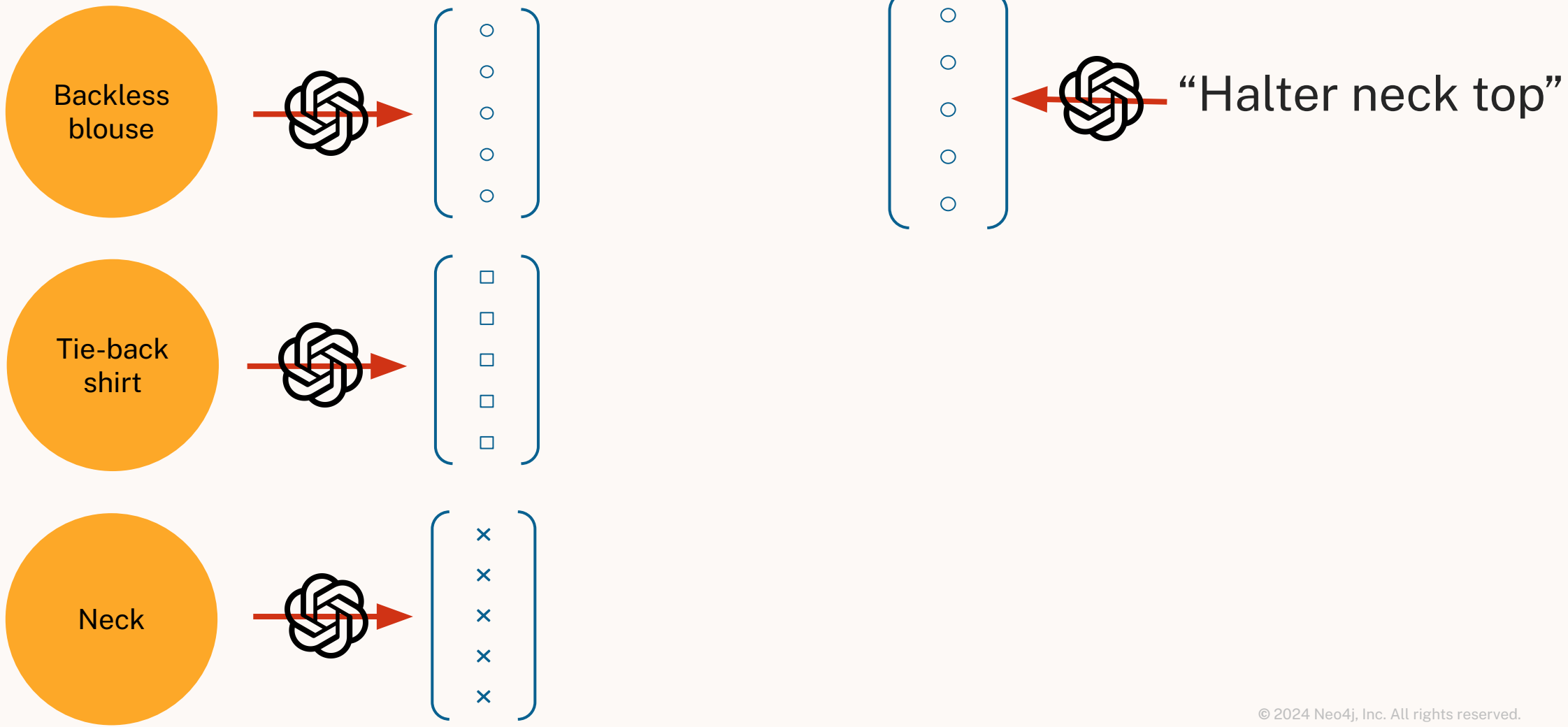
2. **(later) Recommendations** using graph node embeddings

Search & Vectors in Neo4j

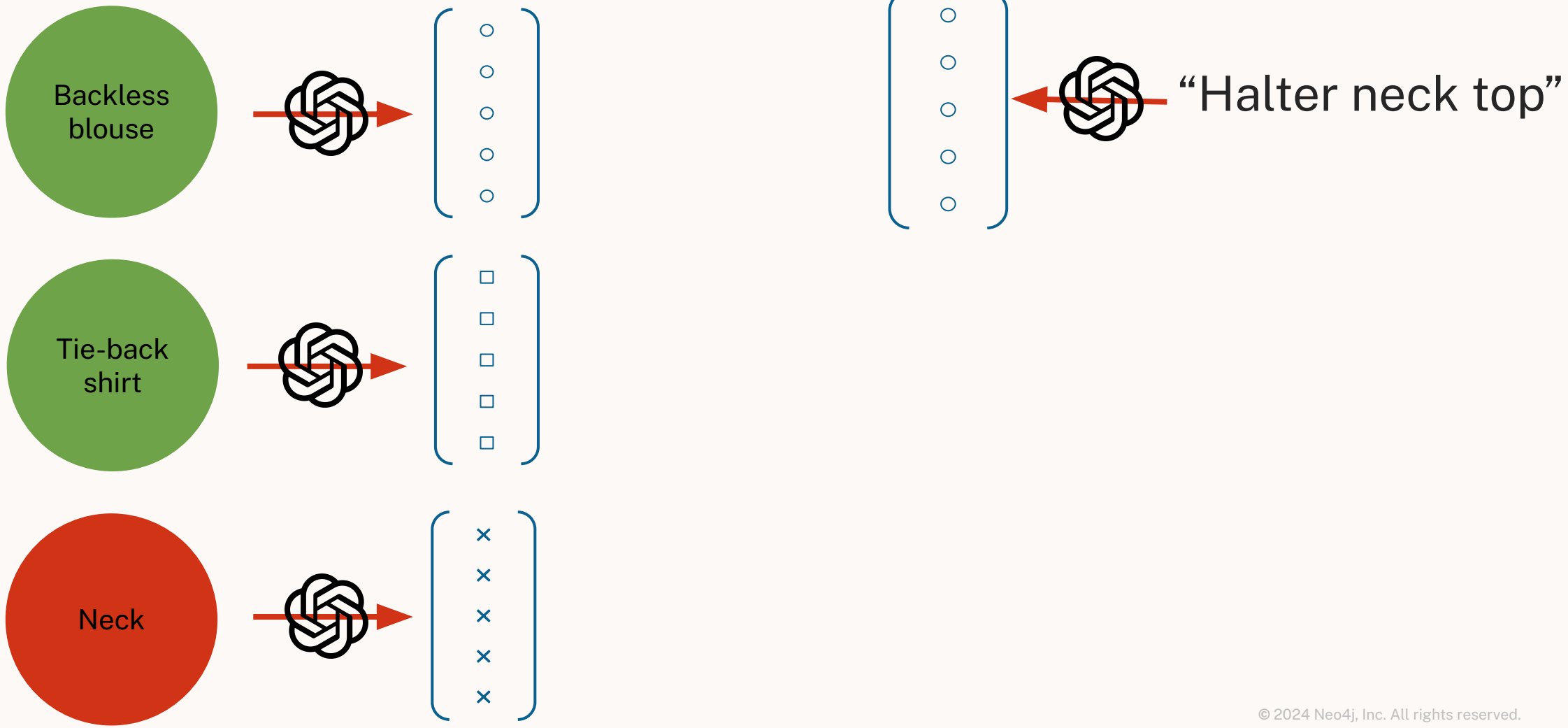
Neo4j makes search efficient through a variety of indexes including

- **Range:** General index for predicates based on equality and range. Numeric, dates, etc.
- **Point:** Predicts on geospatial points like distance bounding boxes, etc.
- **Text:** Predicates on strings like contains, ends with, etc.
- **FullText:** Text search based on tokenization and analyzers
- **Vector:** ANN (Approximate Nearest Neighbor) search on vectors

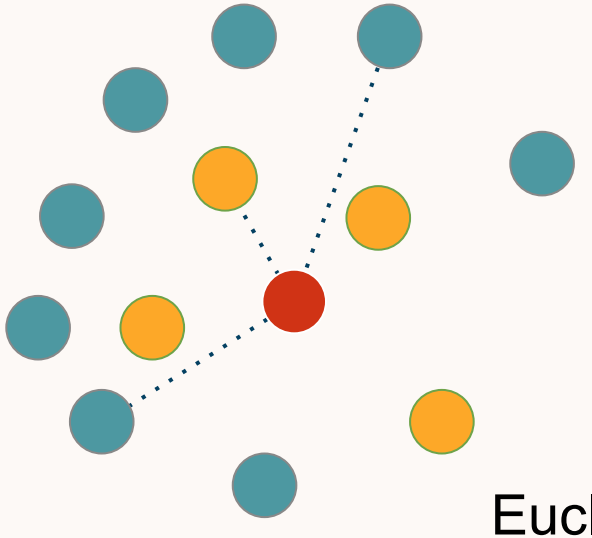
Vector search



Vector search

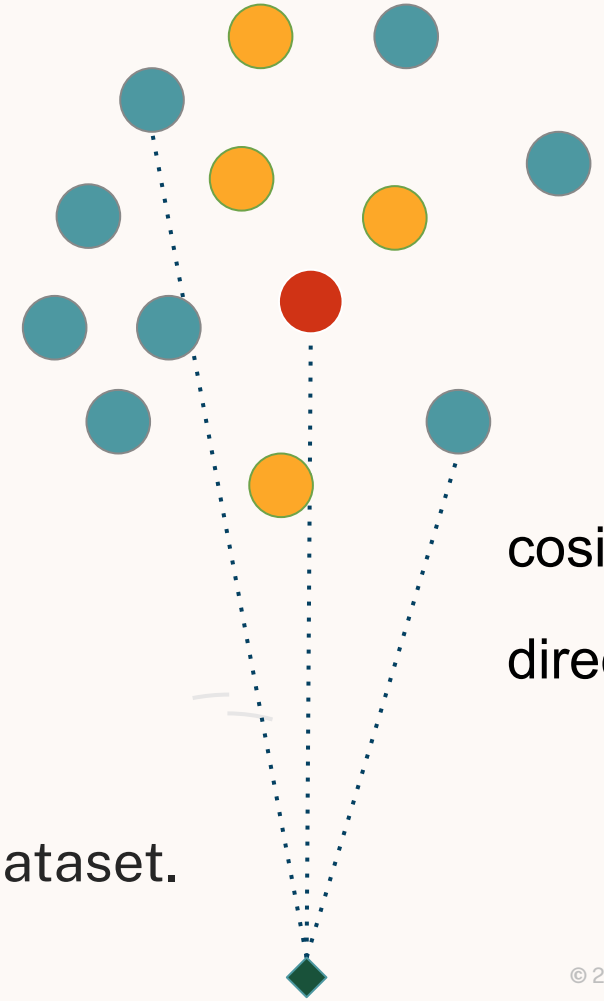


How is similarity measured



Euclidean

distance based

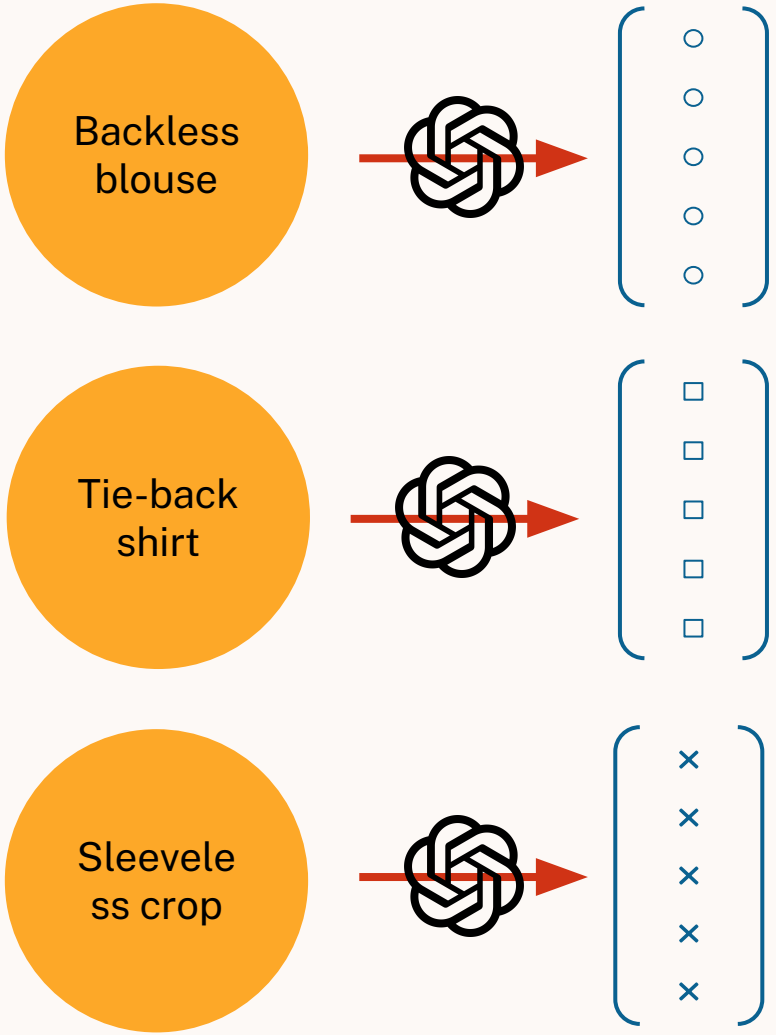


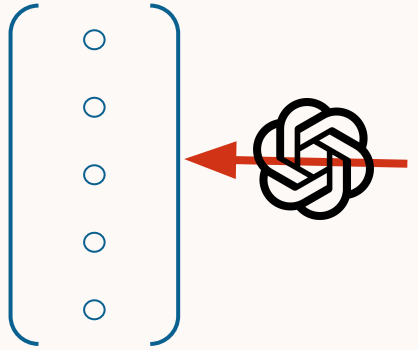
cosine

direction / angle based

The vector Index allow to query vectors on large dataset.
It uses HNSW to perform k-ANN queries

Vector search

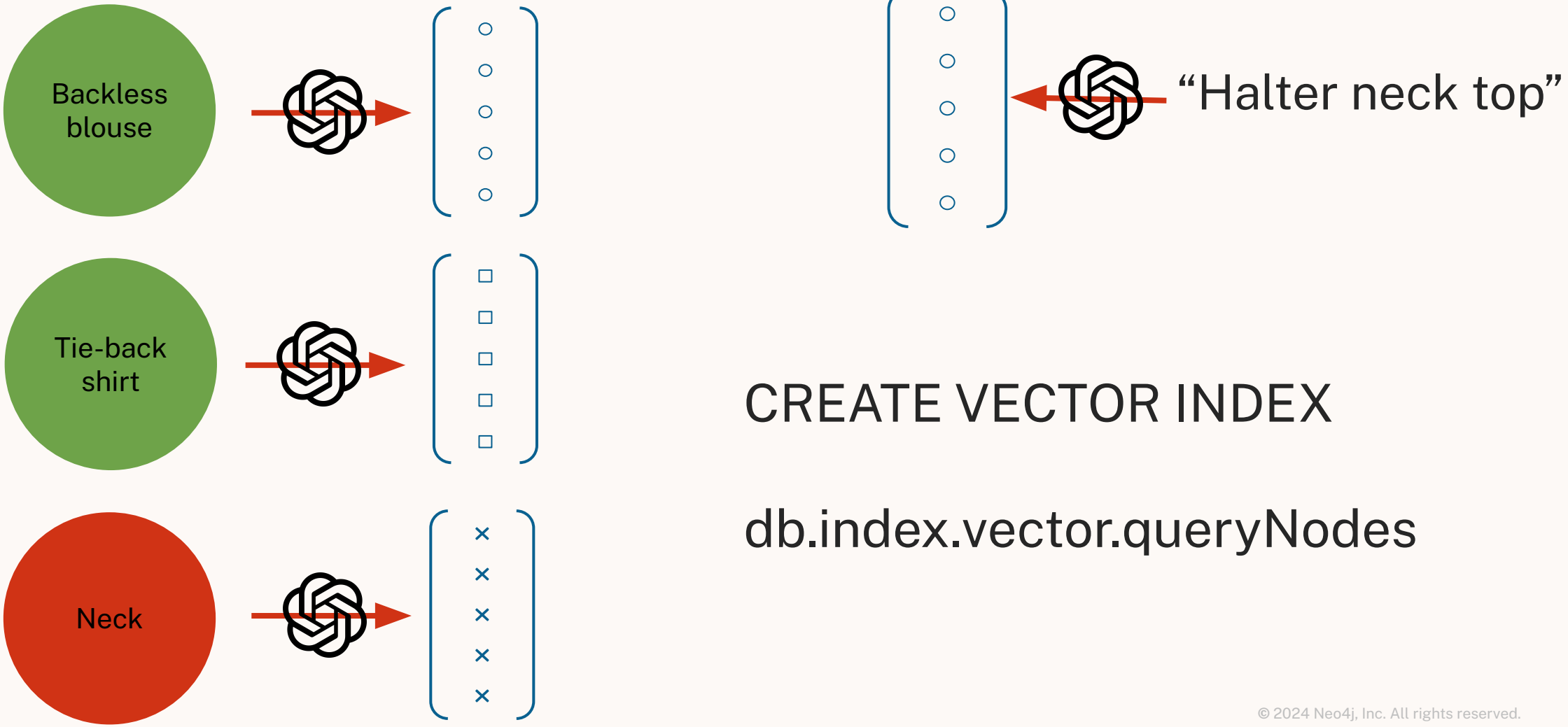




CREATE VECTOR INDEX

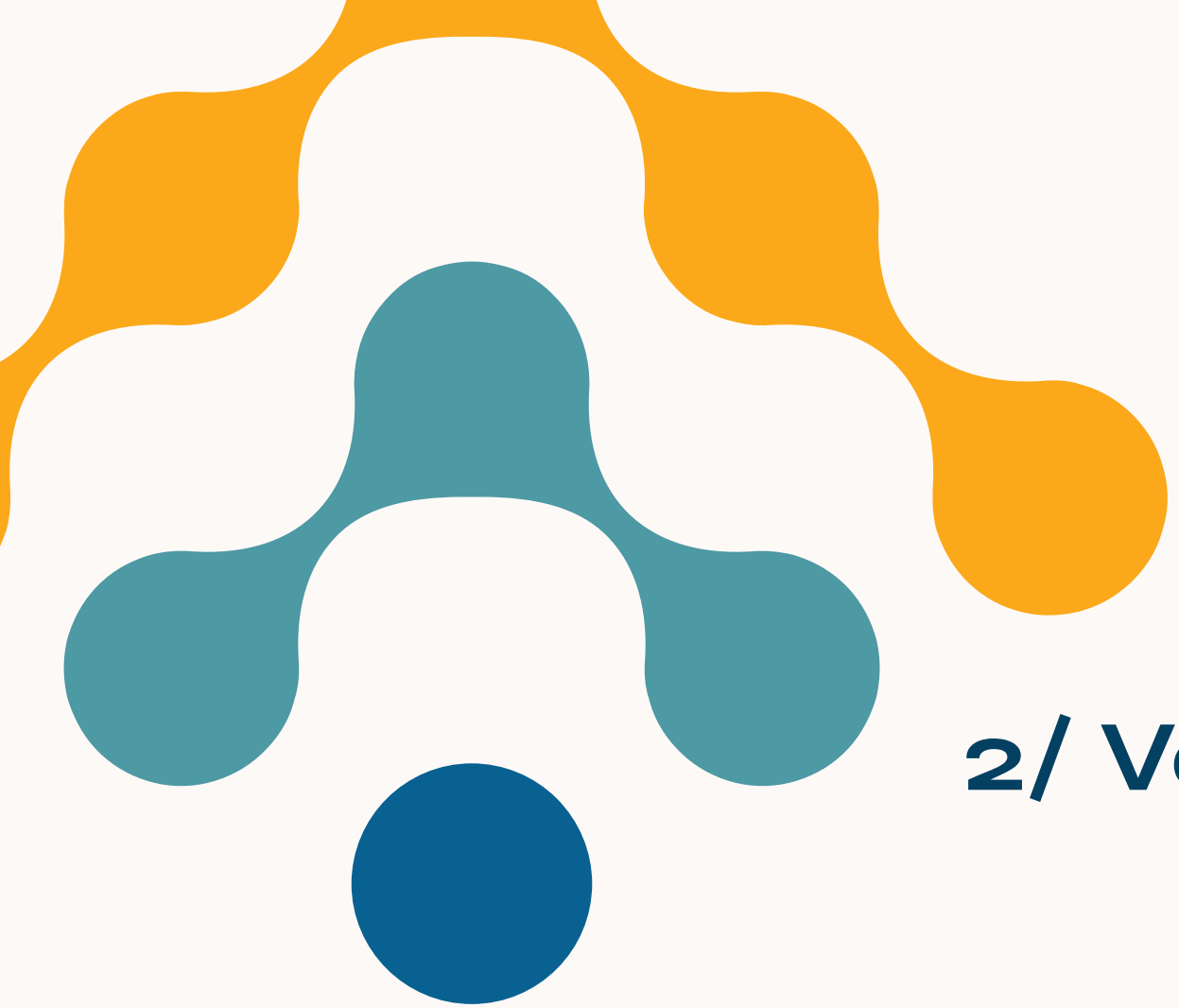
db.index.vector.queryNodes

Vector search



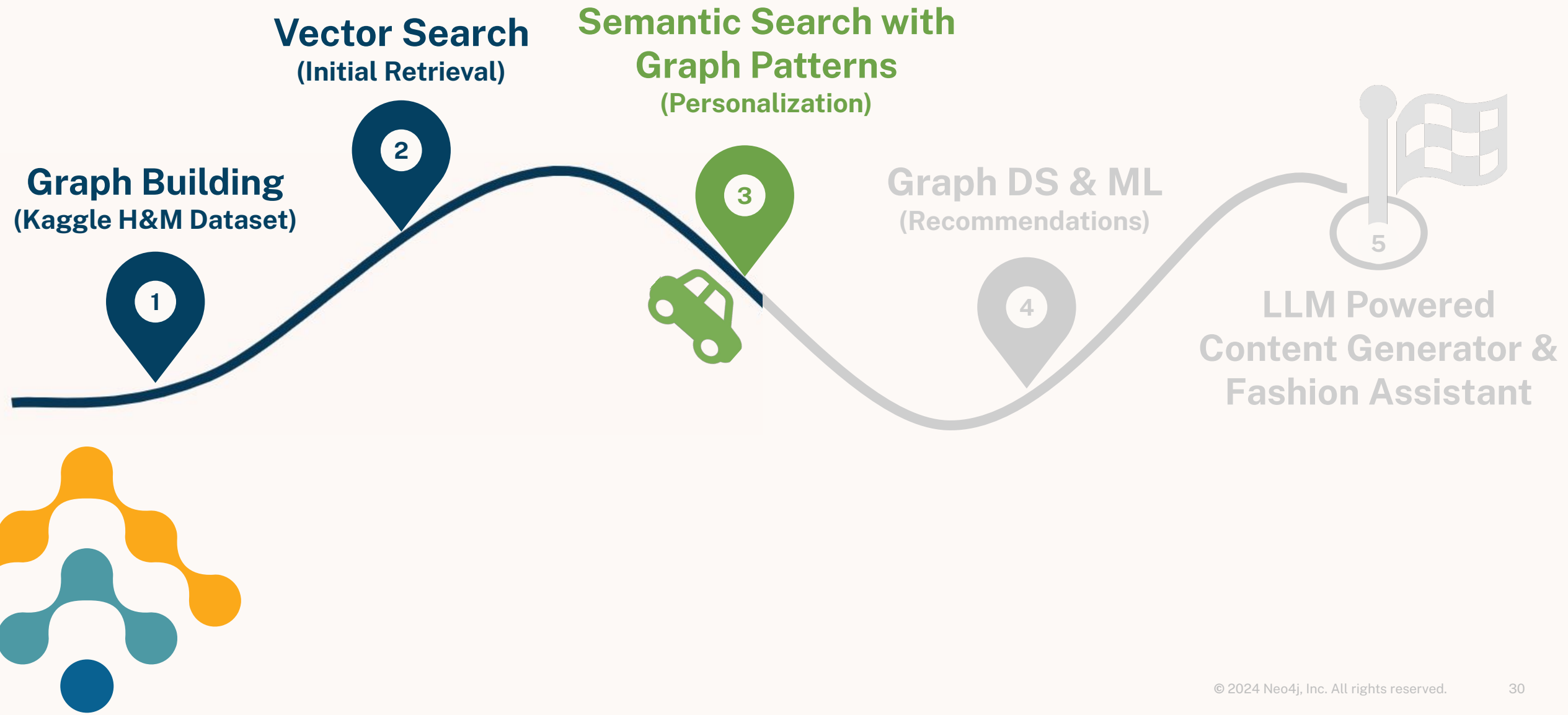
CREATE VECTOR INDEX

db.index.vector.queryNodes



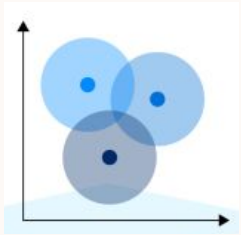
2/ Vector Search

Stop before **Semantic Search with Graph Patterns (Personalization)**



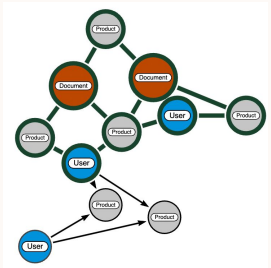
Neo4j & Semantic Search

Vector Similarity Search



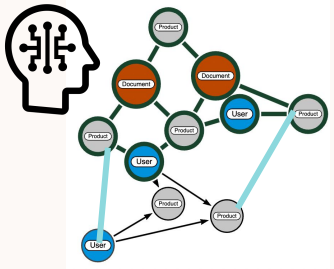
Find relevant documents and content for user queries

Graph Pattern Matching



Find entities associated to content and patterns in connected data.

Knowledge Graph DS/ML

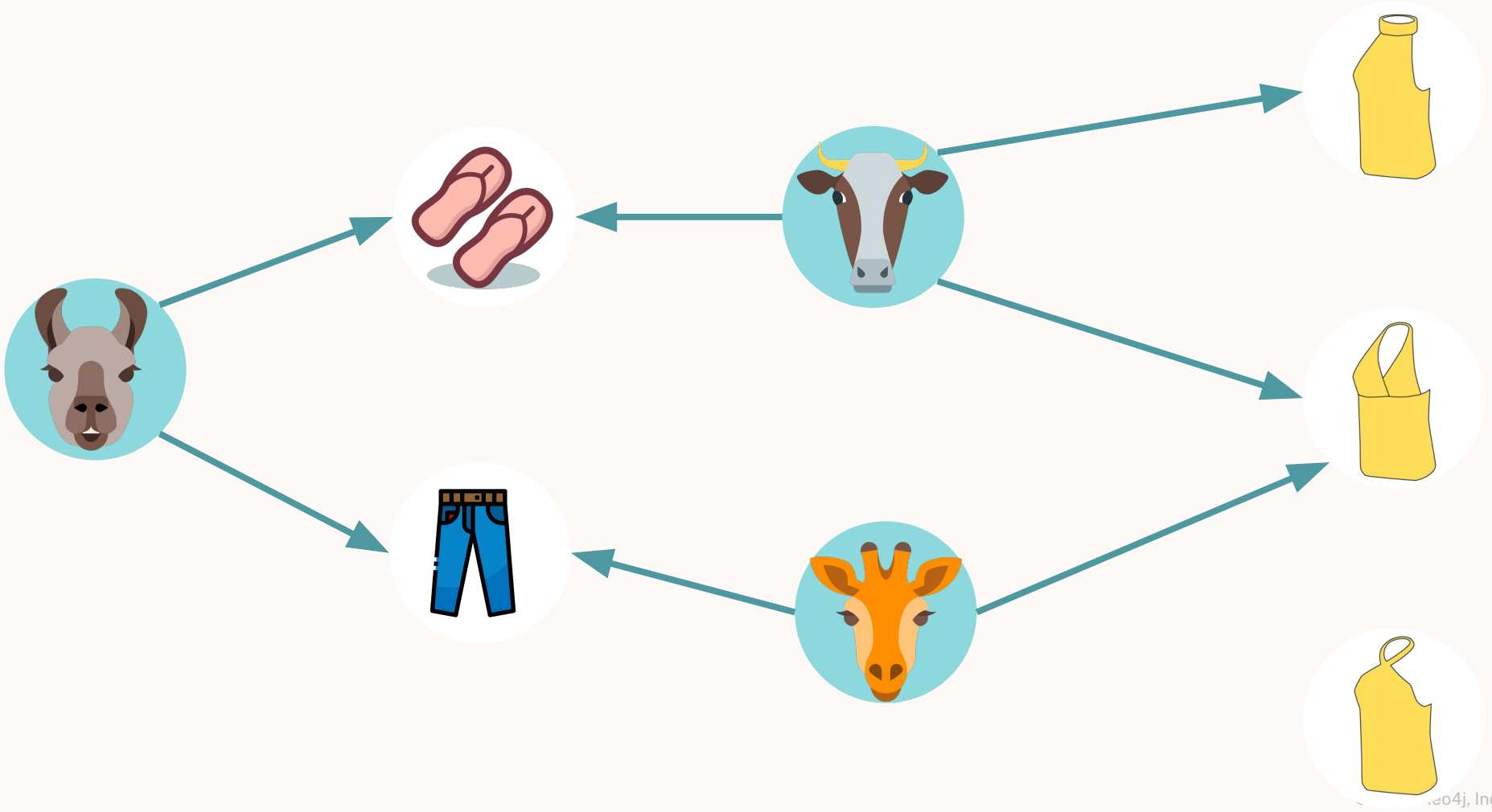


Improve search relevance & insights by enhancing a Knowledge Graph. Use graph algorithms and ML to discover new relationships, entities, and groups.

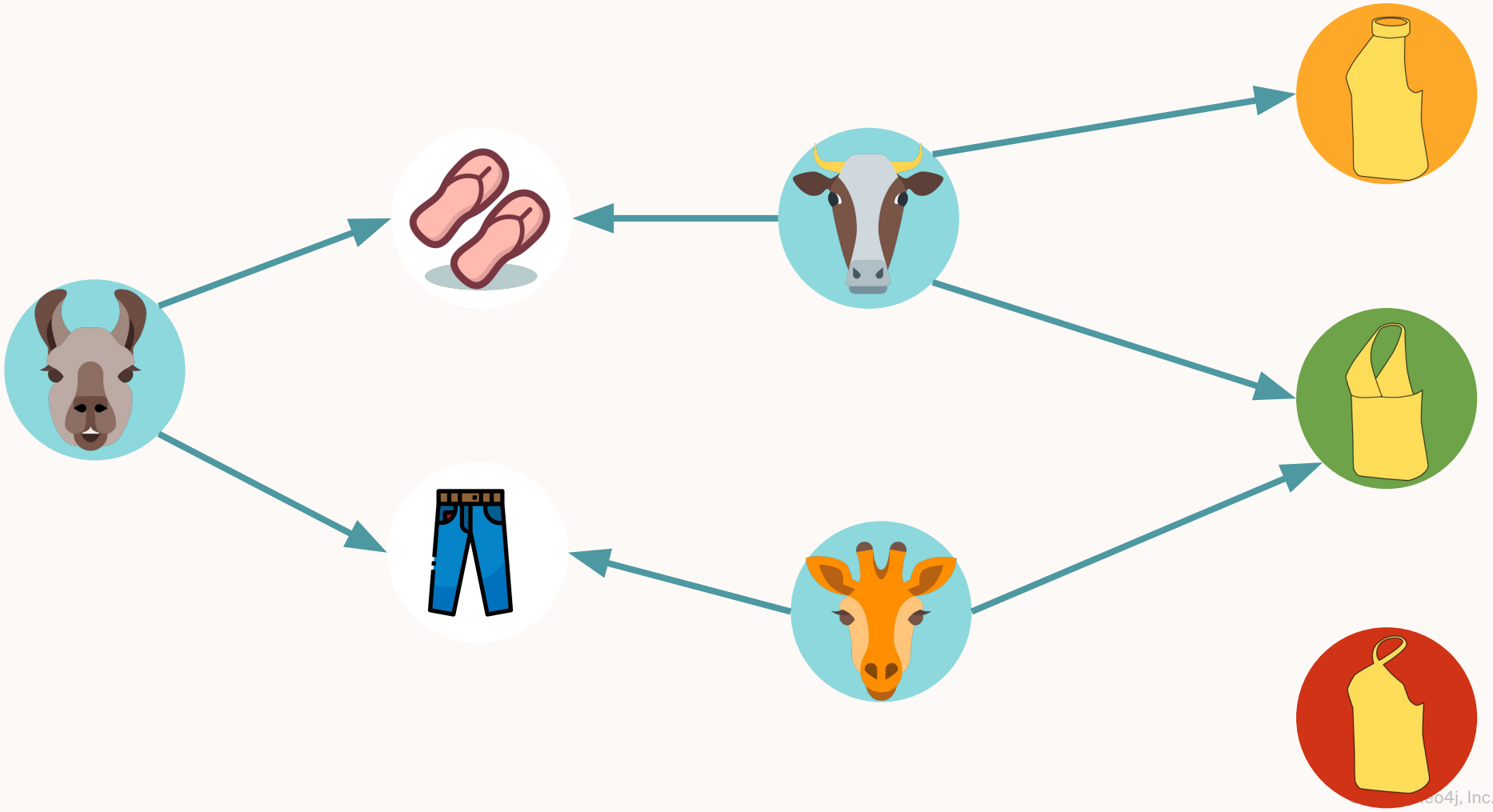
Vector Search

Graph Database

Semantic search with graph pattern



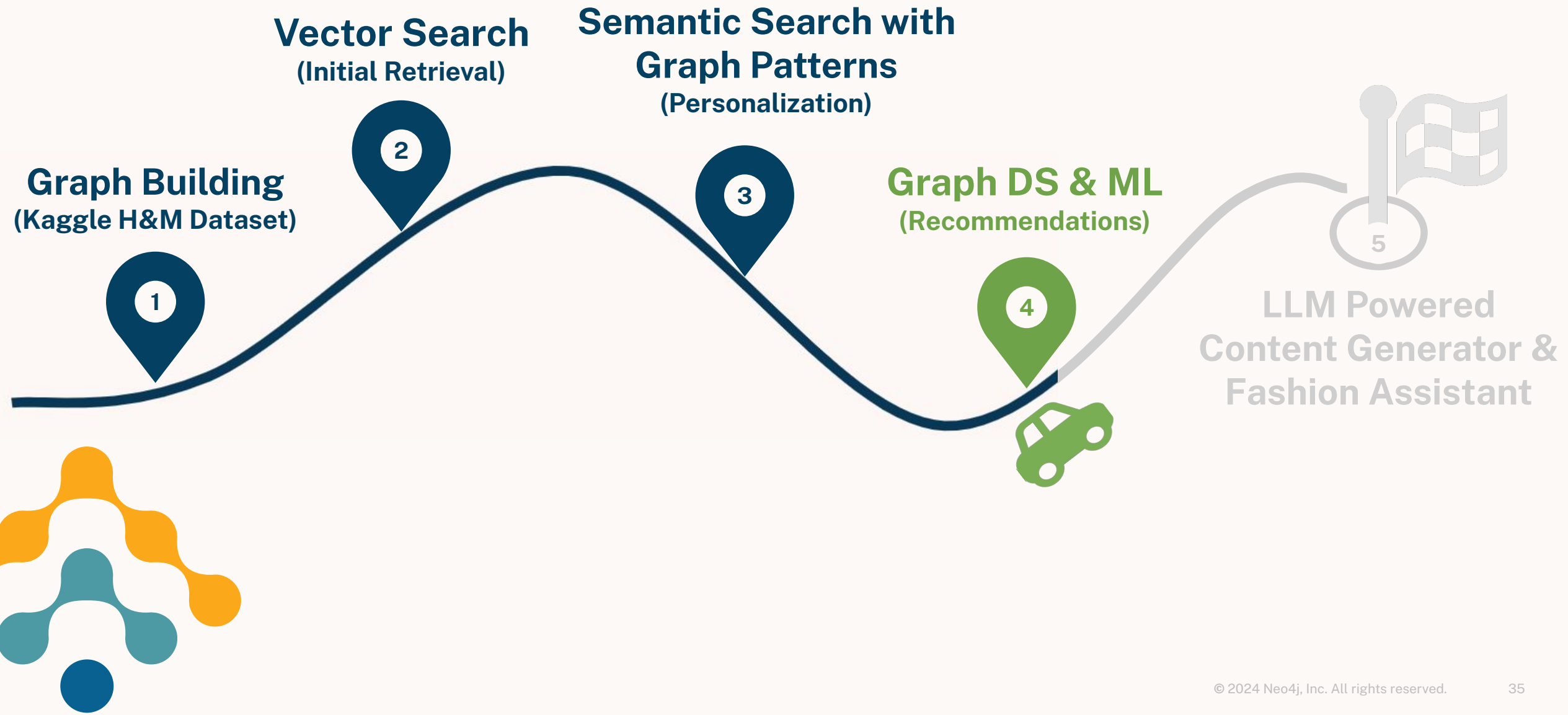
Semantic search with graph pattern



3/ Semantic Search with Graph Patterns

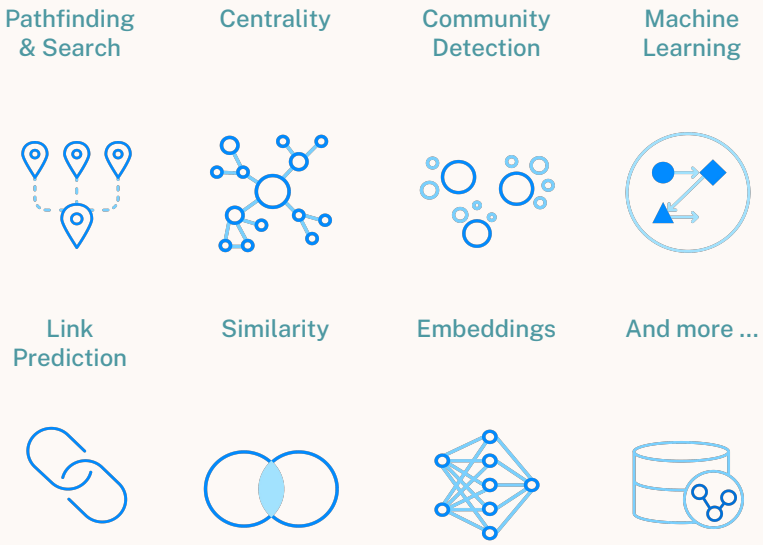
Personalization & Context

Stop before **Graph Data Science & ML (Recommendations)**



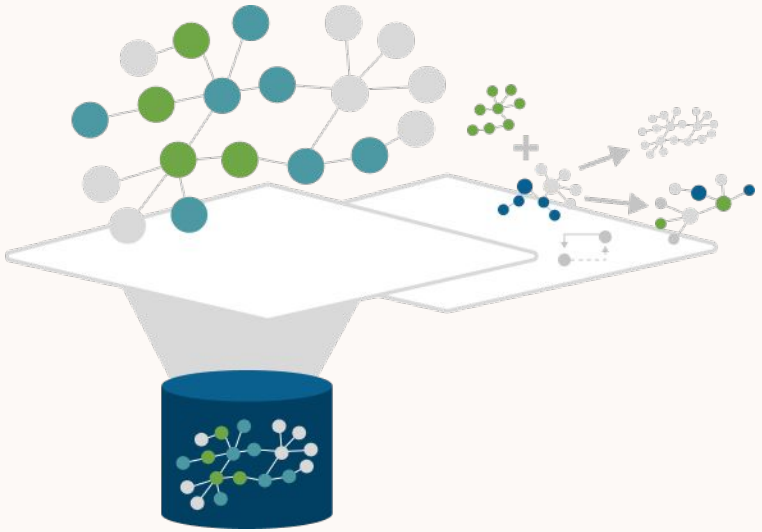
Neo4j Graph Data Science

The Largest Catalog of Graph Algorithms



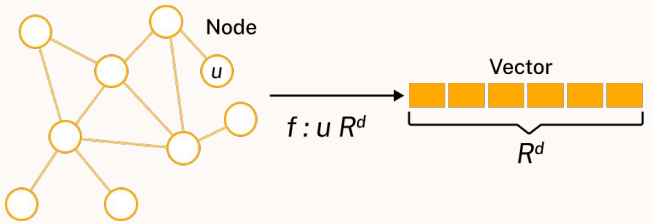
Over 65 pretuned, parallelized algorithms.

Native Graph Catalog and Analytics Workspace

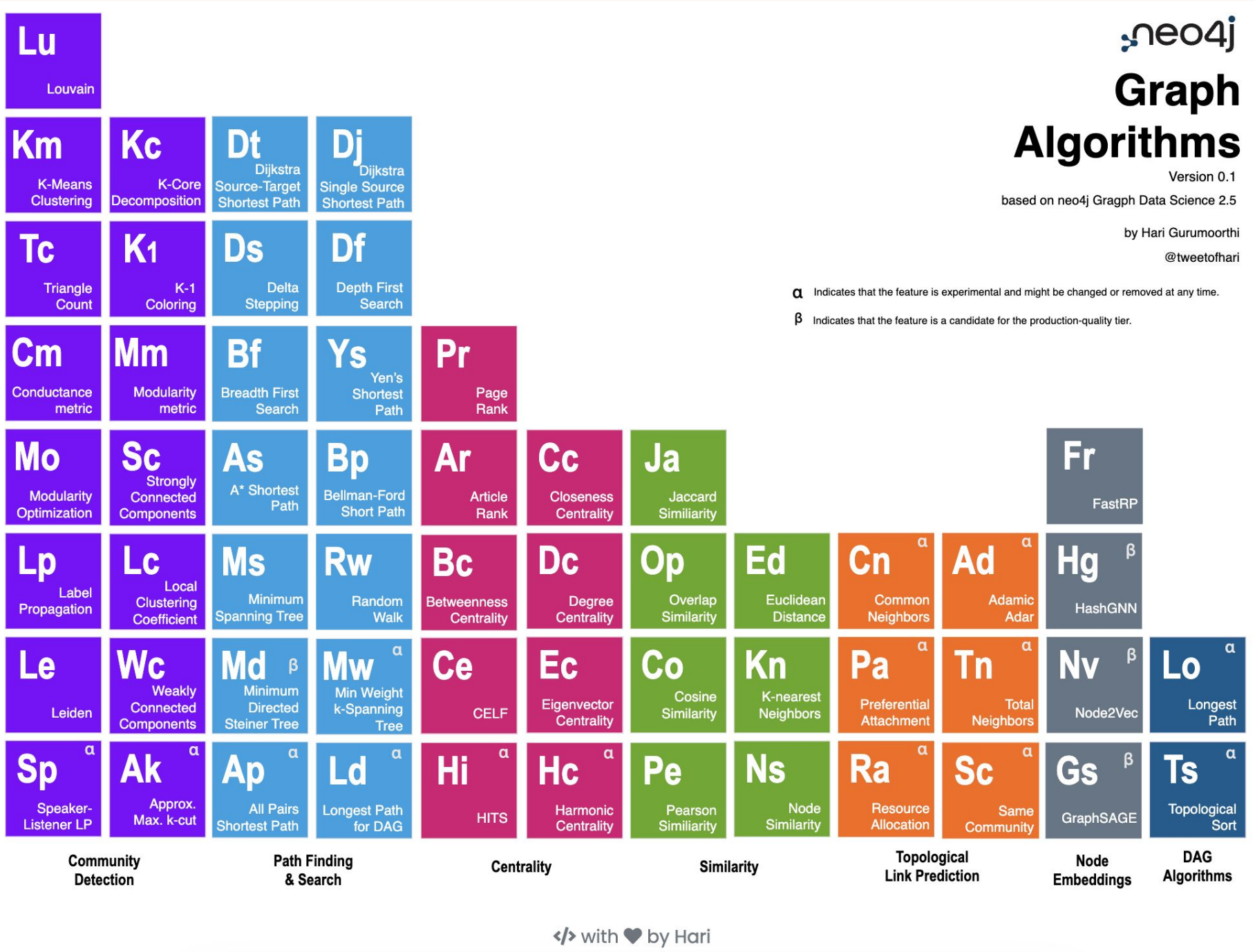


Iterate fast with different data sets, models, and version trained models.

Graph Embeddings for Machine Learning

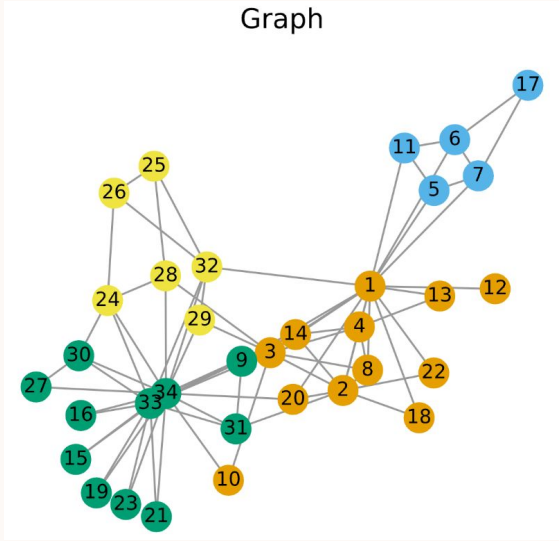


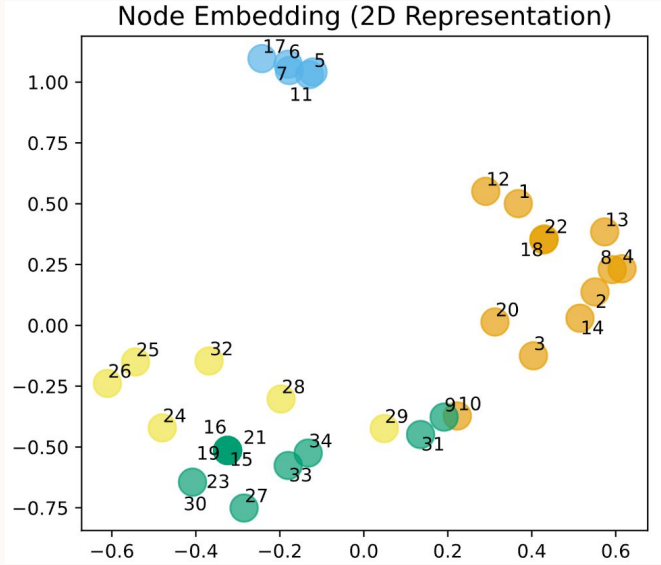
Bring the context of your connected data into a format that other pipelines can ingest.

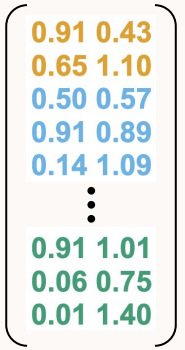


Node Embeddings

*The representation of nodes as low-dimensional **vectors** that summarize their graph position, the structure of their local graph neighborhood as well as any possible **node features***

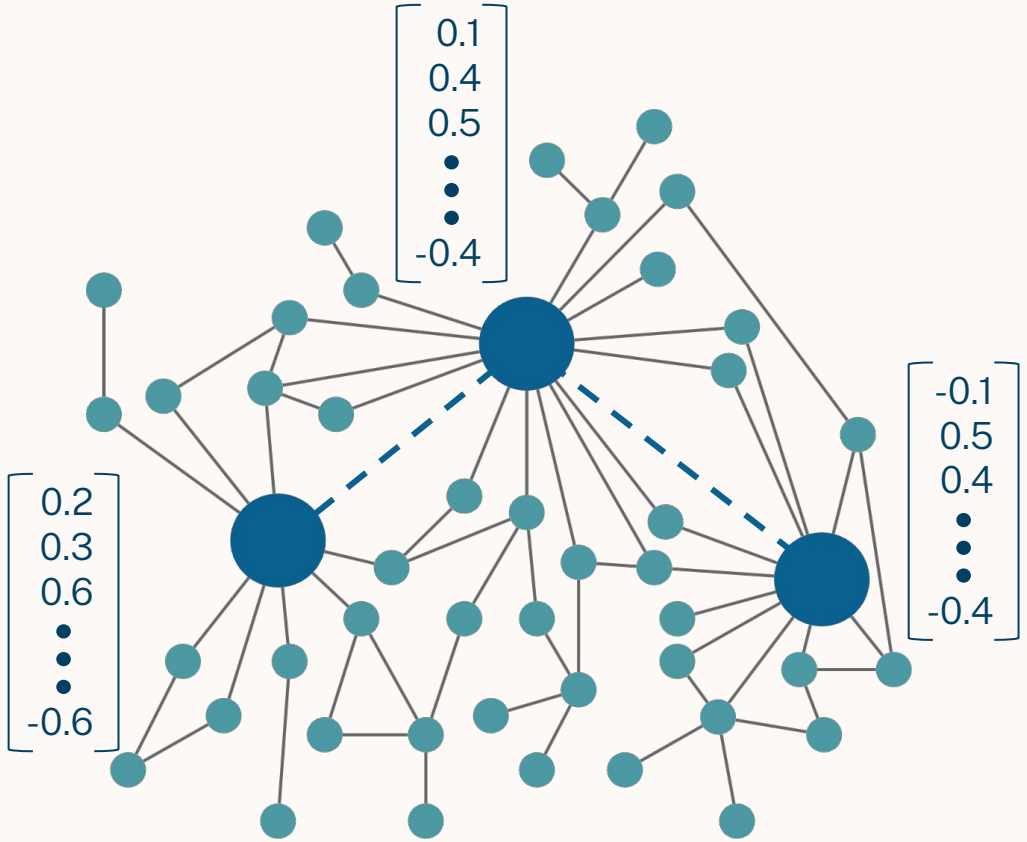






K-Nearest Neighbor (KNN) w/ Node Embeddings

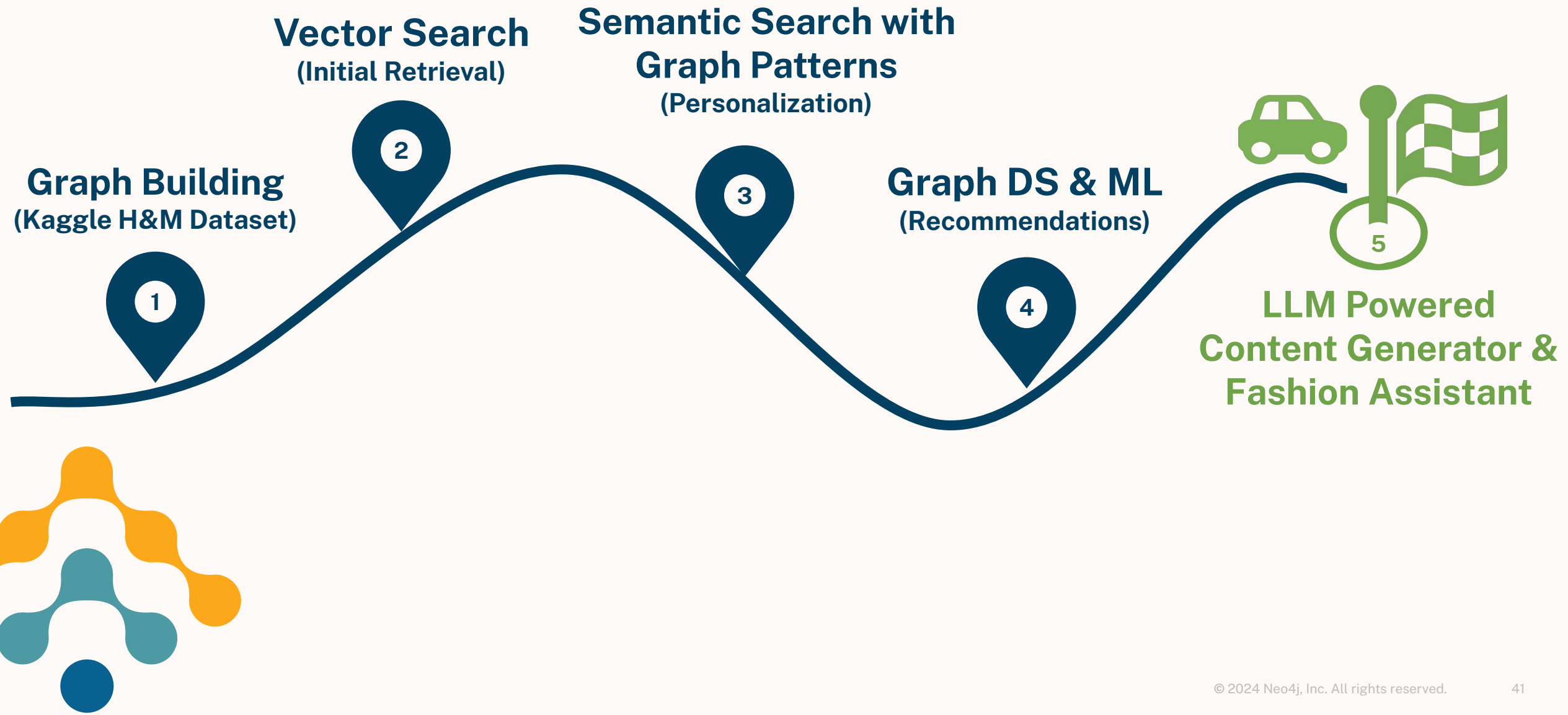
Draw connections between highly interconnected nodes and/or those that have similar roles in the graph



4/ Graph Data Science & ML

Recommendations

Stop before **LLM Powered Content Generator**



Prompting the LLM

You are a personal assistant named Sally for a fashion, home, and beauty company called HRM.Write an engaging email to **{customerName}** , one of your customers, to promote and summarize products relevant for them given:

- The current season / time of year: **{timeOfYear}**

- Recent searches/interests: **{customerInterests}**

Please only mention the products listed below.Do not come up with or add any new products to the list.Each product comes with an https `url` field. Make sure to provide that https url with descriptive name text in markdown for each product.

RelevantProducts:

These are products from the HRM store the customer may be interested in based

on their recent searches/interests: **{customerInterests}**

{searchProds}

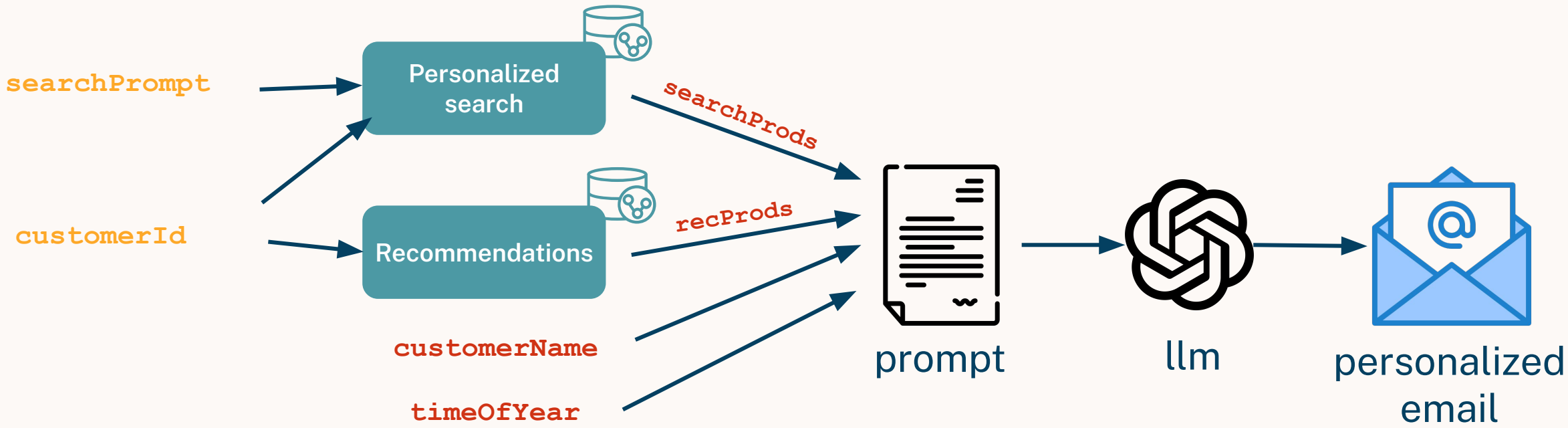
Customer May Also Be Interested In the Following

The below candidates are recommended based on the shared purchase patterns of other customers in the HRM database.

Select the best 4 to 5 product subset from the context that best match the time of year: **{timeOfYear}** and to pair with the RelevantProducts above. For example, even if scarfs are listed here, they may not be appropriate for a summer time of year so best not to include those.

{recProds}

LangChain Chain



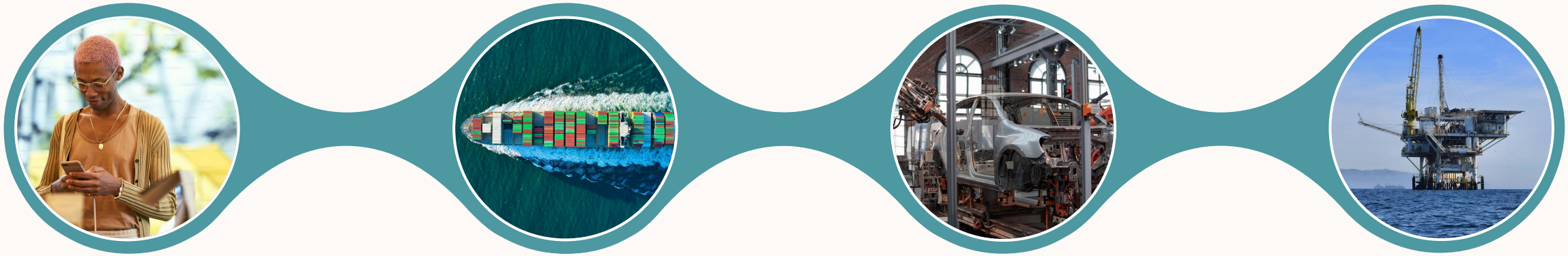

```
{
 searchProds:(searchPrompt , customer_id)| personalizedSearch
 recProds:customer_id | recommendations
 customerName:
 timeOfYear:
}
```

prompt | llm | outputParser

4/ LLM Powered Fashion Assistant

LLM For Generating Grounded Content

Real-World GenAI Breakthroughs Powered by Neo4j



Leveraging AI for customized content at scale

Integrating AI with knowledge graphs for smarter supply chain management

Merging structured and unstructured data for efficient operations

Converting vast amounts of unstructured data into actionable knowledge

1,700+ Organizations Use Neo4j

Banking & Financial Services



Technology


Telecommunications

Energy

E-Commerce

Health & Life Sciences

Thank you!