GraphRAG

GraphRAG

Knowledge graph
- facts, relationships structured representation
- for a domain

```
            manufactured          Car          has
            by                                         Engine   Description=…

        Engine
                   Description=…
```

gollnickdata.de
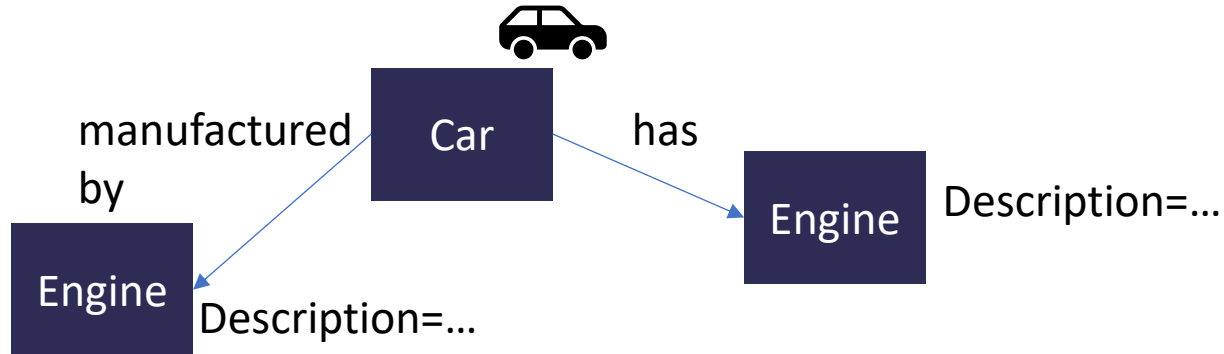
GraphRAG

Knowledge graph has
- nodes and edges

| Property/ Attribute e.g. age, | **Node/ Entity** | **Edge** | **Node** | Property |

Car — has → Engine   Description=…

Property/
Attribute
e.g. age,

Node/
Entity

Edge

Node

Property

gollnickdata.de
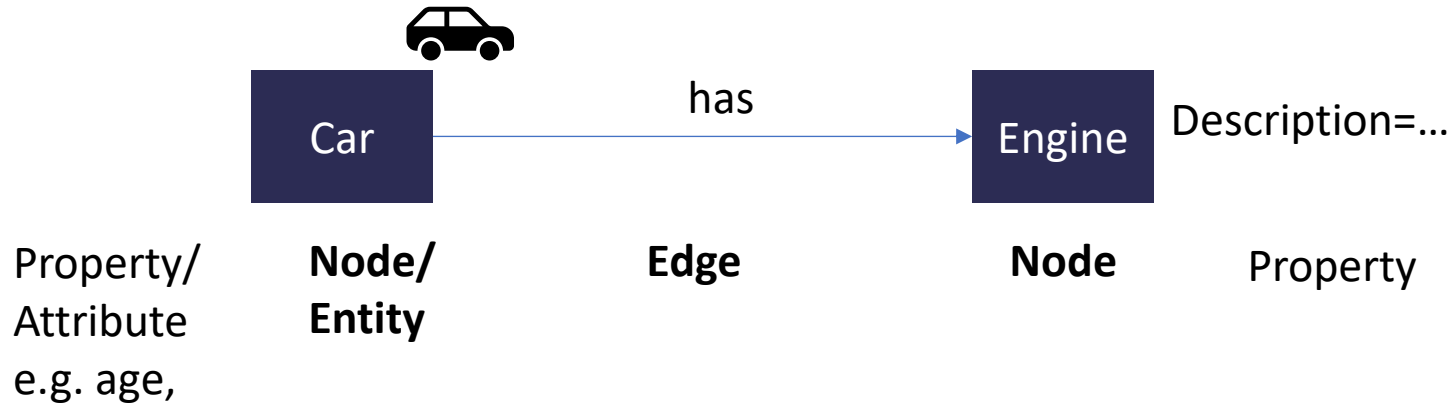
GraphRAG

From Traditional Search to Semantic Intelligence

- traditional search
  - relies on keywords
  - often fails to capture nuances and complex relationships

- knowledge search
  - leverages semantic understanding
  - enables rich information retrieval based on entity relationships

GraphRAG

Retrieval	Reasoning	Generation

GraphRAG retrieves relevant information

performs reasoning over retrieved information, leveraging graph's structure

generates output based on LLM

gollnickdata.de

GraphRAG

Healthcare

- personalized patient care and disease diagnosis

Finance

- fraud detection
- risk assessment
- customer profiling

E-commerce

- personalized recommendations
- intelligent search

Education

- personalized learnings
- knowledge exploration

gollnickdata.de

GraphRAG

- lightweight
- in-memory graphs

- not suitable for large graphs (limited by RAM)

- supports billions of nodes / relationships
- has its own query language
- good visualisation

- requires running db
- Cypher requires learning

gollnickdata.de

GraphRAG

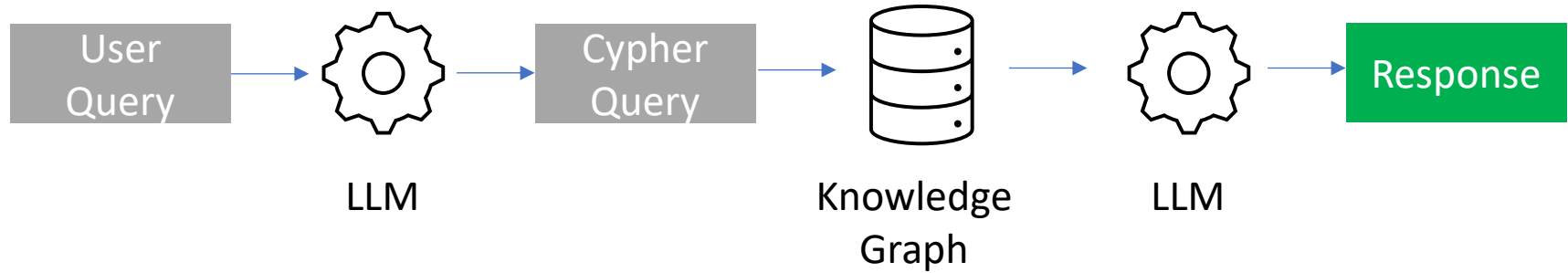| Feature	NetworkX	Neo4j	RDFLib	DGL	igraph	ArangoDB
Type	In-memory graph library	Graph database	RDF-based knowledge graph	Deep learning framework for graphs	Efficient graph analysis library	Multi-model DB (Graph + NoSQL)
Best For	Small to medium-sized graphs	Large-scale persistent knowledge graphs	Semantic web, Linked Data	Graph Neural Networks (GNNs)	Large-scale graph analytics	Multi-model applications
Scalability	Limited (memory-bound)	High (disk-based, cluster support)	High	High	High	High
Ease of Use	Easy (Pythonic API)	Medium (Cypher query language)	Medium (RDF concepts needed)	Medium (GNN knowledge required)	Medium (C-style API)	Medium (Arango Query Language - AQL)
Query Language	Python functions	Cypher	SPARQL	Python API	Python API	AQL (SQL-like)

gollnickdata.de

GraphRAG

Frameworks - Overview

| Feature	NetworkX	Neo4j	RDFLib	DGL	igraph	ArangoDB
Storage	In-memory only	Persistent (disk)	Persistent (RDF triple store)	In-memory (GPU-accelerated)	In-memory & disk	Persistent (document & graph)
Performance	Good for small graphs	Excellent for large graphs	Optimized for RDF queries	GPU-accelerated	Highly optimized	Scalable with indexing
Visualization	Basic (Matplotlib)	Built-in Neo4j Bloom	Limited	No built-in support	Basic (Matplotlib)	Built-in Arango UI
Use Cases	Social networks, dependency graphs	Knowledge graphs, fraud detection	Semantic web, ontology management	GNN research, AI	Large-scale social network analysis	Multi-model use cases

gollnickdata.de

GraphRAG

Knowledge Retrieval: Prompt to Query
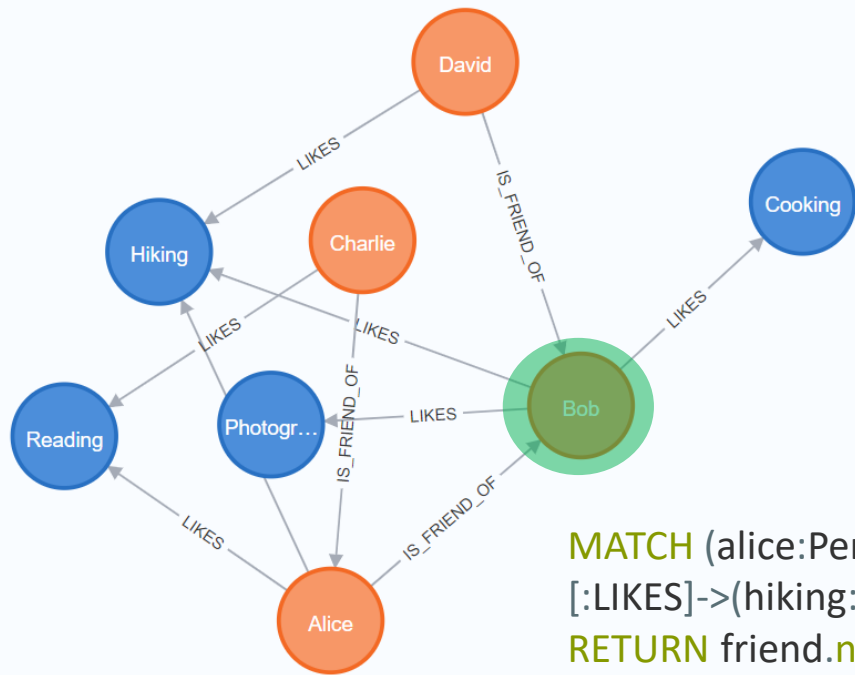
Who are Alice's friends that also like hiking?

without knowledge graph:
- Fetch all persons that are friends of Alice
- Fetch all persons who like hiking
- Filter for overlap

With knowledge graph:
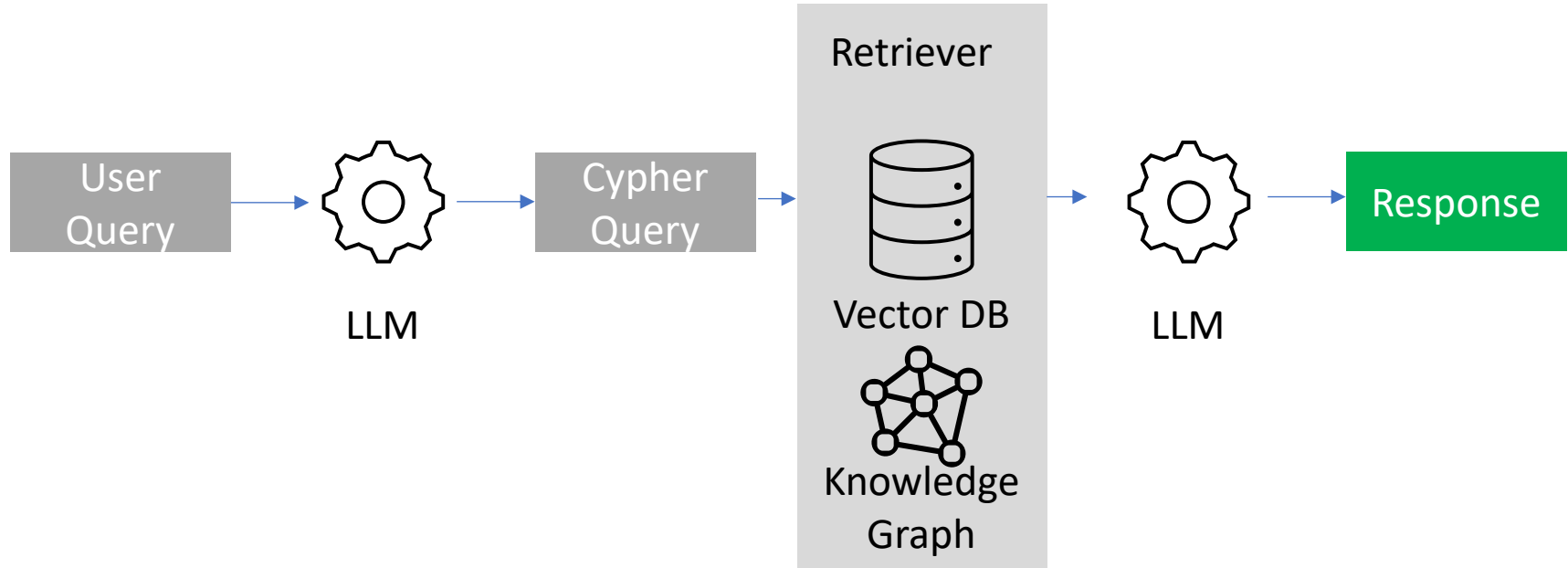- Directly query for the combination

```
MATCH (alice:Person {name: "Alice"})-[:IS_FRIEND_OF]-(friend:Person)-
[:LIKES]->(hiking:Interest {name: "Hiking"})
RETURN friend.name;
```

gollnickdata.de