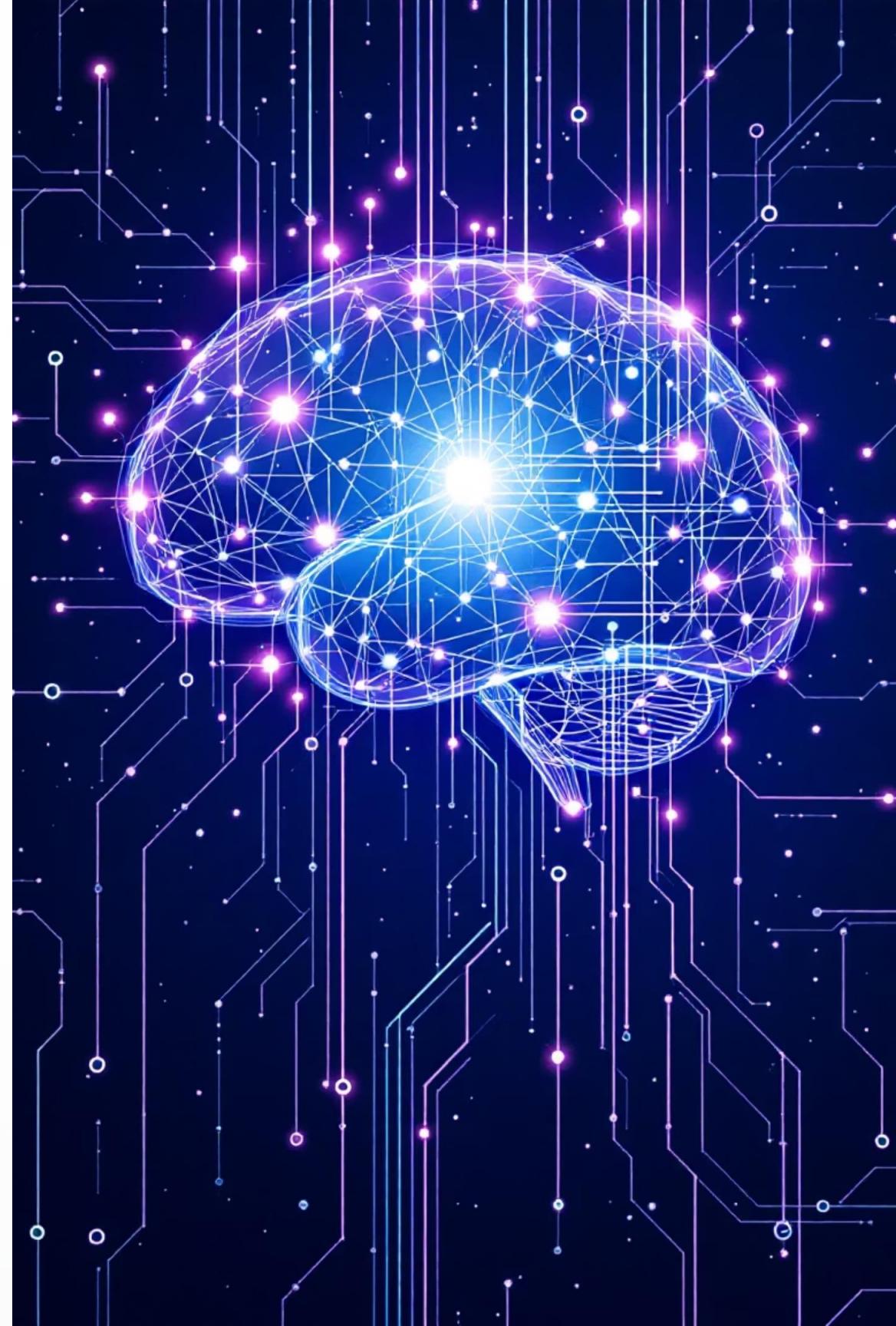


Query Transformation in RAG

(Multi-Hop Retrieval & Q2Q Query Rewriting)



Why Query Transformation?

Complex Queries

LLMs struggle with queries requiring **multiple reasoning steps** or information spread across **different chunks**.

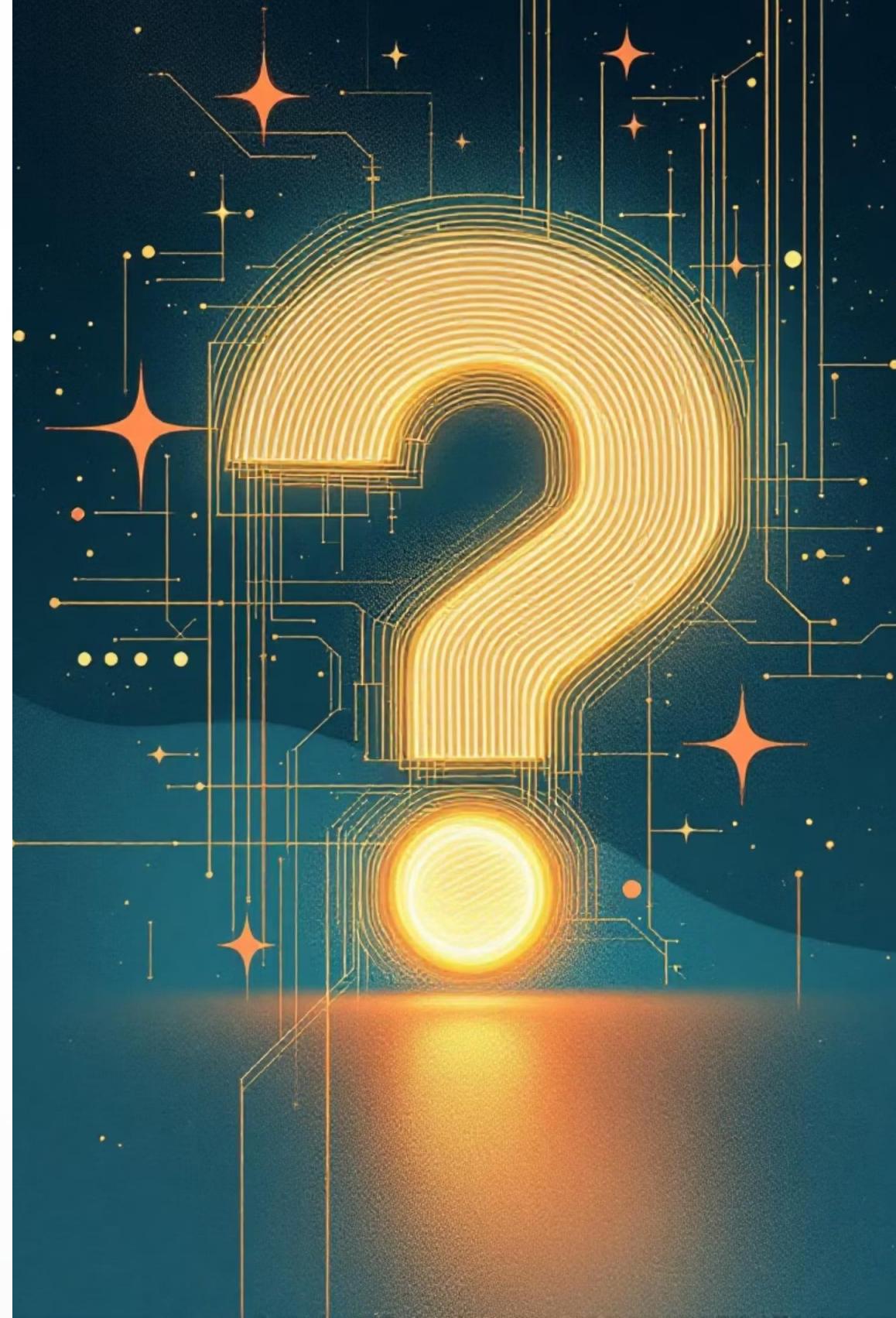
Ambiguity & Incompleteness

Real-world questions are often **ambiguous or incomplete**, leading to poor retrieval.

Missing Context

Direct retrieval methods frequently **miss key context** necessary for a relevant answer.

Query Transformation strategically rewrites queries into more effective forms, addressing these challenges head-on.





Types of Query Transformation

1

Q2Q (Query → Better Query)

Rewording the original query to enhance retrieval quality and relevance.

2

Sub-Query Decomposition

Breaking down a single complex question into multiple, smaller, manageable questions (Multi-Hop).

3

Query Expansion

Adding relevant keywords, synonyms, and entities to broaden the search scope.

4

Query Disambiguation

Clarifying vague or incomplete queries to ensure precise understanding and retrieval.

Q2Q (Query-to-Query Rewriting)

The Large Language Model (LLM) takes the initial user input and intelligently rewrites it into a **retrieval-friendly** query. This ensures the query aligns better with the underlying document structure and language.

Example

User Query: "How is the founder of Microsoft connected to philanthropy?"

LLM Rewritten Query: "Bill Gates philanthropy contributions and charitable initiatives"



- ✓ More precise
- ✓ Better matching to document language
- ✓ Higher retrieval accuracy



Multi-Hop Query Decomposition

When a single question demands **multiple stages of reasoning** or information from disparate sources, we systematically break it down into several discrete, answerable sub-queries.

Example Question:

"How has NASA's collaboration with SpaceX impacted lunar missions?"



Sub-Query 1

What collaborations exist between NASA and SpaceX?

Sub-Query 2

What is the current status of NASA's lunar missions?

Sub-Query 3

How did those specific collaborations influence the progress of NASA's lunar missions?

The LLM retrieves context for each sub-query, then intelligently merges these findings to form a comprehensive, coherent final answer.

How Multi-Hop Retrieval Works



1. Receive Question

The initial complex user query is processed.



2. Decompose Queries

The main question is broken down into smaller, actionable sub-queries.



3. Retrieve Context

Relevant information is fetched for each individual sub-query.



4. Aggregate Evidence

All retrieved contexts are synthesized and combined.



5. Generate Answer

A final, well-reasoned answer is formulated from the aggregated evidence.

✓ Enhanced recall ✓ Higher factual accuracy ✓ Strengthened LLM reasoning capabilities

Examples of When Multi-Hop Helps

Historical Timelines	Information is often spread across multiple disparate pages or documents.	Enables step-by-step reasoning by retrieving events chronologically or thematically.
Cause & Effect	Requires piecing together causality from various sources, which is difficult for direct retrieval.	Sub-queries clarify direct and indirect links between events, leading to comprehensive understanding.
Relationships	Entities or concepts are rarely described in the same document chunk.	Allows separate retrieval for each entity, then linking them contextually.
Comparisons	Documents seldom contain direct comparative analyses of multiple subjects.	Facilitates two or more separate retrievals for each item, then comparison.

Query Transformation Patterns

Clarifying Missing Info

Original: "Who is the coach of the team?"

Transformed: "Coach of *Indian Men's Cricket Team* (current)."

Entity Expansion

Adding full names, crucial dates, and specific locations to enhance query precision.

Context Expansion

The LLM infers and adds relevant, implicit keywords to enrich the query.

Structure Conversion

Transforming complex, convoluted queries into simple, stepwise formulations.



Real World Examples

Example 1 (Political)

Query: "Why was the 42nd Amendment important in India?"

Sub-Queries:

- What specific changes did the 42nd Amendment introduce?
- What were the underlying reasons those changes were deemed influential?
- What long-term impact did it have on India's constitutional structure and governance?

These examples are excellent for classroom demonstrations and practical application exercises.

Example 2 (Historical)

Query: "What caused the fall of the Roman Empire?"

Sub-Queries:

- Investigate the primary economic factors contributing to the fall.
- Analyze the key political factors and leadership crises.
- Examine the significant military factors and external pressures.



Evaluation Framework

Students can utilize this framework to rate each retrieval method and observe the tangible improvements offered by query transformations.

Retrieval Relevance	Low	Medium	High
Answer Accuracy	Low	Medium	High
Hallucination Risk	High	Medium	Low
Evidence Coverage	Low	Medium	High

This hands-on evaluation allows them to empirically observe the significant improvement in RAG performance after applying these transformation techniques.