

Motivation

- LLMs are good at fluent language, but they can struggle with factual and relationship-heavy questions.
- LLMs can do better on these questions when they are connected to an outside source of facts. This is called Retrieval-Augmented Generation, or RAG.
- A standard way to connect LLMs to outside tools and data is the Model Context Protocol (MCP).
- One useful source for these questions is a knowledge graph, which stores facts as connected entities and relationships.
- Knowledge graphs are queried using SPARQL. This project studies whether connecting an LLM to a SPARQL knowledge graph can improve answers.

Research Question

Can connecting an LLM to a SPARQL knowledge graph improve contextual understanding and answer accuracy?

Code and Materials



Research Steps

- Develop the software needed to connect an LLM to a SPARQL knowledge graph using MCP
- Decide which MCP tools should be exposed to the LLM
- Improve the prompt so the LLM uses the MCP tools effectively
- Benchmark the system to test whether graph access improves answer quality

Method/System Overview

- User asks a question
- MCP connects the LLM to graph tools and resources
- The prompt gives the model schema-aware context and safer query generation
- Query executes over RDF knowledge graph
- Retrieved triples/facts are returned as structured context
- LLM generates final answer using graph-grounded evidence
- Outputs are evaluated on reasoning, verifiability, and unsupported claims

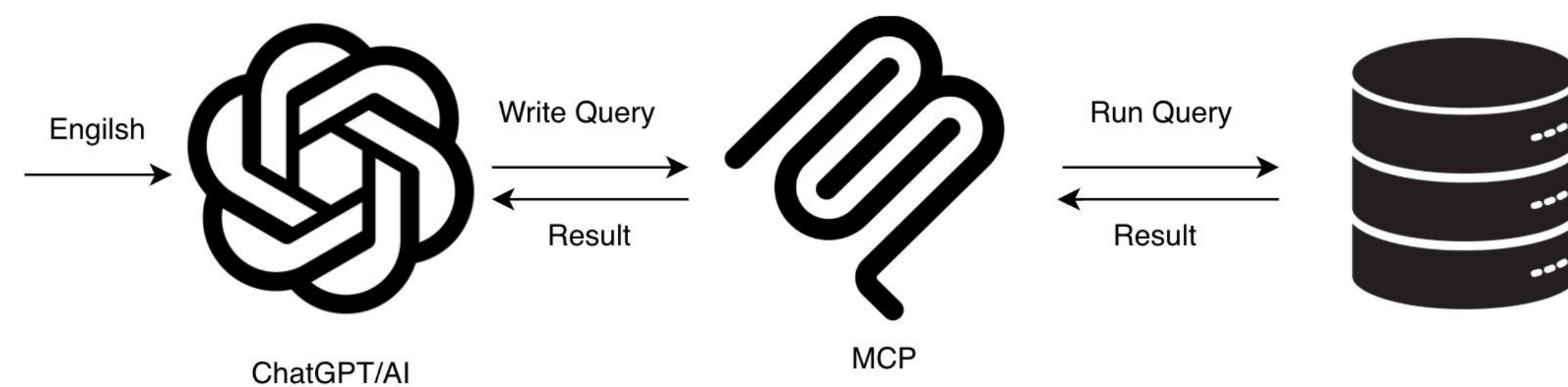


Figure 1. MCP-based question answering pipeline

LLM Responses: Claude

Without

Which reactions interconvert pyruvate and lactate?

The interconversion of pyruvate and lactate is catalyzed by **lactate dehydrogenase (LDH)**, which performs a reversible reaction:

Pyruvate + NADH + H⁺ ⇌ Lactate + NAD⁺

- Pyruvate → Lactate** (reduction): Under anaerobic conditions (e.g., intense exercise), LDH reduces pyruvate to L-lactate, oxidizing NADH to NAD⁺ in the process. This regenerates NAD⁺ to keep glycolysis running.
- Lactate → Pyruvate** (oxidation): In tissues like the liver and heart, LDH runs in reverse, oxidizing lactate back to pyruvate (reducing NAD⁺ to NADH). The pyruvate can then enter gluconeogenesis or the TCA cycle.

This reversible reaction is central to the **Cori cycle**,

With MCP

Reactions Interconverting Pyruvate and Lactate

All reactions are reversible in principle; the arrow reflects the Rhea canonical direction

L-Lactate (S-configuration) ↔ Pyruvate

Rhea ID	Equation	Enzyme (EC)
RHEA:23444	(S)-lactate + NAD ⁺ = pyruvate + NADH + H ⁺	L-lactate dehydrogenase (EC 1.1.1.27)
RHEA:19909	(S)-lactate + 2 Fe(III)-[cyt c] = 2 Fe(II)-[cyt c] + pyruvate + 2 H ⁺	L-lactate—cytochrome c oxidoreductase (EC 1.1.2.3)
RHEA:45816	(S)-lactate + A = pyruvate + AH ₂	L-lactate dehydrogenase (general acceptor, EC 1.1.99.-)
RHEA:55868	(S)-lactate + O ₂ = pyruvate + H ₂ O ₂	L-lactate oxidase

D-Lactate (R-configuration) ↔ Pyruvate

Rhea ID	Equation	Enzyme (EC)
RHEA:16369	(R)-lactate + NAD ⁺ = pyruvate + ↓ H ⁺	D-lactate dehydrogenase (EC 1.1.1.28)

LLM Responses: ChatGPT

Without	With MCP
Which reactions produce pyruvate from phosphoenolpyruvate? Thought for a couple of seconds The reaction is: $\text{Phosphoenolpyruvate} + \text{ADP} + \text{H}^+ \rightarrow \text{pyruvate} + \text{ATP}$ It is catalyzed by pyruvate kinase. A couple of key points: - It is the last step of glycolysis. - It produces pyruvate by transferring the phosphate from phosphoenolpyruvate (PEP) to ADP. - This is an example of substrate-level phosphorylation. In many descriptions, the immediate product is formed via enolpyruvate, which then rapidly tautomerizes to pyruvate.	phosphoenolpyruvate into pyruvate. Thought for 11s Called tool Approved Rhea reactions I found that produce pyruvate from phosphoenolpyruvate are: - RHEA:18381 — dihydroxyacetone + phosphoenolpyruvate = dihydroxyacetone phosphate + pyruvate - RHEA:19997 — phosphoenolpyruvate + H₂O = pyruvate + phosphate - RHEA:23880 — L-histidyl-[protein] + phosphoenolpyruvate = N(pros)-phospho-L-histidyl-[protein] + pyruvate - RHEA:34943 — hydroquinone 0-beta-D-glucopyranoside(out) + phosphoenolpyruvate(in) = arbutin 6-phosphate(in) + pyruvate(in) A few hits in the raw search mentioned both compounds but in the reverse direction, so I left those out.

Results/Takeaways

Table 2. Summary of evaluation outcomes

Metric	Effect
Multi-hop reasoning	Improved
Verifiability	Improved
Unsupported claims	Reduced
Latency	Slight increase

- Knowledge graphs are especially useful for relation-heavy questions
- MCP gives LLMs structured access to graph tools and RDF resources
- Graph-grounded retrieval improves verifiability and multi-hop reasoning
- This comes with a small latency tradeoff but better factual reliability

Acknowledgments

I want to thank Dr. Ashraf Aboulnaga, Humera Sabir and Fatima Farooq for their contributions to this research.