

The background of the slide is a complex network graph pattern. It consists of numerous small blue and black dots (nodes) connected by thin, light blue lines (edges), creating a dense, interconnected web that fills the entire slide area.

Support for Grounding Large Language Models with Knowledge Graph in Graphene

- *Sangamithra Panneer Selvam*

Motivation for Grounding



Ensuring the quality, accuracy, and relevance of the generated output is crucial.



Use-case specific information is not inherently part of their training data.



Time and cost of training.

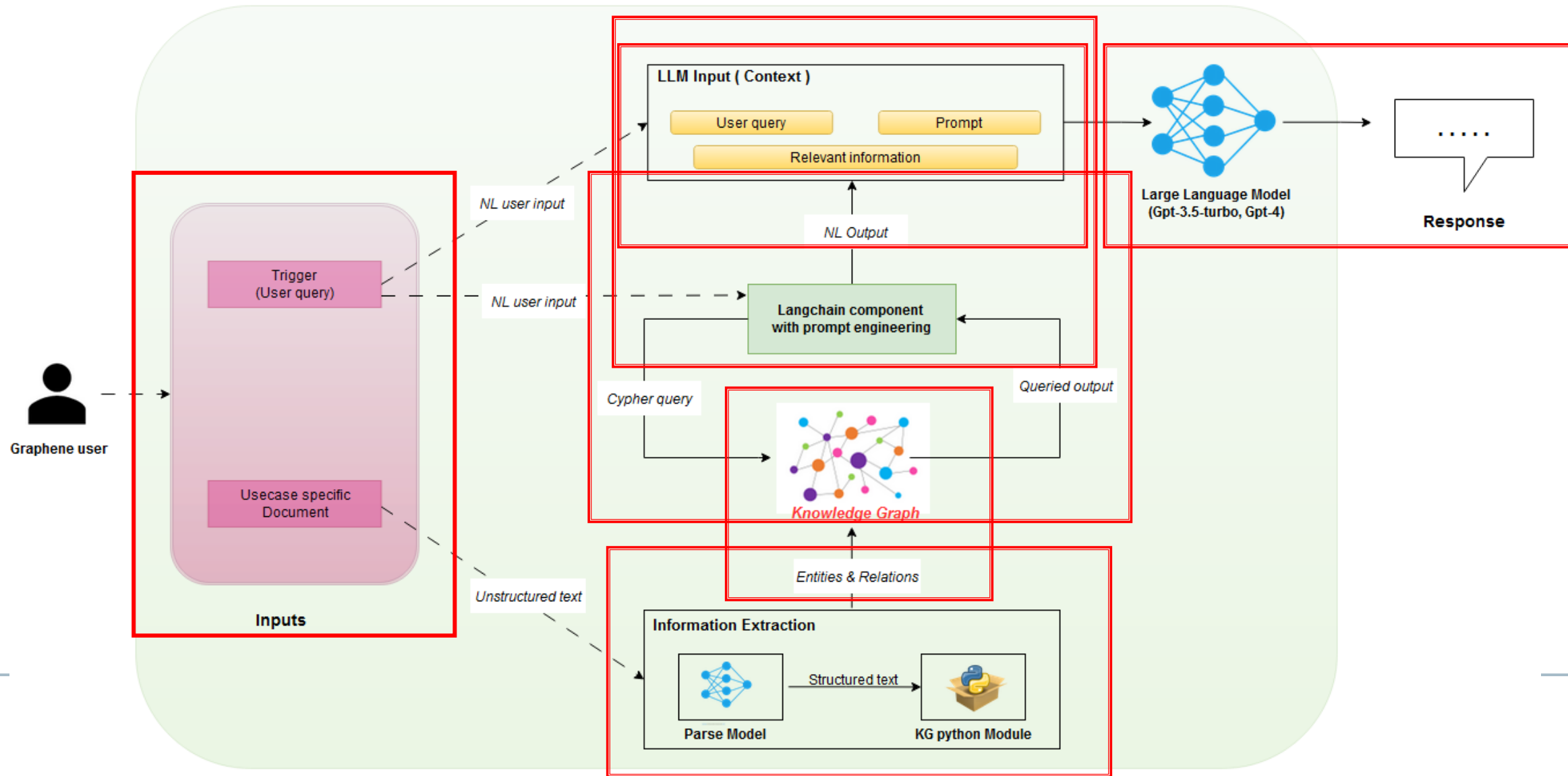


Need for Intelligent capabilities.

LLM

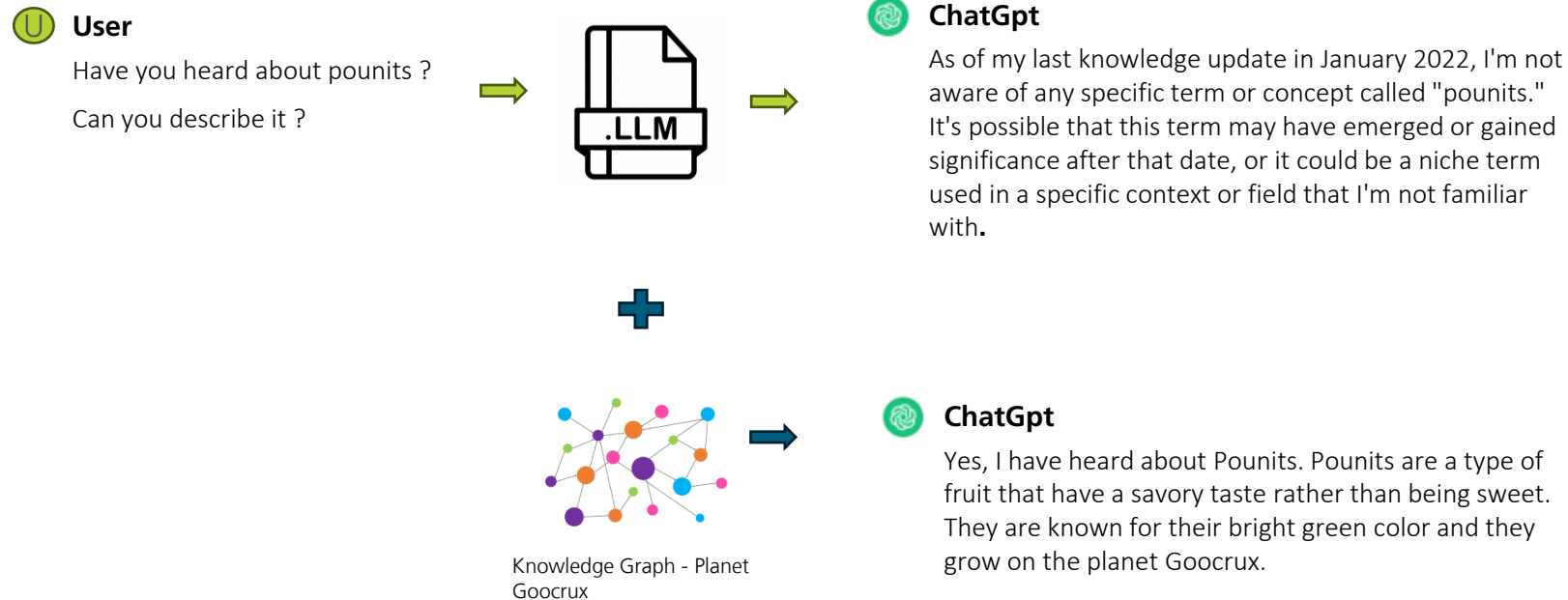


Architecture



Result 1 – Planet Gooocrux

Augment LLM with un-trained information



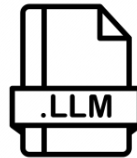
Result 2 - Frozen

Improve Results



User

Tell me about the movie Frozen ?



ChatGpt

Frozen is a 2013 animated musical fantasy film produced by Walt Disney Animation Studios. It is loosely based on Hans Christian Andersen's fairy tale "The Snow Queen." The film tells the story of two princesses, Elsa and Anna, who live in the kingdom of Arendelle. Elsa possesses magical ice powers, which she struggles to control...



Knowledge Graph -Frozen

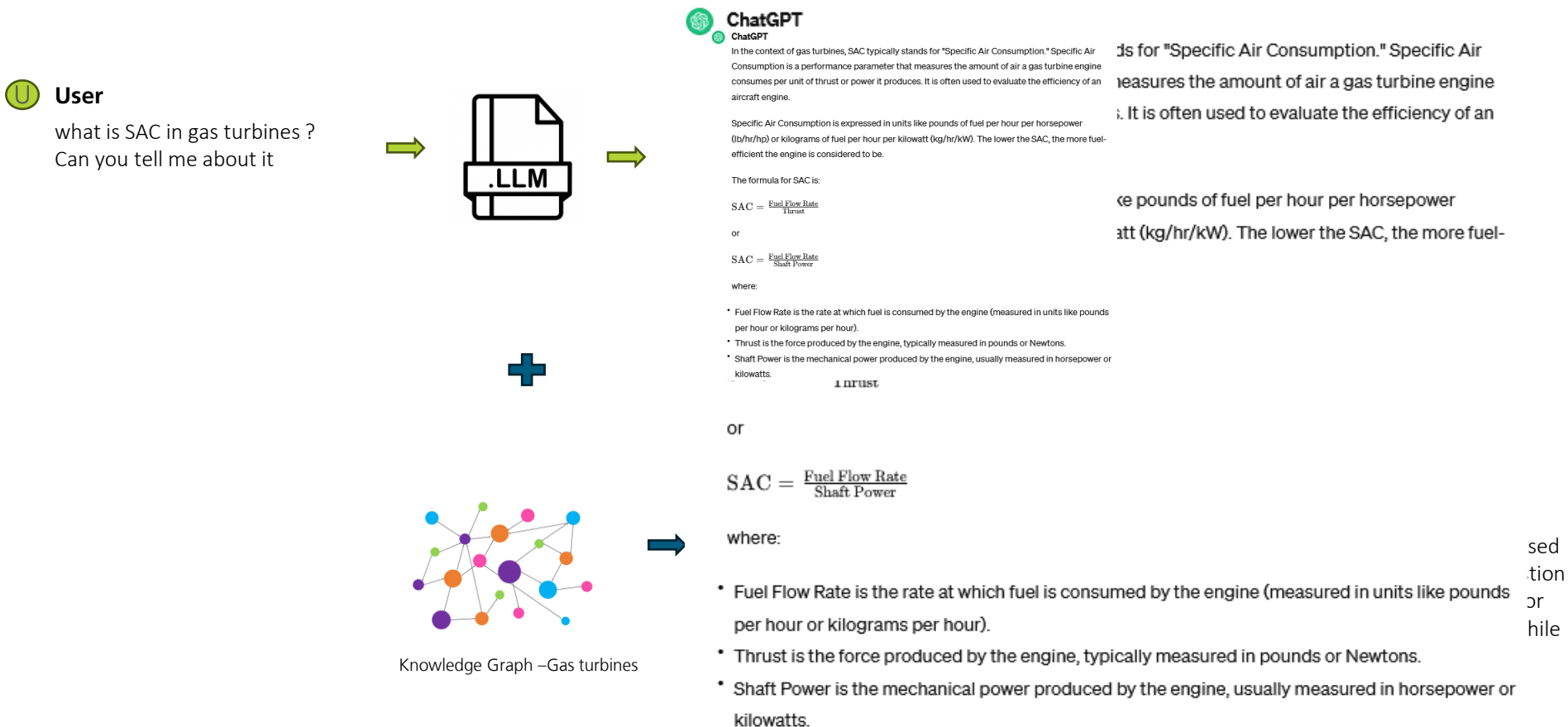


ChatGpt

Frozen is a highly popular animated film produced by Walt Disney Animation Studios and distributed by Walt Disney Pictures. It was released in 2013 and quickly became a cultural phenomenon. The movie premiered at the El Capitan Theatre on November 19, 2013, and had its general theatrical release on November 27, 2013. The film received widespread acclaim from both critics and audiences. It won the Best Animated Feature at the 86th Academy Awards and also received the Best Original Song award for its hit song "Let It Go."...

Result 3 – SAC in Gas turbines

Reduce Hallucinations



Grounding LLM Pipeline - Live Demo

DEV AI-Builder

KI-NRW Playground

dev01.ki-lab.nrw:8080/dashboard?selected_deployment_id=gllm-18310609b1f640b99b5f42e98a809967

KiNRW

gllm

73616e67616d69746872612e70616e6e6565722e73656c76616d40696169732e667261756e686f6665722e6465

Deployment: gllm

Status: Ready

Run

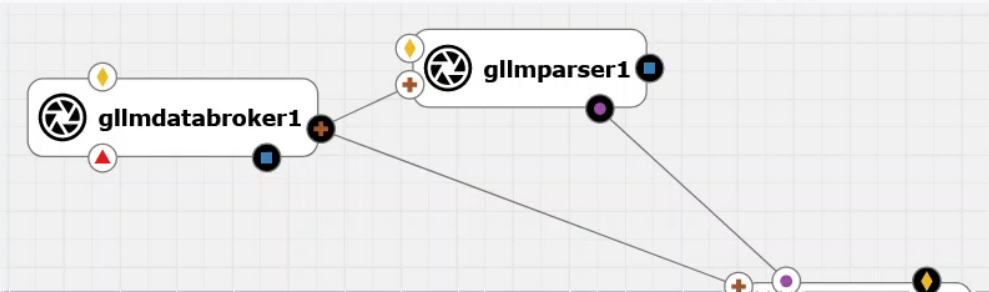
Reset

Delete

Logs

Status Check	Nodename	Status Details	Logs	WebUI/Folder
✓	gllmdatabroker1	view	view	🔗
✓	gllmmodel1	view	view	🔗
✓	gllmparser1	view	view	🔗
✓	orchestrator	view	view	🔗

Solution description

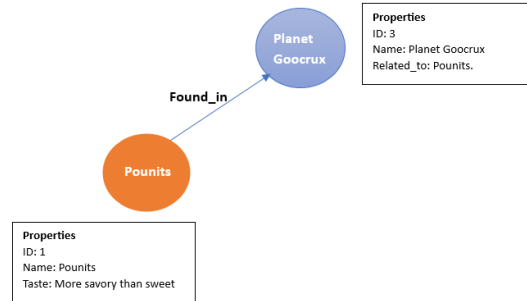


```
graph LR; gllmdatabroker1 --> gllmparser1; gllmparser1 --> gllmmodel1; gllmmodel1 --> orchestrator
```

Grounding is the process of using large language models (LLMs) with information that is use-case specific, relevant, and not available as part of the LLM's trained knowledge. It is crucial for ensuring the quality, accuracy, and relevance of the generated output. While LLMs come with a vast amount of knowledge already, this knowledge is limited and not tailored to specific use-cases. To obtain accurate and relevant output, we must provide LLMs with the necessary information. In other words, we need to "ground" the models in the context of our specific use-case.

Further works on GroundingLLM pipeline

- Modification of Kg



- Python script development to eliminate Langchain
- Comparative study on GPT models with Langchain and python script

Identified Areas of Improvement

✓ Scenario 1:

1. question: 'How does Pounits taste?'
2. usecase document: 'Pounits taste more savory than sweet'
3. KG created : Pounits -> are -> more savroy than sweet
4. cypher generated: `MATCH (p:node {name: 'pounits'})-[:taste]->(t) RETURN t`
5. LLM ouput: Cannot be found

✓ Scenario 2:

1. question: 'How does Pounits taste?'
2. usecase document: 'Pounits are more savory than sweet'
3. KG created : Pounits -> are -> more savroy than sweet
4. cypher generated: `MATCH (p:node {name: 'pounits'})-[:taste]->(t) RETURN t`
5. LLM ouput: Cannot be found



Conclusion & Future scope

- Achieved first running POC.
- Reduce hallucinations, gives better results and augment LLM with untrained information for good output.

Future Scope:

- Conversational memory for LLM for interactive pipeline.
- Embeddings and Embedding store.
- Generic database node for graph and vector storage.
- Open source models and Opengpt-X
- Evaluating LLM via Metrics.

Thank you

