

ERA: A LLM-Ontology informed Goal-based Agent for Fake News Detection



Group 6 | Riccardo, Tijmen, Ruben, Luc, Lars, Vedant

Problem

With the emergence of internet people are more prone to fake news than ever. We want to detect and explain why the given piece of information is factually/logically invalid.

Solution

We propose a solution to identify fake news using a goal based agent which queries our trusted ontology and a LLM to arrive at the decision whether the given information is real or fake.

Result

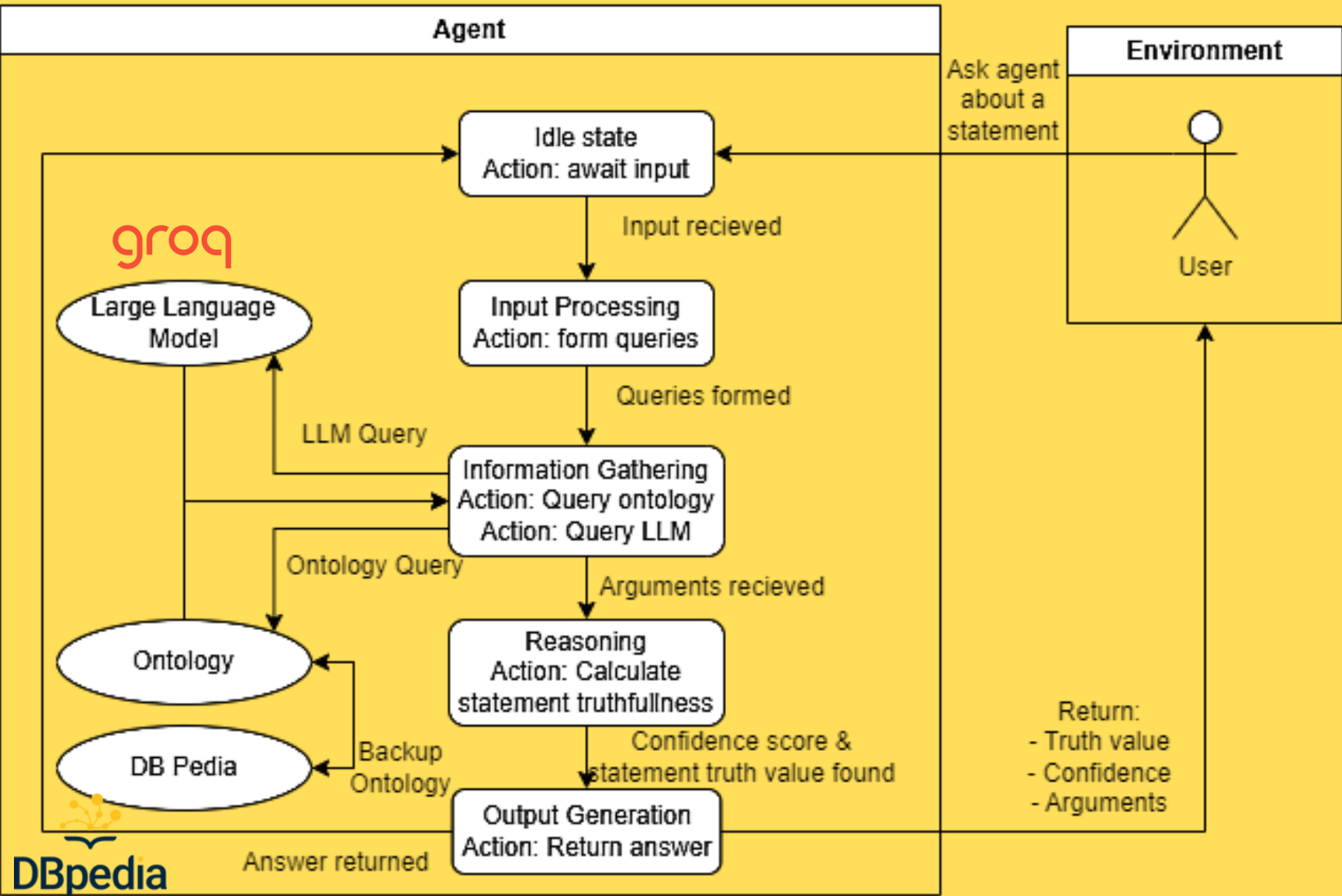
ERA is a goal-based agent designed to evaluate the truthfulness of news by reasoning through various arguments. It provides

- **Verdict on the credibility** of the information
- **Supporting evidence** both for and against the statement, potentially supplemented by **relevant references**.

Scenarios

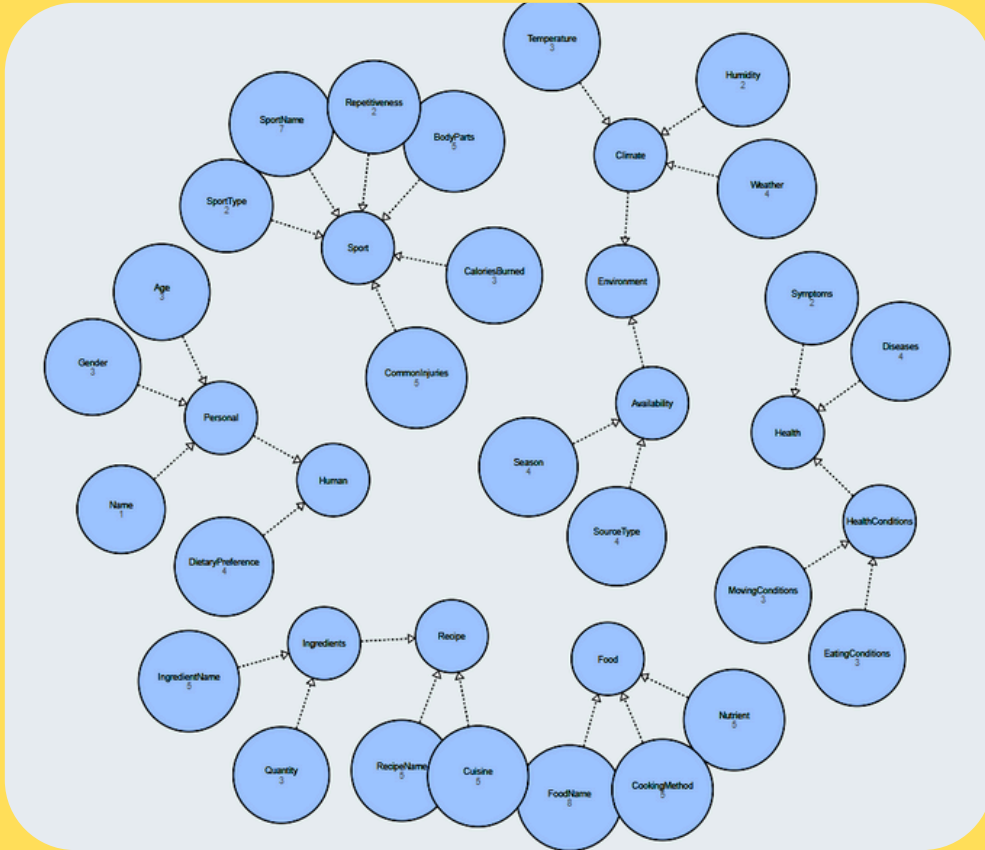
- Running is good for your Health
- Consuming sugary food helps with Diabetes Mellitus

Agent Architecture

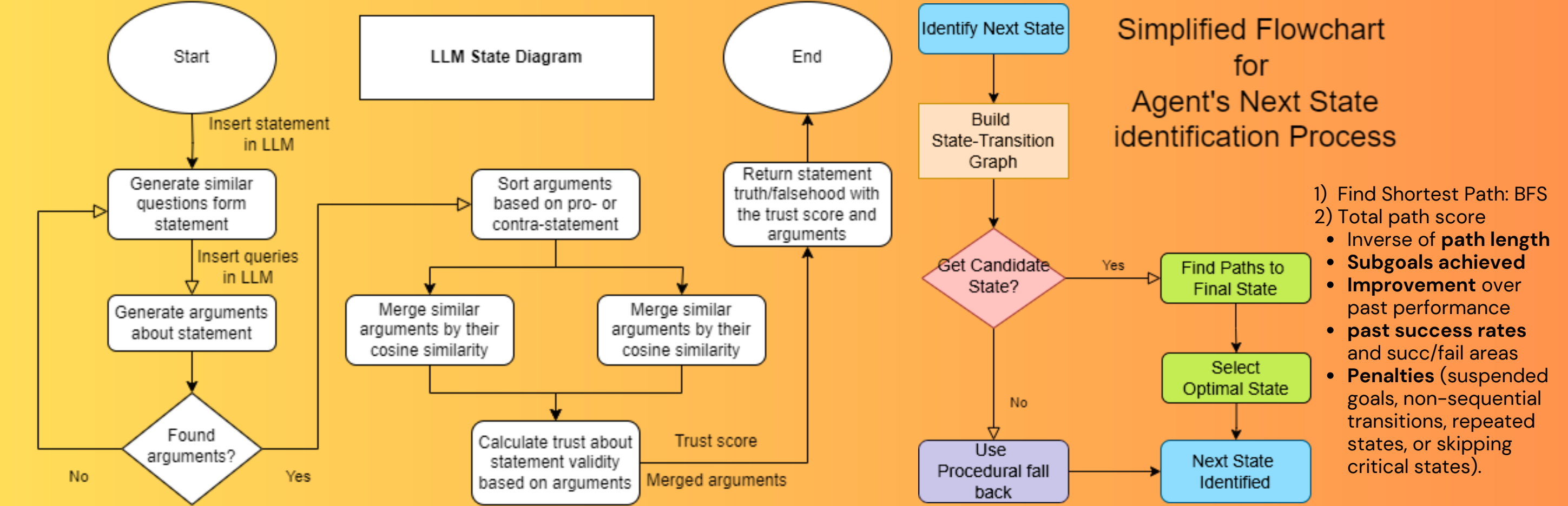


Ontology

- New concepts: Human, Environment
- USE: The Hermit reasoner is queried with the mapped nlp sentence
- Query formulation: retrieve instances belonging to Intersection (all instances for Universal existence), fallback to DBpedia



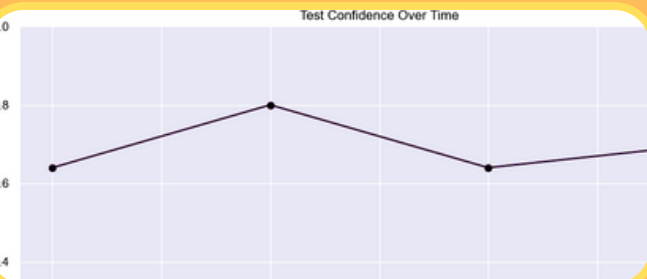
Knowledge Reasoning and Planning capabilities



Performance metrics

ISOT Fake News dataset: (Real-News 21417, Fake-News 23481)

Each news is summarised into news headline (16 train-4 test)



Metric	Value
Final Accuracy	0.75
Average Confidence	0.6949
True Positives	0
False Positives	0
True Negatives	3
False Negatives	1
True Positive Rate (Recall)	0.0
False Positive Rate	0.0
False Negative Rate	1.0
Precision	0
False Discovery Rate	0
Specificity	1.0
Balanced Accuracy	0.5

