Matias Barcelo Treiyer

Dr. Frank Coyle

CS 5393-004

18 November 2024

Final Project Idea

In class, one of our classmates, a black gentleman with an English accent, was telling us about how the performance of AI was a bit underwhelming when it came to translation. I have a buddy who is a medieval history major; the other day he was telling me about how it is possible to find the writings of Cicero in its original Latin. These two experiences made me wonder how well an AI could translate a dead language such as Latin, and whether it could make a knowledge graphs based off its own translations. My idea is to feed a large language model the original Latin of Cicero’s writings, make the model translate it to English, develop RAG and kRAG systems, use the translations with the these developed systems to make knowledge graphs of triples, and to analyze the results.

My source will be a LacusCurtius, a website with a uchicago.edu domain with Roman texts, which I assume is legitimate considering the material, but am not 100% certain is a credible source due to the [authors vague or "dubius" credentials (according to the title page of his own site)](https://penelope.uchicago.edu/Thayer/E/HELP/credentials.html) and the fact I can’t find how it’s related to the University of Chicago besides it being hosted on their domain. I will use their collections of [Cicero’s work in its original Latin](https://penelope.uchicago.edu/Thayer/E/Roman/Texts/Cicero/home.html) for this project.