

Applying Generative AI to Humanities Corpora

Theo Barton, Class of 2026

Purpose

This summer, I had the incredible opportunity to explore the potential of Retrieval Augmented Generation (RAG) systems for humanities research, using the Digital Ricoeur project as the platform to develop the model. Unlike large language models (LLMs), RAG allows for the contextualized integration of specific knowledge bases into the generation process, improving precision, reducing hallucinations, and enabling updates without retraining. In preparation, I studied literature on RAG architecture, vector storage, and humanities-focused AI applications, which informed the choices made during model development. Using Digital Ricoeur, the digitized collection of Paul Ricoeur's works, we created a RAG framework capable of retrieving relevant passages and generating contextualized insights. This project demonstrates the potential of RAG to improve access, interpretation, and engagement with complex philosophical texts, while providing a base for comparing its effectiveness to large LLMs in humanities research.

Corpus Preparation

The initial phase focused on expanding the English-language corpus. This involved digitizing dozens of Ricoeur's primary and secondary works to incorporate into the project database and subsequently incorporated into the English-language corpus into the RAG model.

Retrieval System

A critical part of building the RAG system was testing retrieval methods to maximize the relevance of text selected for the output. Selecting relevant text blocks, or chunks, for a user's query was essential to the RAG model's success. I explored multiple chunking techniques, tested various vector embedding databases, and compared vector distance measurement strategies to ensure the retrieved information held maximum relevance to the inputted question.

Answer Generation

For answer generation, I tested multiple local LLMs to determine which produced the most coherent and accurate responses and finalized our model using Llama. Retrieved chunks were integrated directly into the model's output to ensure grounding in the source material, and I experimented with different prompt structures within the model to refine clarity, depth, and usefulness of the answers.

Results and Broader Significance

For the final product from the summer's work, I created a survey comparing the results of our model to Claude in answering seven questions covering different areas of Ricoeur's corpus. Experts of Ricoeur will grade the generated outputs from the models on accuracy, grounding, interpretive depth, clarity, and usefulness. This ongoing part of the project will deepen understanding of the effectiveness of RAG systems compared to larger scale LLMs when applied to academic research in the humanities. This project was an incredible opportunity to build my critical thinking, problem solving, and technical skills, and I am grateful for those who supported me in this fellowship, especially Fernando Nascimento, Crystal Hall, and Jennifer Snow.

Faculty Mentor: Fernando Nascimento

Funded by the Gibbons Summer Research Fellowship