

Argonne Lab is a government run facility which is under the U.S. Department of Energy. One challenge that Argonne faces is updating and indexing through the vast quantity of government documents that detail policies that Argonne is required to follow. Argonne is interested in using Large Language Models to allow employees to query these documents quickly, but due to privacy issues cannot use proprietary models like Chat GPT-4. Argonne tasked the Data Science clinic team with constructing a graph-based knowledge retrieval model that would be able to retrieve information without relying on proprietary LLM services.

The team worked on creating a pipeline to take an input question and return most relevant document segments. The team used NLP deep learning models to transform and compare the text segments, which then are used to query a pre-constructed knowledge graph that visualizes hidden relationships.

This quarter, the team implemented a query interface that allows users to ask the model real questions (Fig 1). Alternative and scalable methods were explored for each part of the pipeline, including using computer vision to segment documents, using open source LLMs to label relevant segments for testing, and using keywords as an alternate similarity measure (Fig 2).

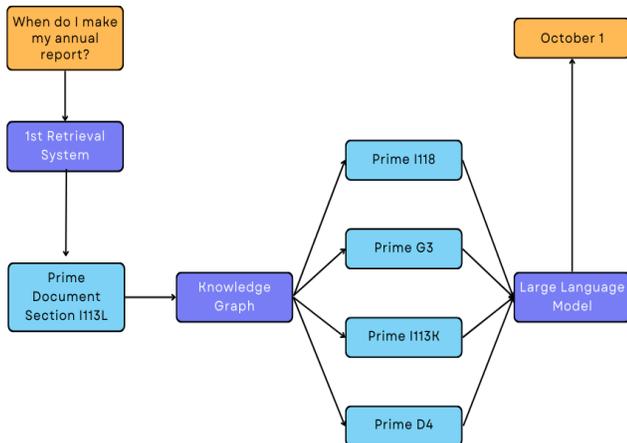


Fig.1 Example Output from Pipeline

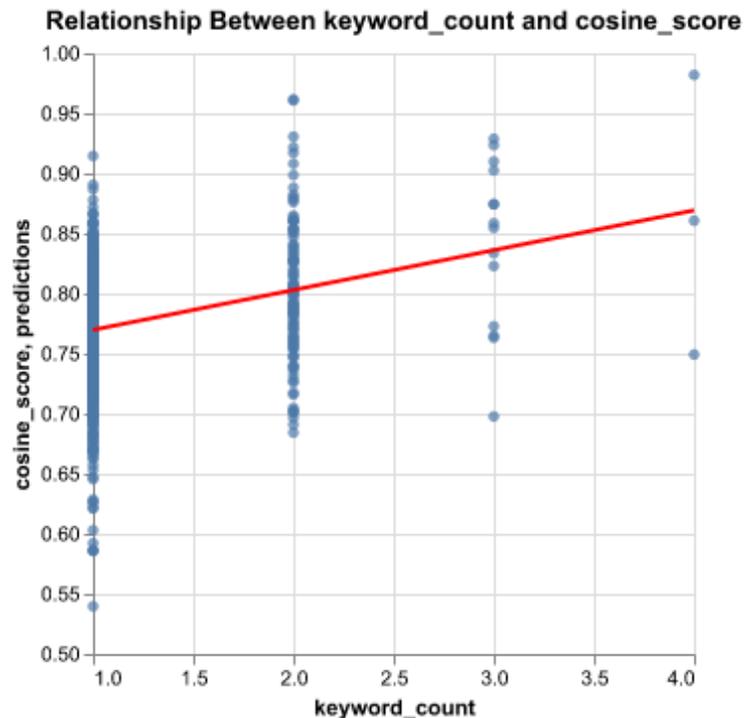


Fig. 2 Keyword Relevance Exploration