

Web-Based Semantic Vocabulary Matcher with LLM Embeddings

Client

Gabe Fierro (gtfierro@mines.edu)

Project

Semantic graphs are a kind of knowledge base used to model large cyber-physical systems like buildings, smart grids and transportation systems. These graphs are increasingly found at the center of IoT orchestration and data analytics platforms where they help software programmatically access data points and other resources. The [Brick ontology](#) is a semantic vocabulary for smart buildings developed by Dr. Fierro. A common task for working with Brick is creating a “mapping” or “translation” between some existing vocabulary and Brick. Dr. Fierro has used [large language model \(LLM\) embeddings](#) perform this task and has incorporated this solution into an open-source tool (be released).

The team’s task will be to create a web interface to this open-source tool that makes it easy for users to upload their vocabulary file and download a “translation” file between the terms in their vocabulary file and the concepts defined by the Brick ontology. The LLM portion of this is largely finished but can definitely be improved. The primary part of the project is developing the web interface so that the whole process can be both *interactive* and *iterative*.

Team Size

3-5 students

Skill Set

The following skills will be used during the project, so some experience with them is recommended (but not wholly necessary):

- experience with 1 or more of the following programming/markup languages:
 - Python, JavaScript, HTML, CSS
- experience developing web frontends (especially if you know Vue.js)
- experience working with LLMs and their APIs, e.g. OpenAI
- experience in designing user experiences on the web

Location

Mines campus

Expected Outcomes

At the end of the project, the team will be expected to release the tool under a permissive open-source license so it can be hosted on a live server and used by researchers.