Raghu Kodanda

RaguKodandaRao@gmail.com

Realtime LLM

Embedding JSON Document with Pathway API Information

Contents

[Preface 2](#_Toc155739496)

[Introduction 3](#_Toc155739497)

[Methodology adopted 3](#_Toc155739498)

[Initial Knowledge Graph 3](#_Toc155739499)

[Initial JSON embedding 4](#_Toc155739500)

[Embedding JSON incrementally 5](#_Toc155739501)

# Preface

In December 2023, IITG (IIT Guwahati) in association with Pathway, provided a three week course on Real time LLM with a hands on Project and helped hundreds of Generative AI enthusiasts to upskill. Further a Discord community was setup, a few online interactions helped the participants towards completing assignments.

As part of my hands on project, I came up with capability to extend the Pathway Framework’s Example Prompt interactions to display a knowledge Graph of the responses. These Knowledge Graphs could be continuously updated based on response text. The project was well received by the reviewing team. Further the reviewers shared feedback to utilize it to build Knowledge Graphs of Pathway’s API references.

This document is my attempt towards incorporating the feedbacks and the subsequent challenges, experiences faced and as a mechanism towards documenting the outcomes.

# Introduction

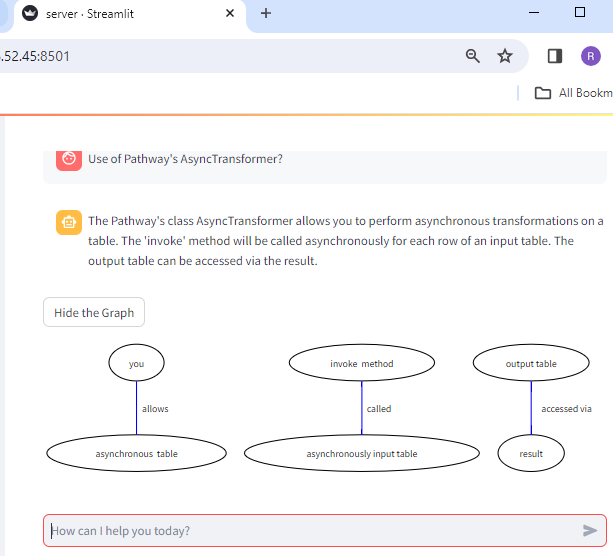
This document covers the methodology adopted to embed JSON documents with Pathway’s API reference information to generate Knowledge Graph from them. It assumes

# Methodology adopted

The Pathway API reference information was first downloaded as a text file for easy referencing, the same was then used to reword phrases for better extraction of Entities & Action words.

For example, occurrences of “pw.AsyncTransformer” was reworded as “Pathway's class AsyncTransformer”. This was done to override references of other AsyncTransformer class references already available in the trained model, for instance that of HuggingFace.

## Initial Knowledge Graph



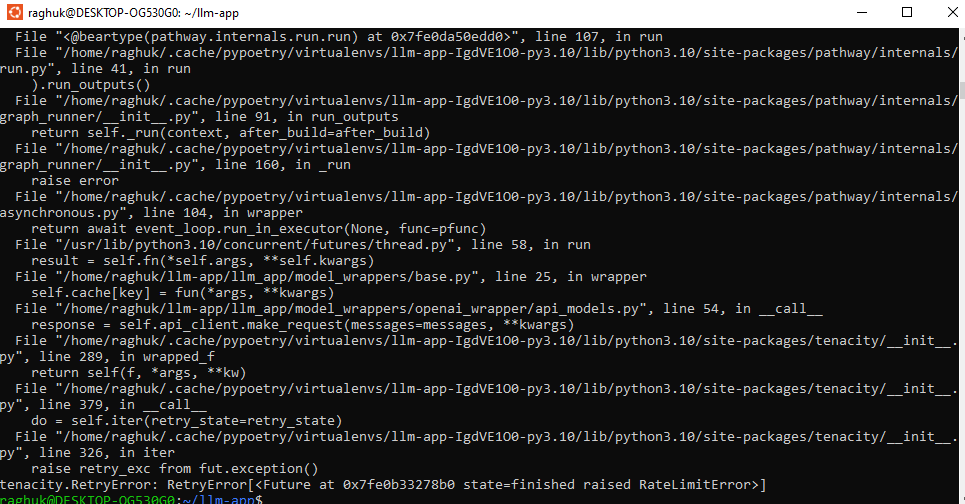
Though various options to convert the Pathway API references using HTML to JSON, Doc to JSON and XML to JSON were explored, each of them had their own limitations and over heads. Also, it needed manual interventions & ultimately the sentences needed manual interventions for better entity & actions extraction.

Hence, each logical context of the Pathway API, example at a class level, was transformed into a JSON manually.

## Initial JSON embedding

*{"doc": "The Pathway's class AsyncTransformer takes an input table and Allows to perform asynchronous transformations on a table. The invoke method will be called asynchronously for each row of an input table, the Output table can be accessed via result. The close method is Called once at the end and is the Proper place for cleanup. The abstract async invoke method is Called for every row of input table. The arguments will correspond to the columns in the input table and Should return dictionary of values matching ‘output schema’. The open method is Called before actual work and is Suitable for one time setup. The Parameters for Open method are capacity sets maximum number of concurrent operations ‘retry strategy’ defines how failures will be handled. The Open method Returns self with result as a Table!"}*

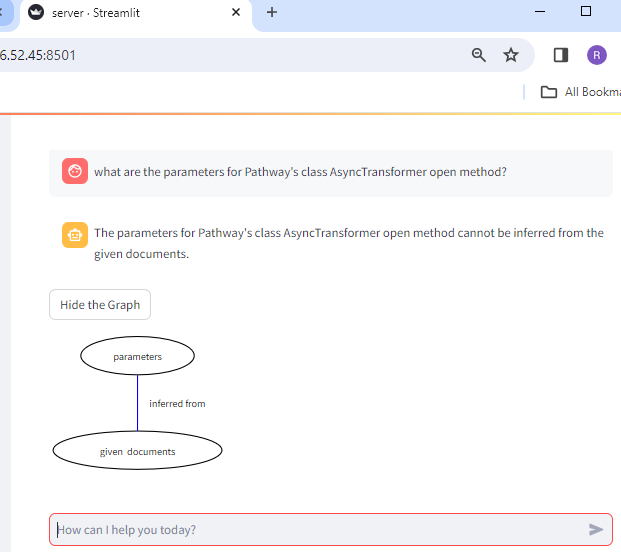
The large chunk of sentences lead to “Rate Limit Errors”



# Embedding JSON incrementally

Given the challenge of “Rate Limit Errors” a smaller JSON document was reworked and the services were restarted.

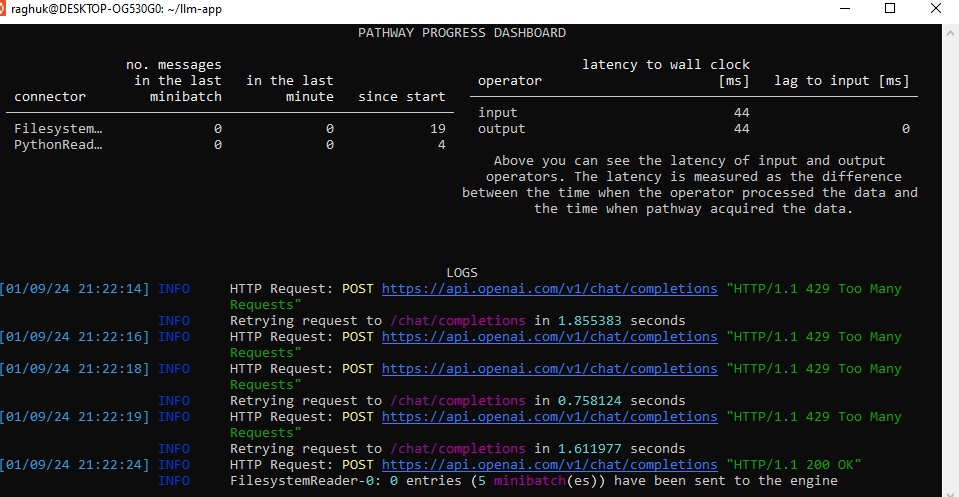
The below screenshot shows a fresh attempt to embed JSON & an interaction prior to the placing the JSON in the data path

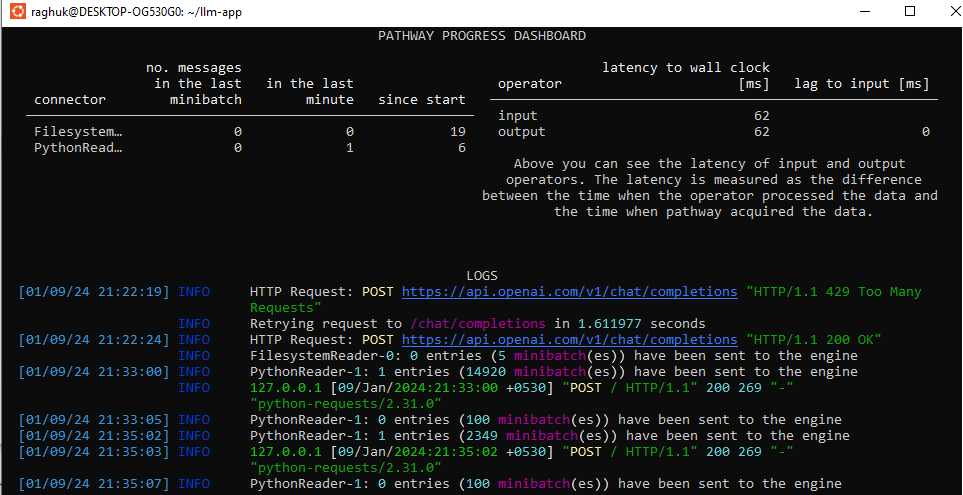


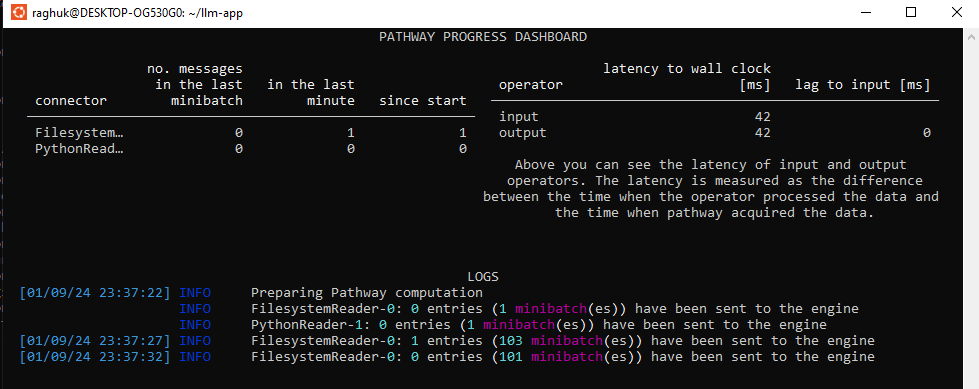
Example of a smaller JSON structure

{"doc": "The Pathway's class AsyncTransformer Open method parameters are 'capacity' sets maximum number of concurrent operations 'retry strategy' defines how failures will be handled. The Open method Returns self with result as a Table"}

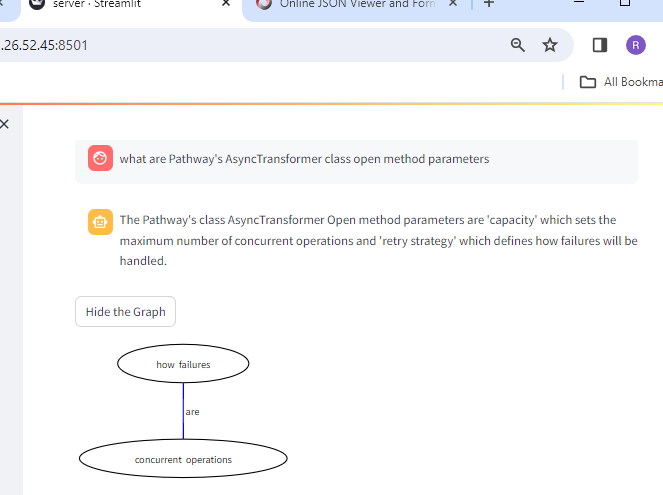
The console output showing the movement on embeddings







The embedded JSON helped to deliver an equivalent textual response. However the JSON document had one long sentence & did not yield the necessary list of entities & action words.



A sample of the reworked JSON is below

*{"doc": "The Pathway's class AsyncTransformer Open method parameters are Capacity and Retry Strategy. The Capacity parameter sets maximum number of concurrent operations. The Retry Strategy parameter defines how failures are handled. The Open method Returns self with result as a Table"*

<<Pending further updates>>