AI Integration for Web Development

I. Introduction to API Calling (Day 1)

Large Language Models (LLMs) are built upon massive computational systems due to their requirements to process large data each second. Because of this, companies usually host their own private servers where it runs all data processing and matrix calculations. An API acts as a key for developers to access just the model to work with, without risking any tamperings of the main servers.

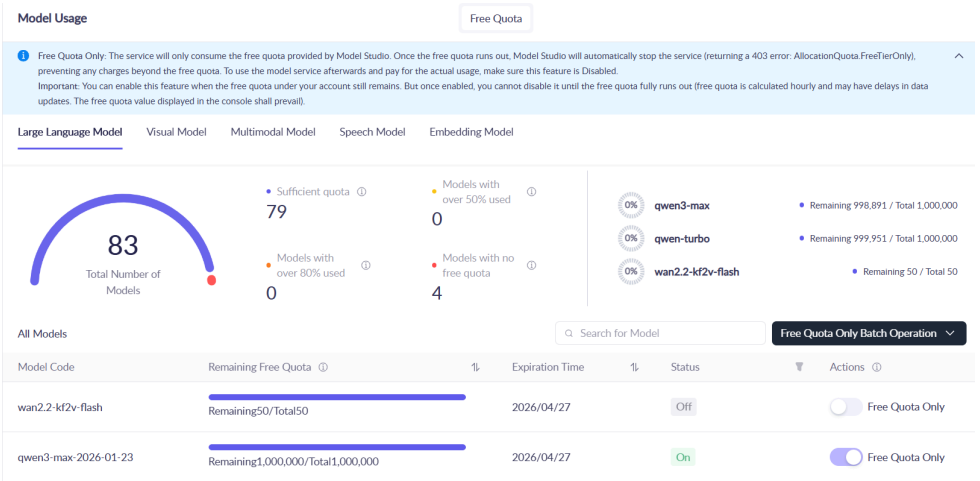
A. API Functions Towards Interface Standards and Security

An API sets the standard for all users on how they interact with the LLM. Developers using the API key to access the model obtain a standardized collection of functions and tools tailored to the use of that specific model. We can think of this as adding an extra library to include in an IDE, which lets developers write predictable code.

Having access only to the model also isolates your data and code from other developers and vice versa. Additionally, any improvements made on the AI model behind the API would not have any affect on a developer's code.

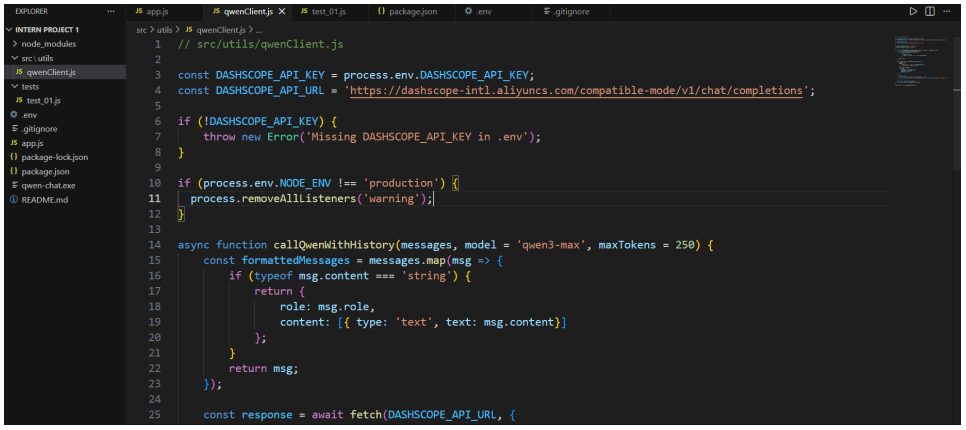
B. Basics of API Key Application

In order to carry out a test of working with an existing LLM, I chose the company Alibaba which has an abundance of working AI models comparable to OpenAI's chatGPT. After creating a free account and obtaining an API key, we gain access to more than 70 AI models with a **free** quota of 1 million output tokens.

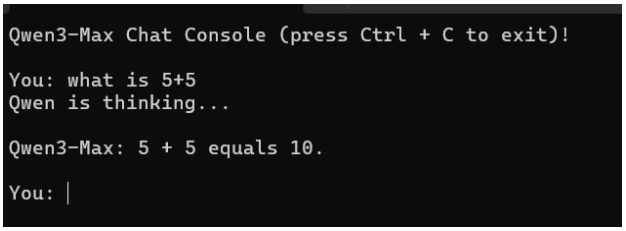


Before working with any of those models, an essential step is to store the private API key in the environment variables or a .env file. This step ensures security on the data from other developers and avoids getting banned from the server for any abuse.

After creating a project folder, we use java script to create a default qwenClient.js file to call Qwen. This enables all other applications to access the model without code repetition. It should handle authentication, message formatting specific to Qwen, error handling, etc. Standardizing the input format is important since most LLMs use a multimodal input format.



qwenClient.js exports a function ‘callQwenWithHistory’ which works to analyze a string of inputs using the default ‘qwen3-max’ model. Calling the function from any external .js file like the example below can be used to create a command line interface (CLI) where Qwen is integrated in. In this case, we can hold a conversation with the AI model with an expected output under 500 tokens.



II. AI Integration In Inputting Product Attributes (Day 2)

A. Problem Scenario