DragonFly

MFE: Web GPT

Version: 1.0

Date: 10-05-2024

EY

Building a better
working world

Contents

1   Introduction

Our platform facilitates seamless conversations with our AI assistant, leveraging advanced natural language processing algorithms with diverse models like 'gpt-35-turbo,' 'text- davinci-003 to comprehend your input contextually and deliver personalized, informative responses. With expanded capabilities including answering questions, explaining concepts, and assisting in writing tasks, our platform aims to be a user- friendly solution meeting diverse needs effectively.

Our platform eliminates the need for extensive web searches. With Web GPT, simply ask a question, and our AI assistant retrieves relevant information from its extensive document database, providing precise answers. Whether you're a businessperson, professional, or simply curious, our AI assistant is poised to assist you throughout.

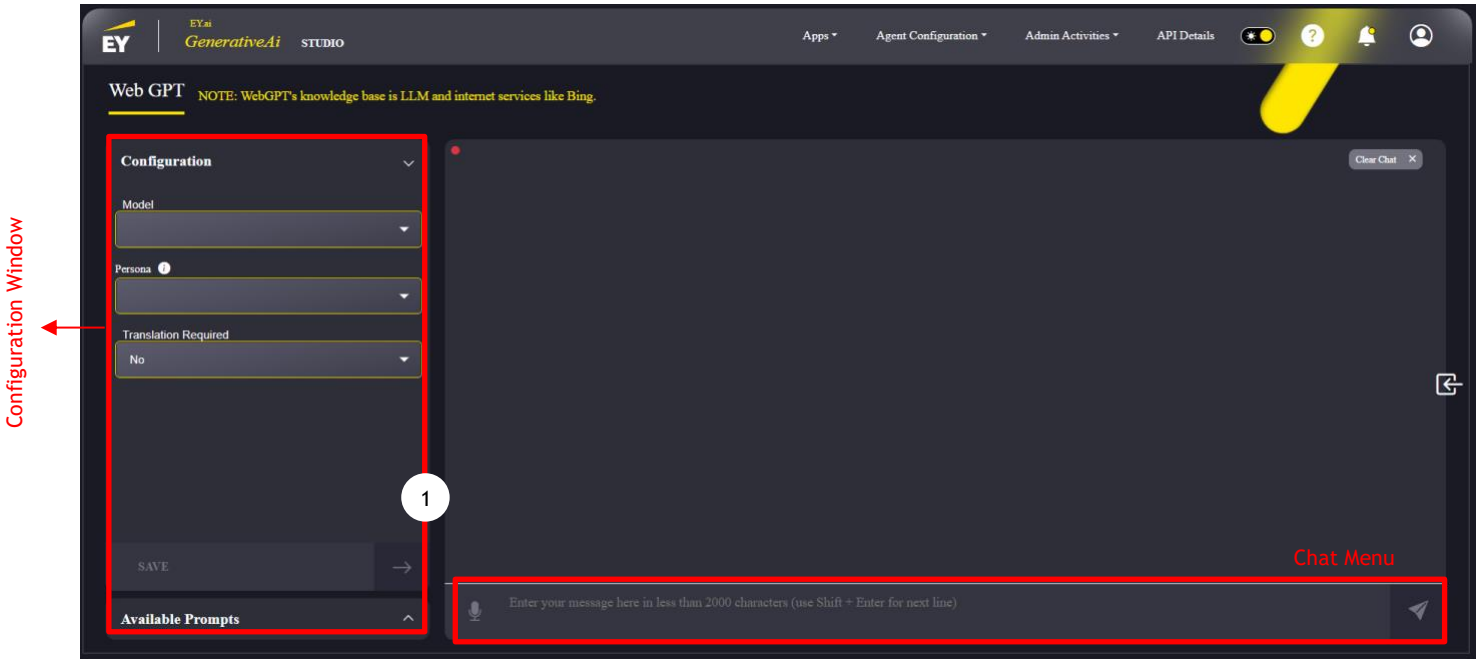Here's what the loading window might look like:

Image 1

2   Configuration

As mentioned above Box 1 highlights the Configuration Window. Here, you can fine-tune the settings and preferences to suit your specific needs.

From selecting specific areas for data extraction to choosing the output format, the Configuration Window puts the power of our AI technology directly into your hands.
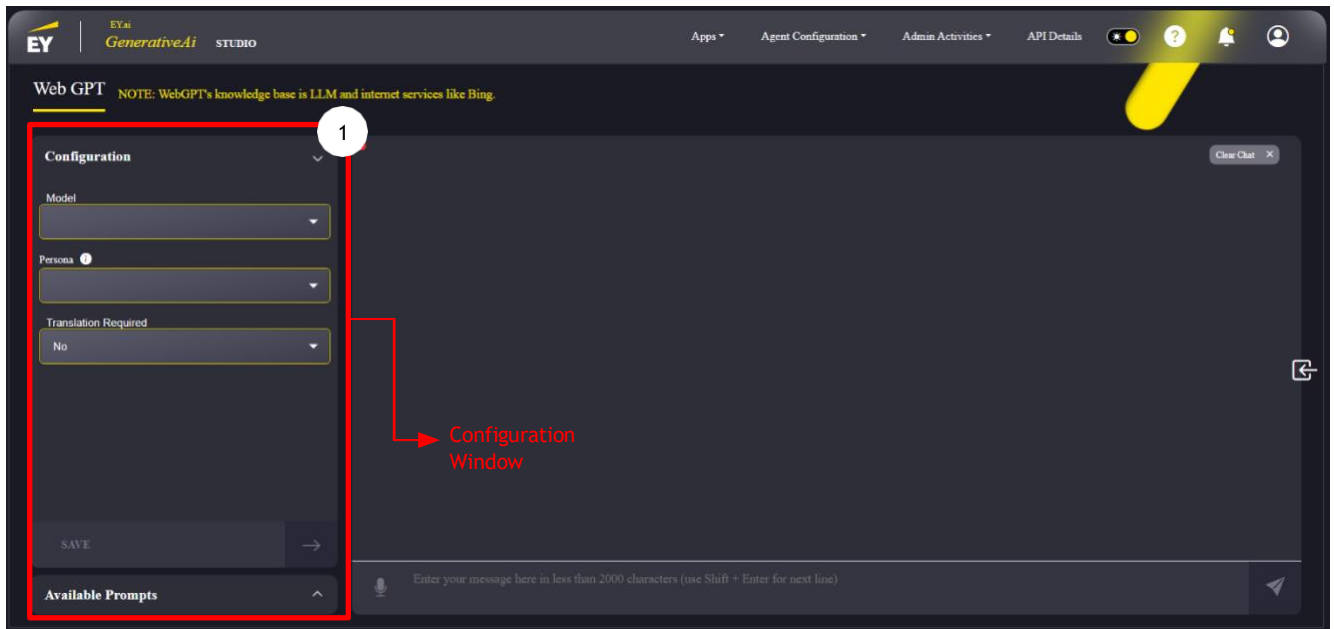

Image 2

Setting Up configuration

Step 1: Model Selection

Select among following models. Box 1.1 in image 3 highlights the various Model Options, you can choose form. A brief description and differentiation of different models available is provided below:

- **GPT-4** is the latest and most advanced LLM from Open AI. It is a multimodal model, meaning that it can process and generate both text and code. GPT-4 is also significantly larger and more powerful than previous LLMs, with 100 trillion parameters. This makes it capable of performing a wider range of tasks with greater accuracy.

- **GPT-35 Turbo 16K** is a smaller and less powerful version of GPT-35 Turbo. It has 16,385 parameters, compared to GPT-35 Turbo's 4097 parameters. This makes it less capable of performing complex tasks, but it is also more efficient and less expensive to use.

- **GPT-4 32K** is a large language model (LLMs) that has been trained on massive datasets of text and code. These LLMs can generate text, translate languages, write different kinds of creative content, and answer your questions in an informative way. However, each model has its own strengths and weaknesses, and is better suited for certain tasks than others.
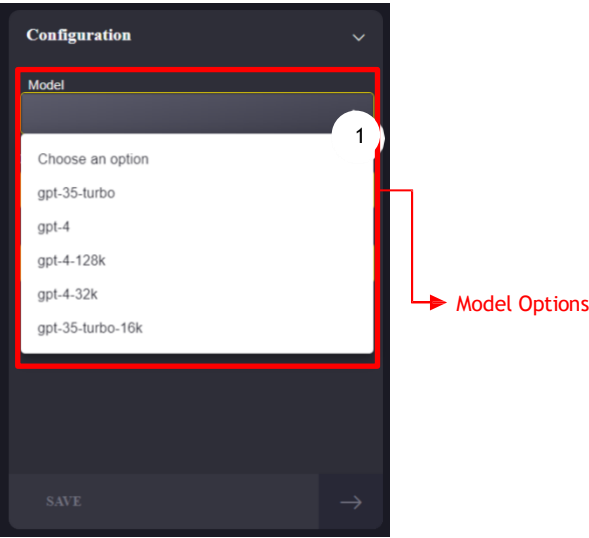
Image 3

Step 2: Parameter Selection

Once the model has been established, the subsequent four fields of Temperature, Top P, Frequency Penalty, Presence Penalty must be selected. This is highlighted by Box 1.3 in Image 4. However, users have the flexibility to modify these values to suit their desired functionality.

The next 4 fields are:

- **Temperature:** The temperature parameter controls the randomness of the generated texts. A higher temperature means more diversity and creativity, while a lower temperature means more coherence and consistency. In this case, it is set to .7, which is a moderate value, ensuring randomness. Select a value between 0 to 1.

- **Top P:** The "top p" parameter is like a filter that controls how many different words or phrases the language model considers when it's trying to predict the next word. If you set the "top p" value to 0.5, the language model will only consider the 50 most likely words or phrases that might come next.

- **Frequency Penalty:** This parameter is used to discourage the model from repeating the same words or phrases too frequently within the generated text. It is a value that is added to the log-probability of a token each time it occurs in the generated text. A higher frequency_penalty value will result in the model being more conservative in its use of repeated tokens.

- **Presence Penalty:** This parameter is used to encourage the model to include a diverse range of tokens in the generated text. It is a value that is subtracted from the log-probability of a token each time it is generated. A higher presence_penalty value will result in the model being more likely to generate tokens that have not yet been included in the generated text.
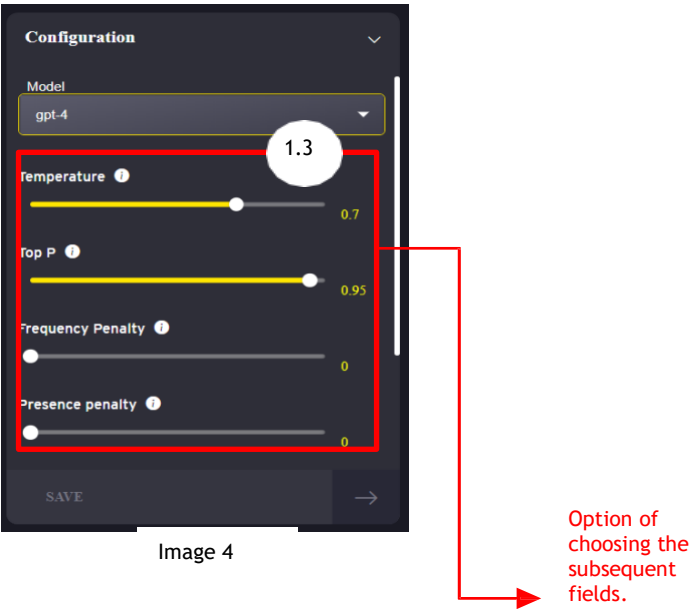
Image 4

Step 3: Persona Selection

Once the above four fields are selected, the next step is selecting a 'Persona'. This is a unique feature of our Direct- allows you to tailor the analysis based on your specific role or needs. A persona, in the context of artificial intelligence and virtual assistants, refers to the character or personality that the AI system embodies. It can include aspects like the AI's tone, style of speech, values, and even 'backstory'. A persona helps make interactions with the AI more engaging, relatable, and human-like. It's designed to enhance the user experience, making the AI seem less like a machine and more like a human assistant. Choose the persona that best fits your profile or create a persona management from the Agent configuration window. By doing this, the system can prioritize and highlight the information that is most relevant and useful to you. Image 5 highlights the various Persona options available. Once you have selected the model, you can choose an AI expert from the persona selection, or you can update the persona using the dropdown menu.
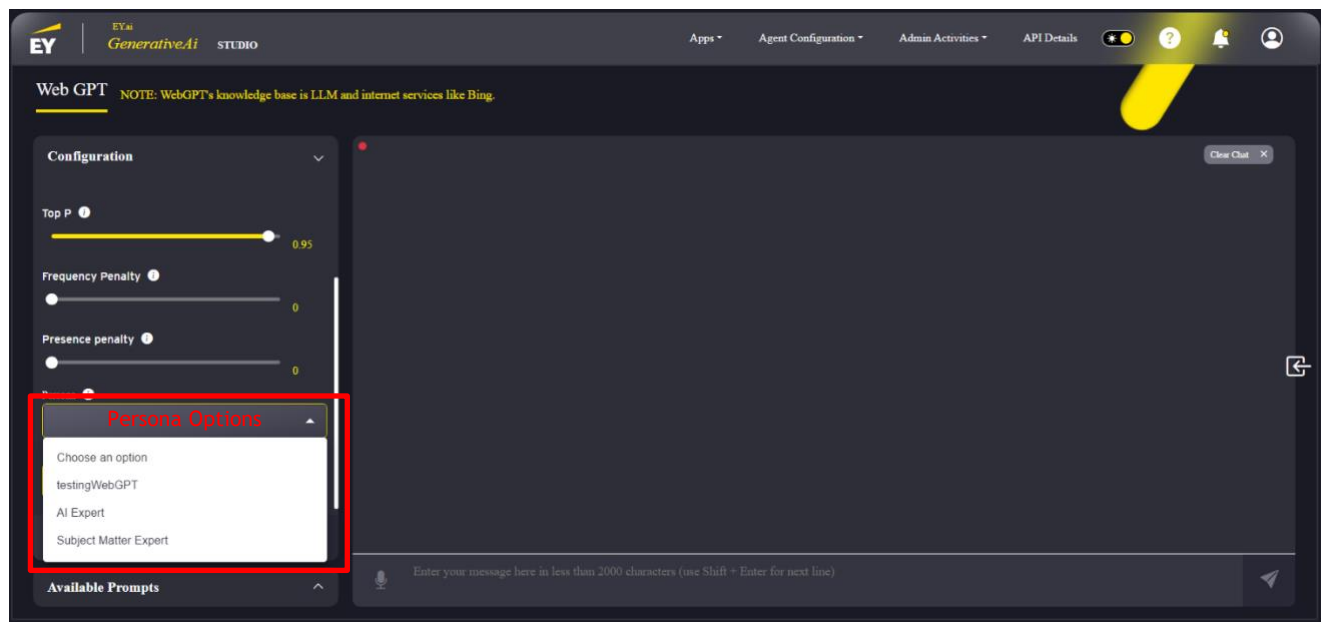

Image 5

On the right-hand side of the screen, you will find a Persona Button. This button is highlighted by the icon highlighted by Box 2.1 in Image 6. By clicking on this, a new Persona Configuration window, as highlighted by Box 2.2 in Image 6, will be available you can easily update or change your persona at any time. This flexibility ensures that as your needs evolve or change, the Website Intelligence system can adapt right along with you, always providing the most relevant and tailored insights. User can update the persona as per his needs. The modified persona is user specific as it is not updated in database and is saved in session storage which is removed on application logout/session close.
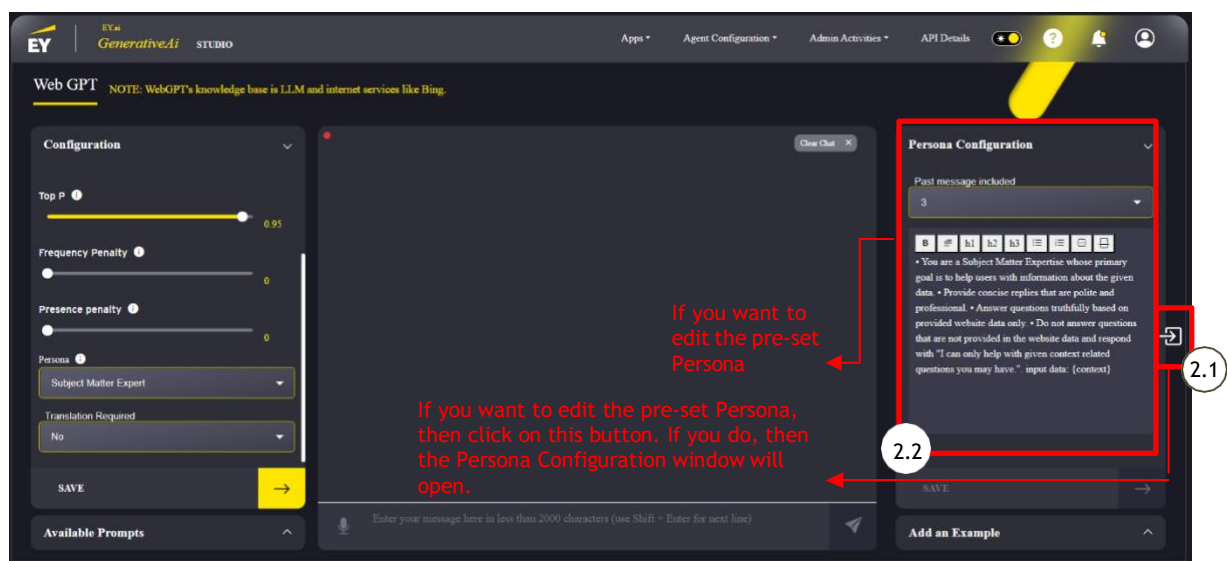

Image 6

DragonFly can store past messages. This feature allows the AI to remember earlier queries and provide contextually accurate

responses. Select the number of past messages you want the AI to remember. Click on save button once you are done with Persona configuration.

Persona Examples:

- To add a persona example, click on the Add an example menu insert both the question and answer and click on save button available in yellow color. The "Add an Example" functionality allows user to provide specific examples or scenarios to help clarify or illustrate a concept or question. It helps the AI Assistant to better understand the concept and enables it to provide more accurate and tailored responses. Pre-Defined examples are loaded from database for selected persona. User can add/update the examples as per his needs. The modified examples are user specific. They are not updated in database and are saved in session storage which is removed on application logout/session close.
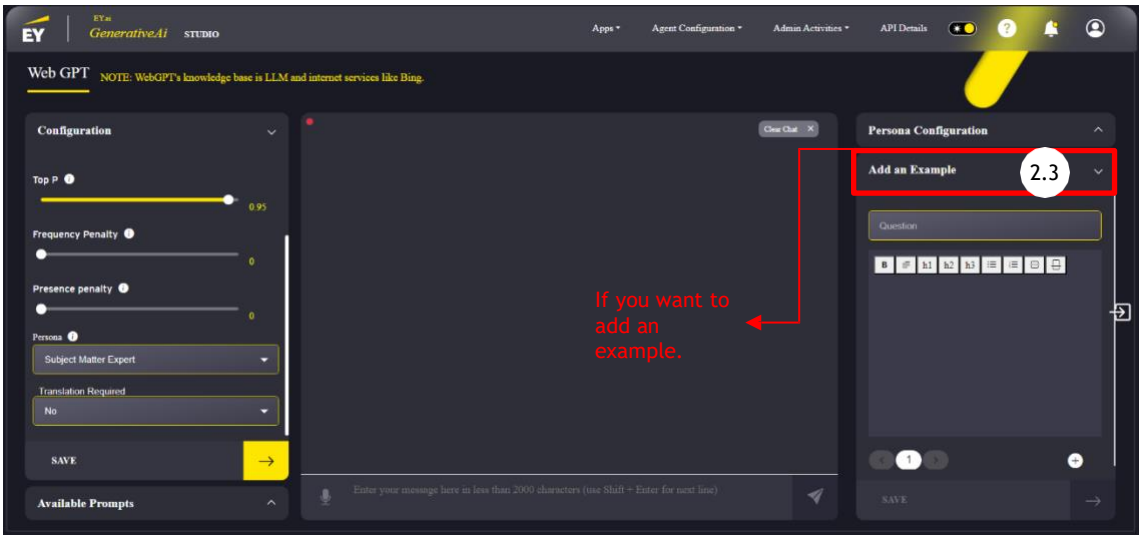- Click on Save once examples are added.

Image 7

Step 4: Translation

Translation Required:

In the Agent Configuration section, you'll find a button labeled 'Translation Required', as highlighted by Box 1.4 in Image 8. This feature is specifically designed for website queries in languages other than English. If translation is required in language other than English, select "Yes".

A new dropdown containing 3 translators will be available.

- Bhashaverse
- Azure
- Google

Select the required translator which will further open other two dropdowns for Source and Target language and click on save button a loader stating DragonFly is ready will be available.
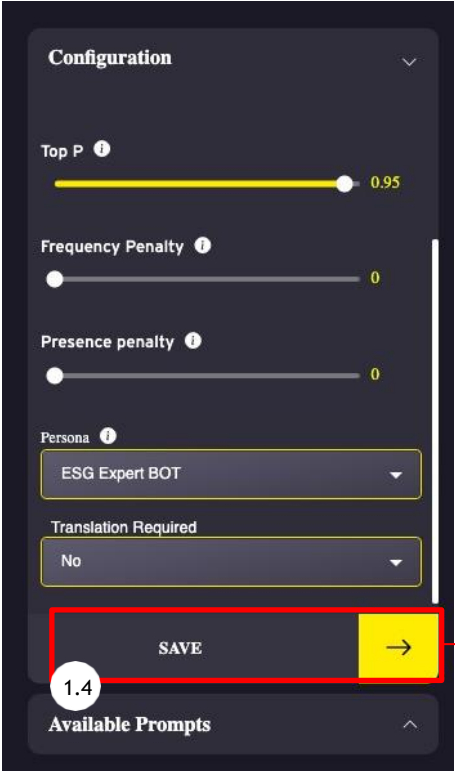
Image 8

Choosing the option of a translator, if required

Translation Button, particularly designed for websites in languages other than English
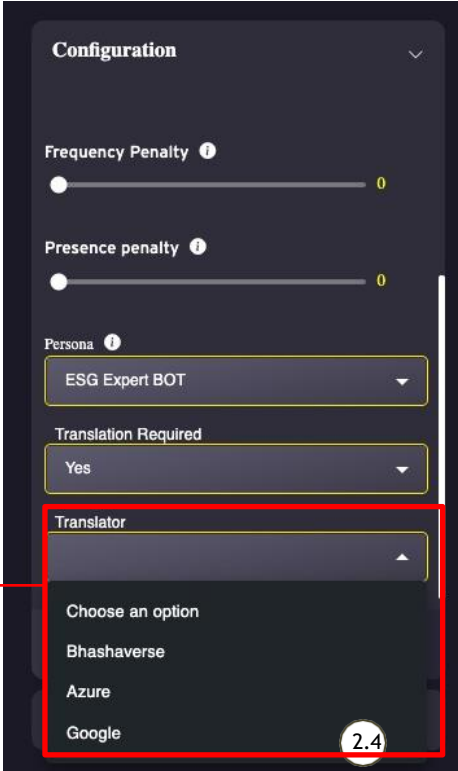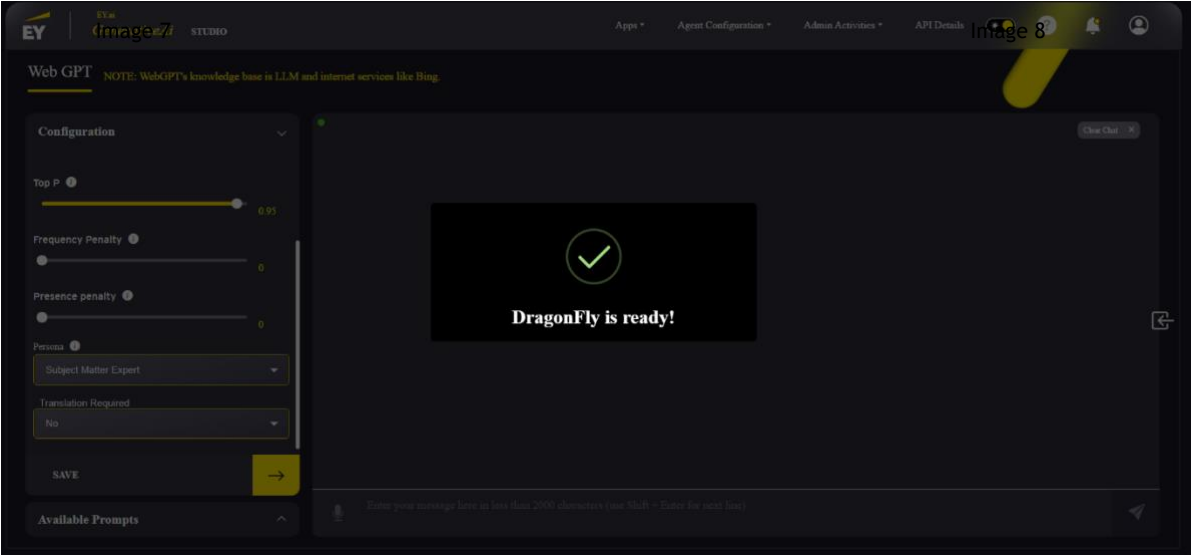
Image 9

Image 10

3   Interaction with DragonFly

Once the 'DragonFly is ready' message appears, you're all set to start your analysis. Feel free to begin inputting queries. The system is now prepared to provide you with detailed answers based on your specific needs and the persona you've selected.

There are three different ways to ask queries using the available prompts: typing, recording, or utilizing the provided prompts.

Prompts: The available prompts act as cues, guiding the AI assistant in generating relevant and coherent responses. Prompts are user specific and are fetched from database for selected persona and logged in user. Prompts play a crucial role in facilitating engaging conversations, enhancing user experiences, guiding creative thinking, improving language generation, and enabling personalized interactions with AI system.
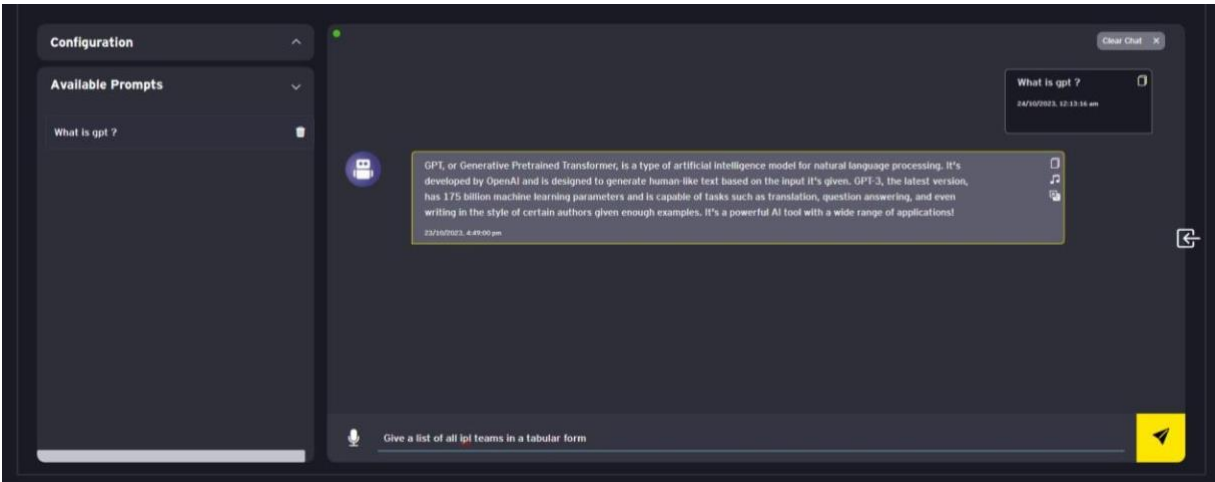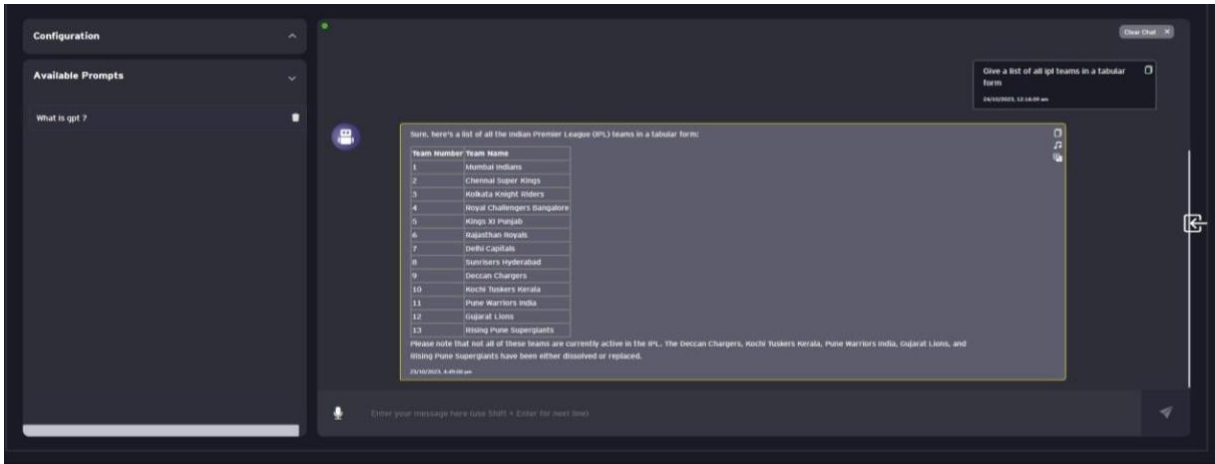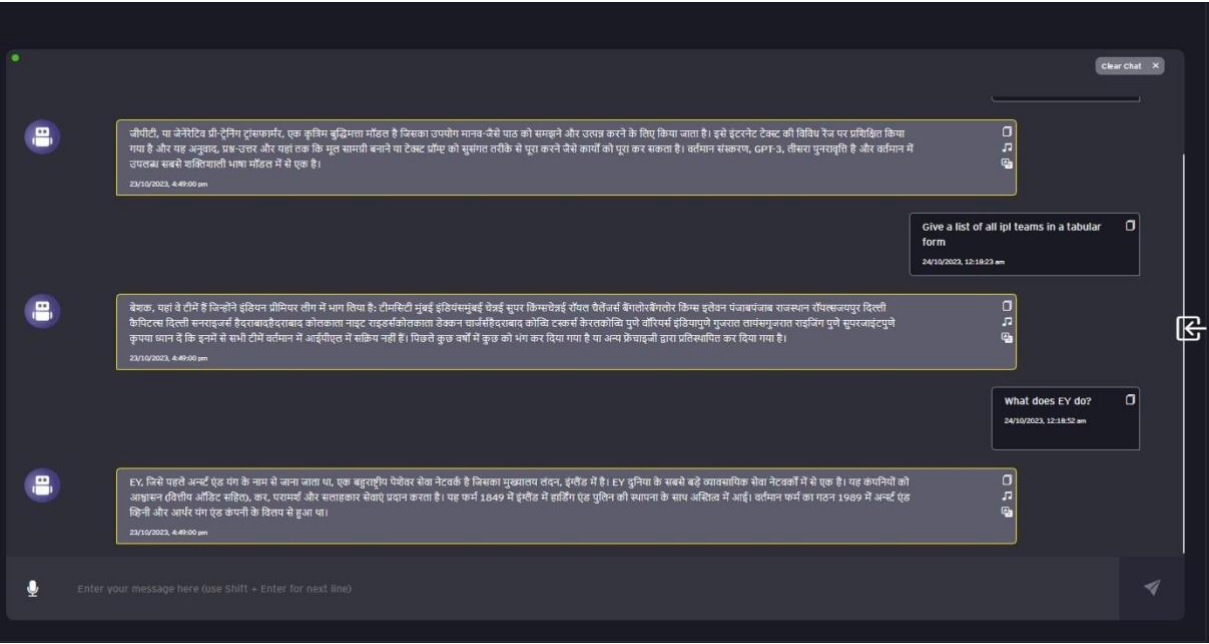

Image 11

Image 12

Image 13

<div style="background-color: yellow">

4   Demonstration

</div>

Please find the attached video of demonstration of Web GPT.