Usage instructions:

1. Launch the product via 1-click. **Please wait until** the instance passes **all** status checks and is running. You can connect using your Amazon private key and '**ubuntu**' login via your SSH client.

To update software, use: **sudo apt update && sudo apt upgrade -y**

Note: Because this an API server you will also need to **open** the "Inbound Rules" (under security) ports **5000-5001 (0.0.0.0/0)**

2. **Important: You will need to create your API key.**

For More Information on OpenAI see: **https://openai.com/api/**

3. **Once you are logged into the server, create and activate your Virtual Environment, use:**
   python3 -m venv gpt-venv

   source gpt-venv/bin/activate

4. **Once you're inside the virtual environment. Navigate to your project directory**

   cd gpt-app

   > *(Optional)*
   >
   > To change the port. Currently set to **5001**, run
   >
   > **nano app.py**
   >
   > Change the port to for example: **5002**
   >
   > **Save & exit**

5. **Set the environment variable in your shell, use the following command**

   export OPENAI_API_KEY='replace-with-your-api-key'

- ***For example, it would look something like this…***

export OPENAI_API_KEY='sk-proj-jx-1JL7r3jVCtrTlfArPEAcHLtFG9_eOBF_JF4tb45l9F9roFpoYtTj2JR9pgEMK6k89HjT3BlbkFJzKM-Eo3wffe5sHMIp8mO0mnysHN7W4Eoeq-Pt2arP2bL621KpeU58LXvL0EjCvKG3TTTDpE6EA'

- **To check the API the key, run:**

  echo $OPENAI_API_KEY

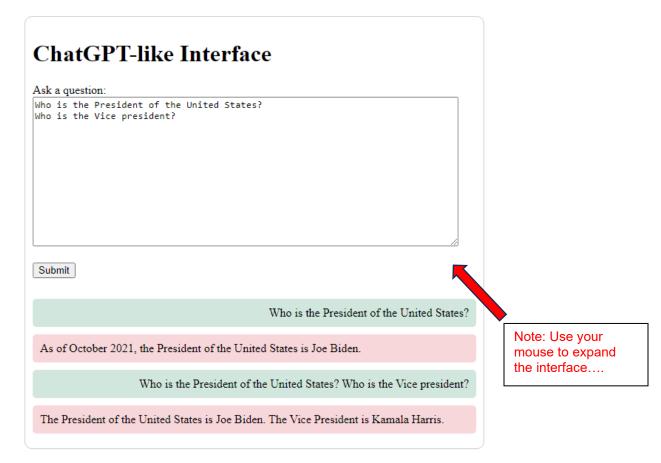6. **Start your Flask application!**

   python app.py

   Leave instance running

7. **Access the App via Browser.**

   - You should now be able to access the Flask app using your EC2 instance's public IP on port 5001: or **what port you set above**

     http://<your-instance publiciv-ip>**:5001**

     For ex: http://363.323.35:5001

Note: Use your mouse to expand the interface….

Additional Info:

The Python code, model, etc. can be found here. Be sure to be inside the virtual environment:

source gpt-venv/bin/activate

cd gpt-app

nano app.py

Note: **To make changes or <u>customize</u> the front-end GUI** (the HTML code):

nano templates/index.html

<u>AWS Data</u>

- Data Encryption Configuration: This solution does not encrypt data within the running instance.

- User Credentials are stored: /root/.ssh/authorized_keys & /home/ubuntu/.ssh/authorized_keys

- Monitor the health:

  - Navigate to your Amazon EC2 console and verify that you're in the correct region.

  - Choose Instance and select your launched instance.

  - Select the server to display your metadata page and choose the Status checks tab at the bottom of the page to review if your status checks passed or failed.

<u>Extra Information:</u> (Optional)

Allocate Elastic IP

To ensure that your instance **keeps its IP during restarts** that might happen, configure an Elastic IP. From the EC2 console:

1. Select ELASTIC IPs.
2. Click on the ALLOCATE ELASTIC IP ADDRESS.
3. Select the default (Amazon pool of IPv4 addresses) and click on ALLOCATE.
4. From the ACTIONS pull down, select ASSOCIATE ELASTIC IP ADDRESS.
5. In the box that comes up, note down the Elastic IP Address, which will be needed when you configure your DNS.
6. In the search box under INSTANCE, click and find your INSTANCE ID and then click ASSOCIATE.
7. Your instance now has an elastic IP associated with it.
8. For additional help: https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/elastic-ip-addresses-eip.html