Proxy API
Example of use: Image generation

Alejandra Caggiano

As we continue to learn more about the APIs offered by GeneXus Enterprise AI, we are now going to learn about the Proxy API. This API acts as a transparent proxy for companies to easily connect to several Artificial Intelligence models from different providers through a single access point. It is possible to access any supported LLM with a single SDK.

Proxy API

➢ Seamless Access to LLMs.

➢ Simplified API Requests.

➢ Costs and performance monitoring.

➢ Future-Proof.

Let's take a look at its standout features:

First of all, **seamless access to LLMs**.

With the OpenAI SDK as a single interface, you can get immediate access to a wide variety of supported LLMs, including those from OpenAI, Azure, Amazon and Google. It is also a transparent proxy, as it works seamlessly in the background, automatically routing the various requests to the designated AI models and without the need for code changes.

Another feature to note are the **simplified API requests**.

The GeneXus Enterprise AI Proxy API applies specific API Tokens, eliminating the requirement for individual API keys for each AI provider. It also maintains a consistent API request format, regardless of the AI model being accessed.

Another key feature is **cost and performance monitoring**.

It is possible to obtain information on the performance and profitability of each model without having to integrate multiple monitoring solutions.

Lastly, it is f**uture-proof** because its integrations remain intact as support for more LLMs is added, saving future development time.

Proxy API

URL format

https://api.qa.saia.ai/proxy/{AIProviderName}/{hostRelativeProviderURL}

Header Authorization

Should be replaced by the GeneXus Enterprise AI Token api

Here are some examples of how to use this API.

To integrate the use of this GeneXus Enterprise AI proxy, no programming changes are required. Modifications simply involve alterations of two elements in each request:

1. On one hand, the base URL used to invoke AI providers must reference the GeneXus Enterprise AI component. The format of the URL to be used is as follows:

https://api.qa.saia.ai/proxy/{AIProviderName}/{hostRelativeProviderURL}

where this part, as we already know, corresponds to the content of the BASE_URL variable.

2. On the other hand, the Authorization header included in each request must be replaced with the GeneXus Enterprise AI API Token previously provided, without the need to specify the invoked provider.

Proxy API

https://api.openai.com/v1/chat/completions

https://api.qa.saia.ai/proxy/openai/v1/chat/completions

https://api.replicate.com/v1/predictions

https://api.qa.saia.ai/proxy/replicate/v1/predictions

In this way, the request and the response received from the call will be the same as when invoking the provider directly.

For example, if the original URL invoked is:

https://api.openai.com/v1/chat/completions

To use the GeneXus Enterprise AI HTTP Proxy, the call must look like this:

https://api.qa.saia.ai/proxy/openai/v1/chat/completions

For the following original URL:

https://api.replicate.com/v1/predictions
…when using the GeneXus Enterprise AI HTTP Proxy, it should be like this:

https://api.qa.saia.ai/proxy/replicate/v1/predictions

Proxy API: Scenarios that we can solve

➤ Integrate GeneXus Enterprise AI with third-party SDKs.

➤ Call DALL-E-2 and DALL-E-3 to generate images.

➤ Call GPT-4 to interpret images.

What specific scenarios can we solve?

- GeneXus Enterprise AI can be integrated with third party SDKs,

- Interact with artificial intelligence systems to generate images,

- Interact with GPT-4 to interpret images.

Example of use: Interact with DALL-E-2 to generate images

DALL-E-2 - cURL

```
curl --location 'https://api.saia.ai/proxy/openai/v1/images/generations' \
-H 'Content-Type: application/json' \
-H 'Authorization: $SAIA_APITOKEN" \
-d '{
    "model": "dall-e-2",
    "prompt": "a halloween pumpkin",
    "size": "1024x1024"
}'
```

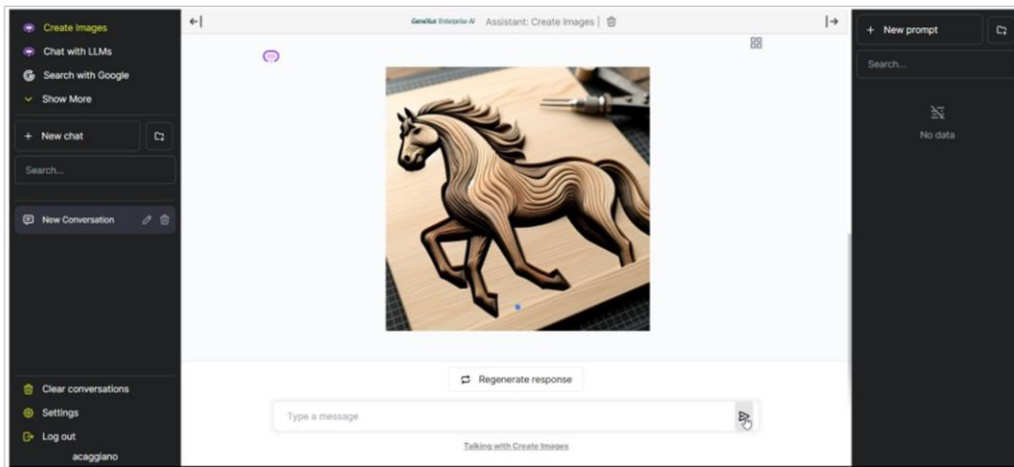https://api.qa.saia.ai/proxy/openai/v1/images/generations

As an example of use, let's see how to interact with DALL-E-2 to generate images. This AI system can create realistic images and art from a description in natural language.

If we look at the cURL sample, we note that we need an API token and that the corresponding URL in our context will be as follows:

https://api.qa.saia.ai/proxy/openai/v1/images/generations

Example of use: Interact with DALL-E-2 to generate images

Therefore, from Postman we define the POST.

Now we define the authorization and indicate the token. It is important to mention that in our previous examples we have always used the default API token. However, it is advisable to define API tokens for the use of each API and in this way monitor the statistics of the requests and associated costs.

So from the platform we create a new API token... copy... and paste it.

Good. We go to the Body tab, Raw, JSON, and paste the corresponding sample.

Let's ask for the image of a white tiger.

We click on Send. As a response, it returns the URL of the generated image.

So we access it.

Now let's ask for the image of a green lion with a size of 512x512S.

OK, so now we know how to generate any image by interacting through this API.

Creation of images from the GeneXus Enterprise AI Playground

But, do we have a way to create images directly from GeneXus Enterprise AI?

Yes; let's go to the platform and select Playground.

Here we see the Create images option.

Let's ask for the image of a wooden horse to be generated.

What can we do now? We can save it by right-clicking on it, or access its URL.

For that, we select this Copy option, which returns the complete answer with the URL of the generated image.

GeneXus by Globant

GeneXus™
by **Globant**