

Proxy API

Exemplo de uso: Geração de imagens



Alejandra Caggiano

Avançando um pouco mais no conhecimento das apis que GeneXus Enterprise AI oferece, vamos conhecer agora a Proxy Api. Esta API surge como um proxy transparente para que as empresas se conectem facilmente a vários modelos de Inteligência Artificial de diferentes provedores, através de um único ponto de acesso. É possível acessar qualquer LLM compatível com um único SDK.

Proxy API: Características que se destacam

- Acesso perfeito aos LLM
- Solicitações de API simplificadas
- Monitoramento de custos e desempenho
- Preparado para o futuro

Vejamos as características que se destacam:

Em primeiro lugar, o **acesso perfeito aos LLM**

Com o SDK de OpenAI como interface única, é possível obter acesso imediato a uma grande variedade de LLMs compatíveis, incluindo aqueles de OpenAI, Azure, Amazon e Google.

É também um proxy transparente, pois funciona sem problemas em segundo plano, encaminhando automaticamente as diferentes solicitações para os modelos de Inteligência Artificial designados e sem a necessidade de realizar alterações no código.

Outra característica a destacar são as **solicitações de API simplificadas**

A Proxy api de GeneXus Enterprise AI aplica API Tokens específicos, eliminando a necessidade de chaves API individuais para cada provedor de inteligência artificial.

Também mantém um formato coerente de solicitação de API, independentemente do modelo de Inteligência artificial acessado.

Outra característica fundamental é o **monitoramento de custos e desempenho**

É possível obter informação sobre o desempenho e a rentabilidade de cada modelo sem a necessidade de integração de múltiplas soluções de monitoramento.

E por último, está **preparado para o futuro**, pois à medida que é adicionado suporte para mais LLMs, sua integração permanece intacta, o que significa poupar tempo de desenvolvimento futuro.

Proxy API

URL format

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

Header Authorization

Should be replaced by the [GeneXus Enterprise AI Token api](#)

Vejamos alguns exemplos de como usar esta API

Para integrar o uso deste proxy de GeneXus Enterprise AI não são necessárias alterações no nível de programação. As modificações envolvem simplesmente alterações de dois elementos em cada solicitação:

- Por um lado, a url base utilizada para invocar provedores de inteligência artificial deve fazer referência ao componente de GeneXus Enterprise AI. Então, o formato da URL a ser utilizada é o seguinte:

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

onde esta parte, como já sabemos, corresponde ao conteúdo da variável BASE_URL

- E por outro lado, o header Authorization presente em cada request deve ser substituído pelo APIToken de GeneXus Enterprise AI fornecido previamente, sem a necessidade de especificar o provedor invocado.

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>

Desta forma, a solicitação e a resposta recebida da chamada serão as mesmas de quando se invoca diretamente o provedor.

A título de exemplo, se a url original invocada era a seguinte

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

Para utilizar agora o HTTP Proxy de Genexus Enterprise AI, a chamada deve ser assim:

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

Para a seguinte url original

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

ao utilizar o Proxy HTTP de GeneXus Enterprise AI, deveria ser assim

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

Proxy API: Cenários que podemos resolver

- Integrar GeneXus Enterprise AI com SDKs de terceiros.
- Interagir com sistemas de inteligência artificial para gerar imagens
- Interagir com GPT-4 para interpretar imagens

Bem. Quais cenários concretos podemos resolver?

- Podemos integrar GeneXus Enterprise AI com SDKs de terceiros,
- Interagir com sistemas de inteligência artificial para gerar imagens
- Interagir com GPT-4 para interpretar imagens

Exemplo de uso: Interagir com o DALL-E-2 para gerar imagens

DALL-E-2 - cURL

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

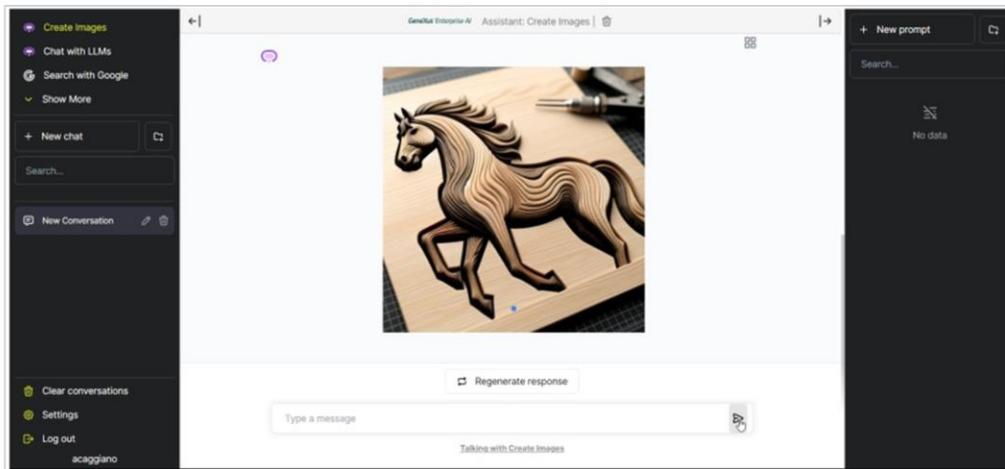
<https://api.qa.saia.ai/proxy/openai/v1/images/generations>

Como um exemplo de uso, vamos ver como interagir com o DALL-E-2 para gerar imagens. Se trata de um sistema de inteligência artificial que pode criar imagens e arte realistas a partir de uma descrição em linguagem natural.

Se observarmos o cURL sample, vemos que precisamos de um api token e que a url correspondente em nosso contexto será a seguinte:

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

Criação de imagens a partir de GeneXus Enterprise AI



Pois bem, temos como criar imagens diretamente a partir de GeneXus Enterprise AI?

Sim. Vamos para a plataforma e selecionamos Playground.

Vemos aqui a opção Create images.

Vamos pedir que seja gerada a imagem de um cavalo de madeira.

O que podemos fazer agora? Podemos salvá-la clicando com o botão direito ou acessar sua url.

Para isso clicamos nesta opção Copy, que retorna a resposta completa com a url da imagem gerada.

GeneXus[™]
by **Globant**

training.genexus.com