

Desenvolvido por: Nathalia Trazzi e Guilherme Indiciate
Período: Agosto/2025

Chat com Agentes do Watson Orchestrate

Esse asset fornece um serviço HTTP baseado em **FastAPI** que permite a comunicação com agentes do **IBM Watson Orchestrate (WxO)** por meio de dois endpoints:

- **/get_response**: Retorna uma única resposta de um agente WxO.
- **/stream_response**: Transmite a resposta do agente em tempo real usando **Server-Sent Events (SSE)**.

Ele aceita a entrada do usuário e a encaminha para um agente WxO específico, retornando a resposta do agente ou transmitindo-a conforme chega.

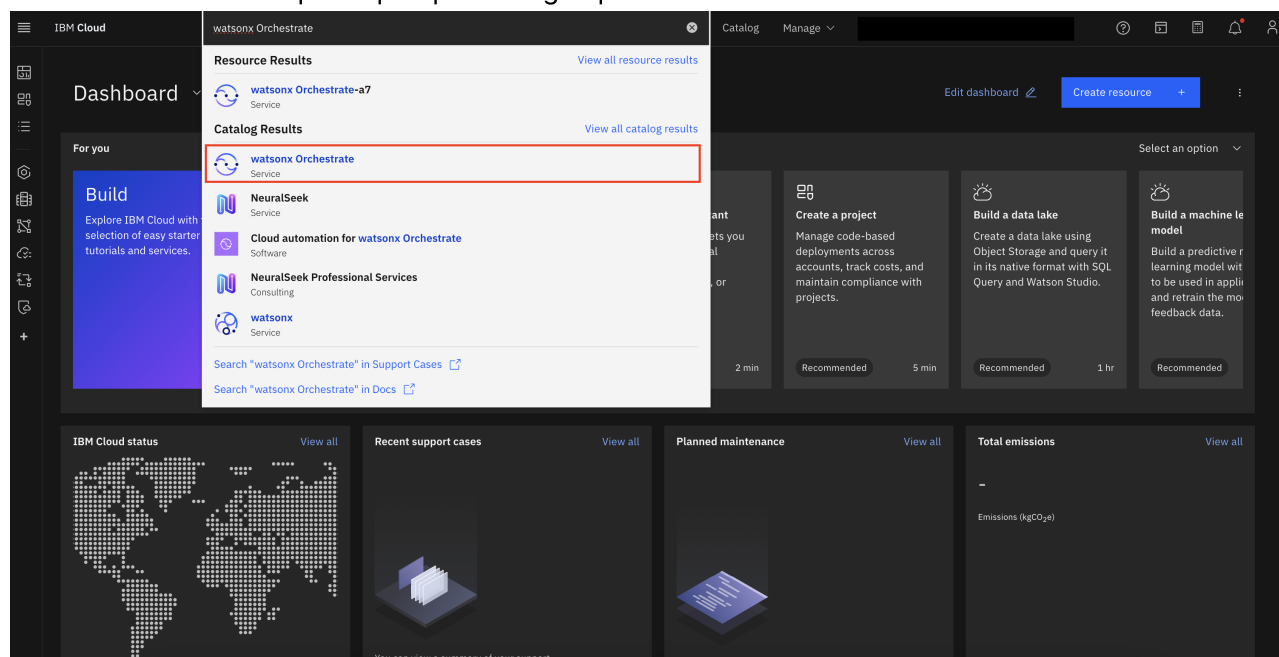
Funcionalidades

- Conversar com agentes do IBM Watsonx Orchestrate
- Endpoints FastAPI: **POST /get_response** (resposta única) e **POST /stream_response** (streaming)
- Transmite resposta da API WxO em tempo real
- Fácil de implantar com dependências mínimas

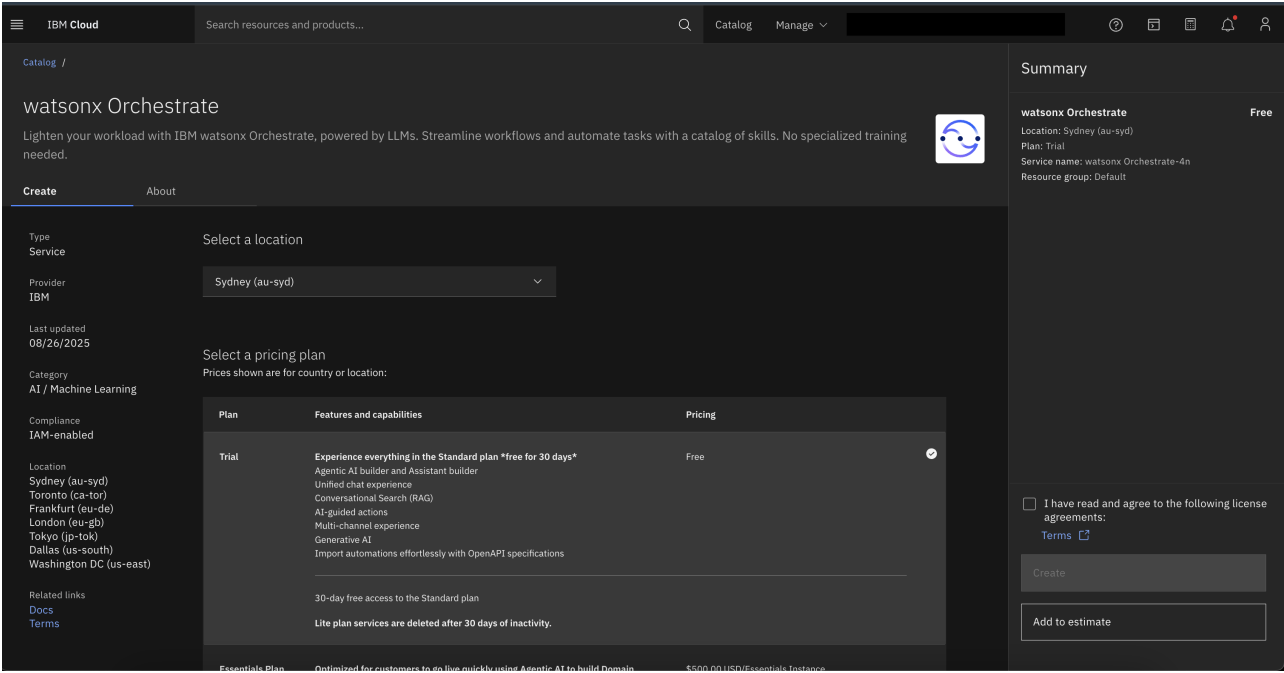
Como obter as chaves para se comunicar com o Watsonx Orchestrate via API

Para obter as chaves **IAM_API_KEY** e **WXO_INSTANCE_ID**, siga esses passos abaixo:

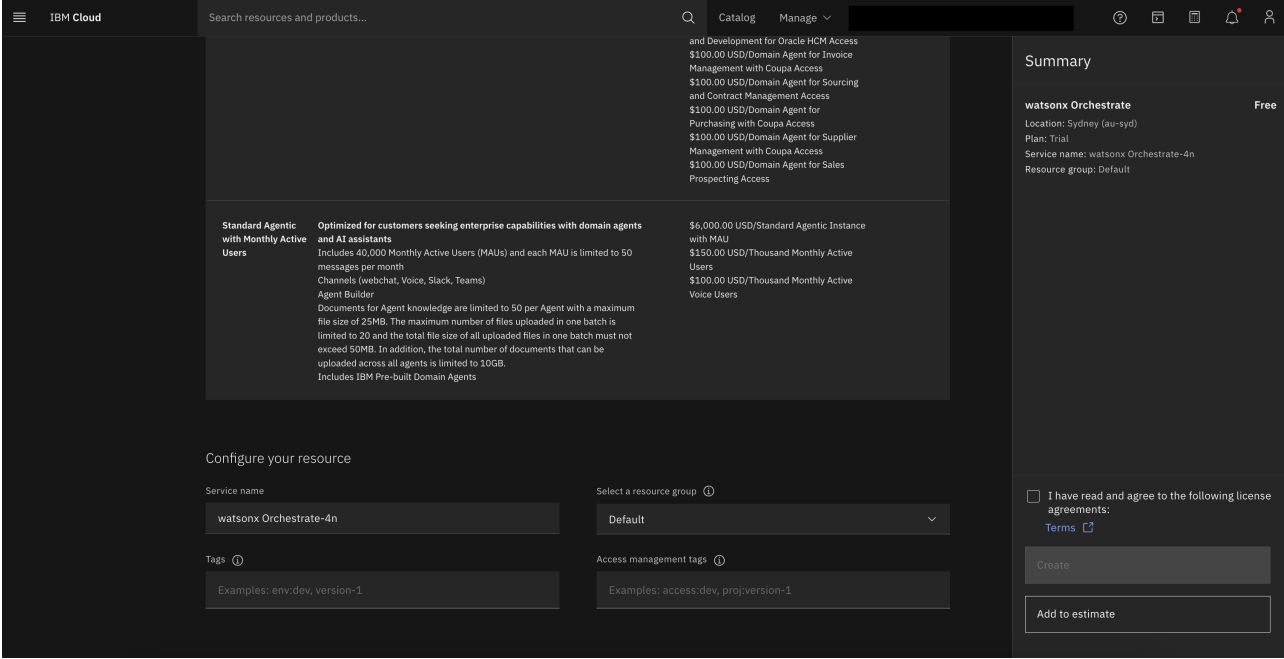
1. Se direcione até o campo de pesquisa e digite por **watsonx Orchestrate**.



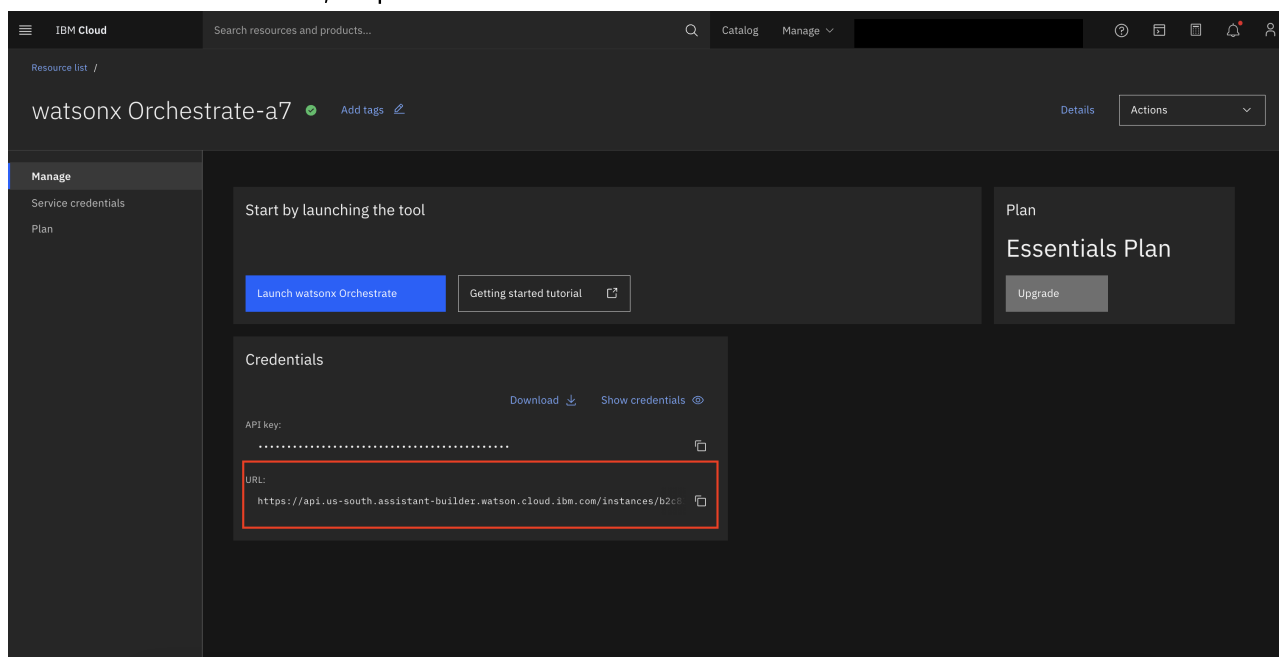
2. Escolha a **location** de sua instância e o **plan** de sua escolha.



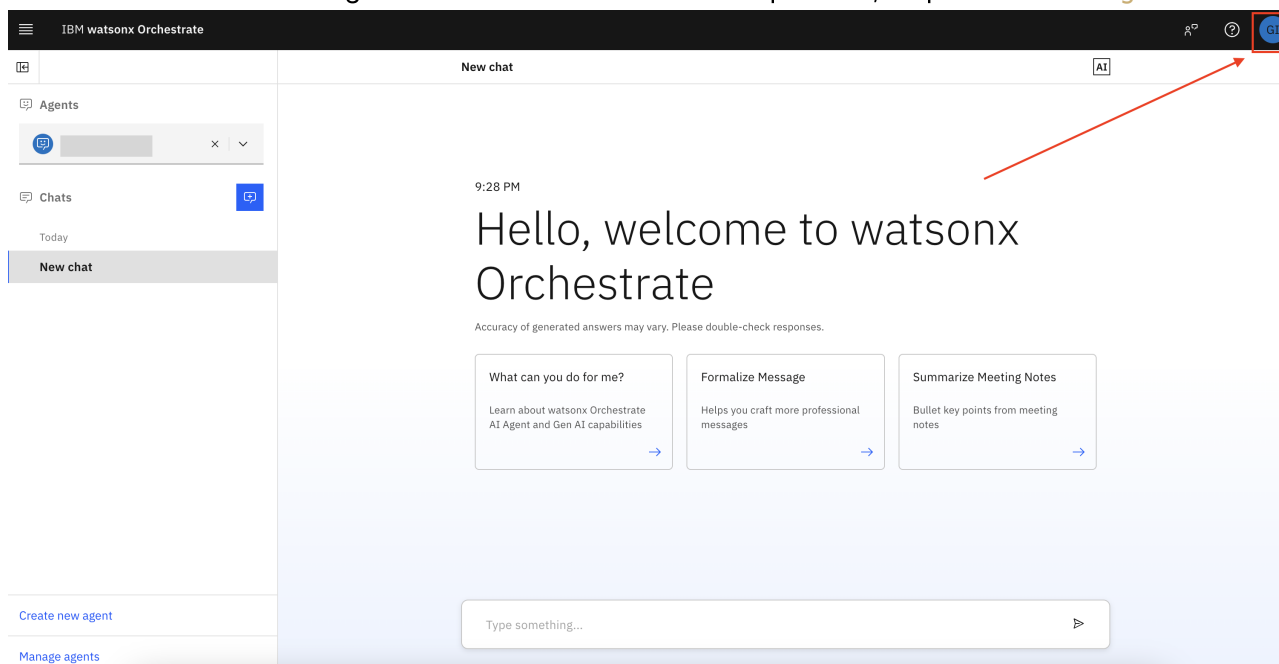
3. Por fim, digite o nome do serviço e o seu **resource group** (Default).



4. Com sua instância criada, clique em **Launch watsonx Orchestrate**.

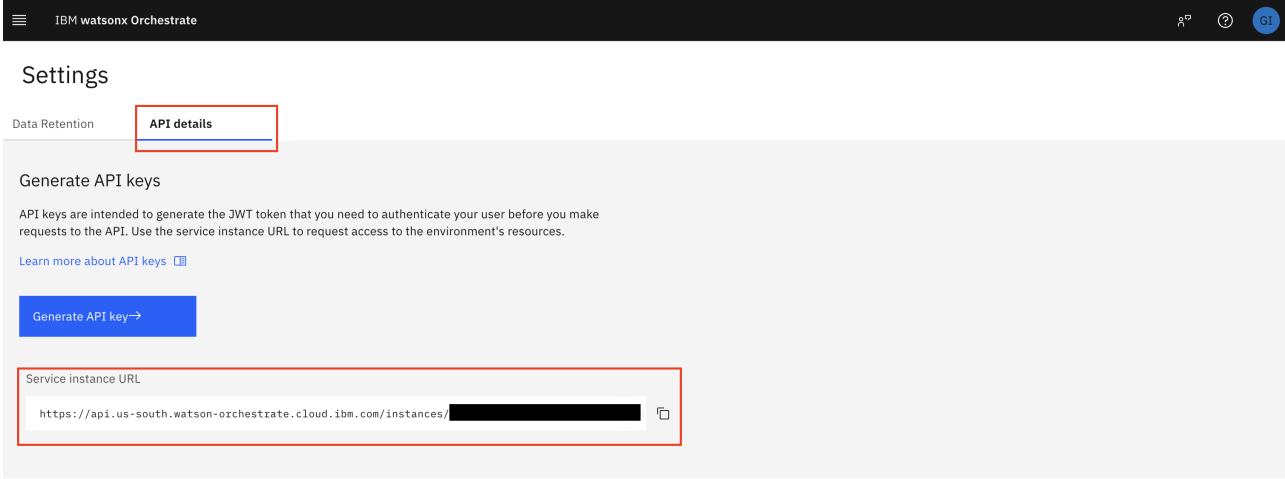


5. Se direciono ao ícone da sigla de seu nome e sobrenome e após isso, clique em **Settings**

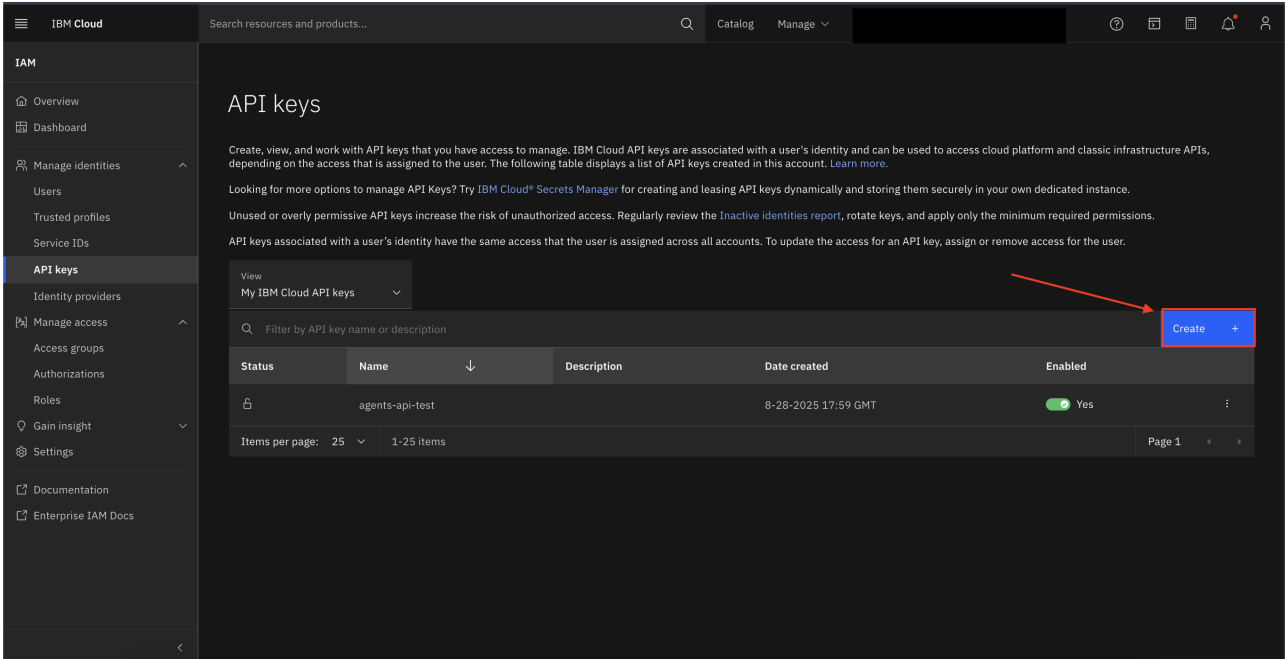


6. Settings aberta, entre na aba **API details** e armazene o ID da **Service instance url** como demonstrado abaixo:

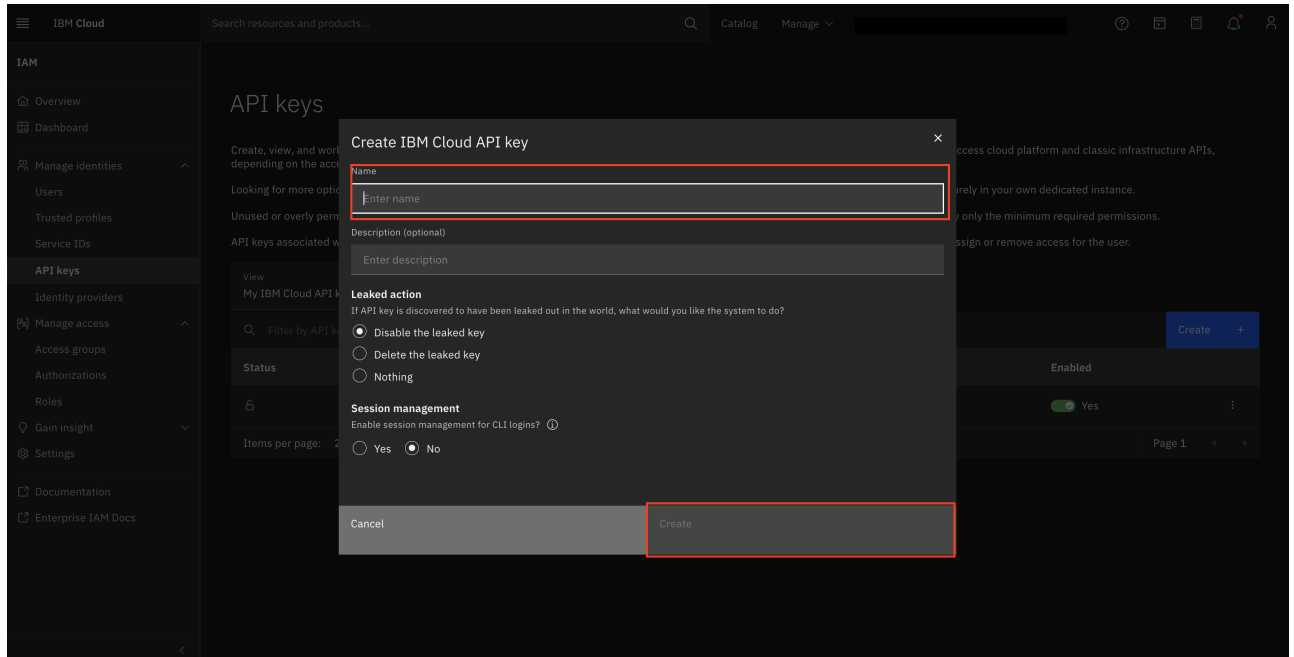
```
https://api.us-south.assistant-builder.watson.cloud.ibm.com/instances/{ARMAZENE_ESSE_ID}
```



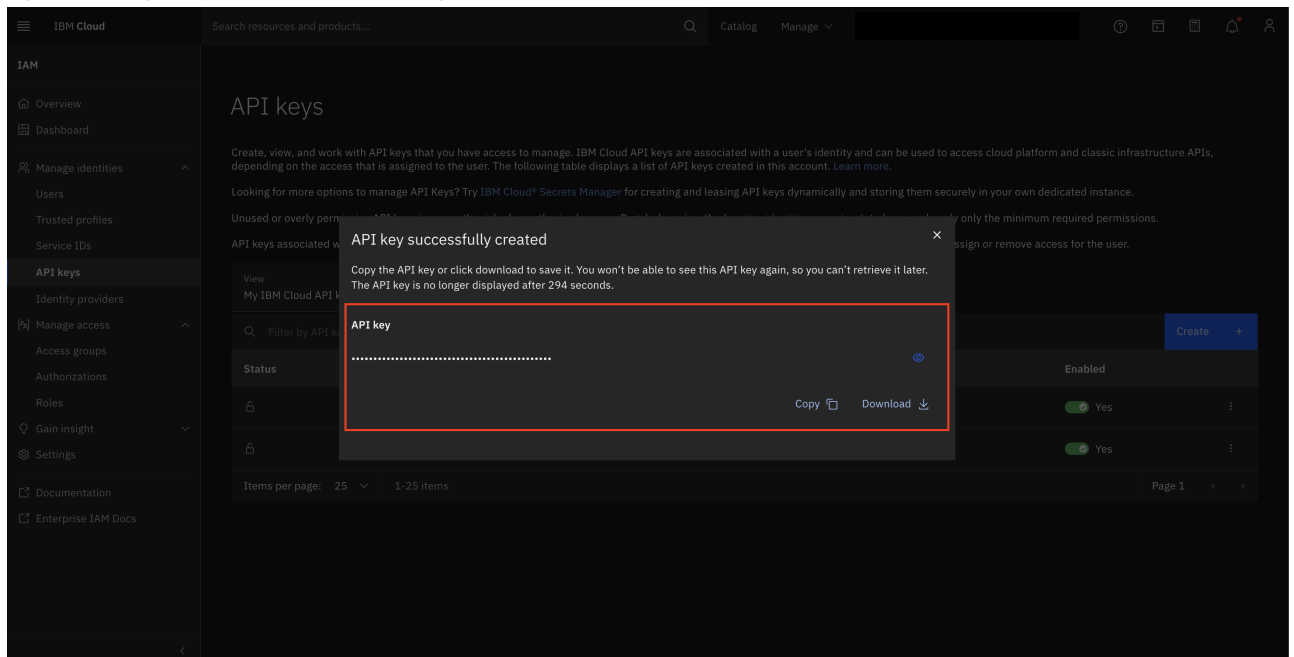
- 7. Também na mesma página, clique em **Generate API key**.
- 8. Você será redirecionado para o catálogo onde você cria, visualiza e trabalha com chaves de API que você possui acesso para gerenciar.



9. Aqui você deve digitar o nome e descrição (opcional) de sua chave de API.



10. Nesta tela, você terá até 5 minutos para copiar, salvar ou fazer o download de sua chave de API. Após esse período, não será mais possível visualizá-la novamente.



11. Após todos esses passos, você já possui as duas credenciais necessárias para a execução da interação via backend com o chat do seu agent (watsonx Orchestrate).

3. Copie o arquivo `.env_example` para `.env` e preencha com suas credenciais:

```
IAM_API_KEY=sua_ibm_cloud_api_key
WXO_INSTANCE_ID=seu_wxo_instance_id
```

4. Executar o servidor (localmente - terminal):

```
uvicorn app:app --reload
```

O servidor estará disponível em: <http://localhost:8000>

Uso da API

GET /list_agents

Recupera a lista de agentes WxO disponíveis.

Requisição:

GET /list_agents

Resposta:

```
{
  "agents": [
    {
      "agent_id": "12345678-aaaa-bbbb-cccc-1234567890ab",
      "name": "Sample Agent"
    },
  ]
}
```

Conversas e thread_id

Ao iniciar uma nova conversa, faça sua primeira chamada para `/get_response` ou `/stream_response` **sem** fornecer um `thread_id` (deixe vazio ou omita da requisição). A resposta incluirá um valor `thread_id` value.

Para continuar a conversa no mesmo contexto, inclua o `thread_id` retornado como parâmetro em todas as requisições subsequentes.

Isso permite que o agente mantenha o contexto entre as mensagens.

POST /get_response

Envia uma mensagem para um agente WxO e recebe uma resposta única.

Requisição (parâmetros de query):

POST /get_response?message=Hello&agent_id=your-agent-id[&thread_id=optional-thread-id]

Observações:

- Para a primeira mensagem, não inclua `thread_id`.

- Para mensagens seguintes, defina `thread_id` com o valor retornado na resposta anterior.

Resposta:

```
{
  "response": "The weather is sunny with a high of 26°C.",
  "thread_id": "686b5163-04a0-4687-bf86-618c3c0f59ab"
}
```

POST /stream_response

Envia uma mensagem para um agente WxO e recebe a resposta de forma contínua (SSE).

Requisição (parâmetros de query):

POST /stream_response?message=Hello&agent_id=your-agent-id[&thread_id=optional-thread-id]

Observações:

- Para a primeira mensagem, **não** inclua `thread_id`.
- Para mensagens seguintes, defina `thread_id` com o valor retornado na resposta anterior

Resposta:

Content-Type: `text/event-stream`

Cada bloco de dados do evento:

```
{
  "response": "streaming response text",
  "thread_id": "686b5163-04a0-4687-bf86-618c3c0f59ab"
}
```

Variáveis de Ambiente

Variável	Descrição
<code>IAM_API_KEY</code>	Chave de API do IBM Cloud IAM
<code>WXO_INSTANCE_ID</code>	ID da instância do Watsonx Orchestrate

Notas

- O endpoint `/stream_response` usa **streaming** (SSE) para interagir com o agente WxO, processando respostas em tempo real.

- Certifique-se de que seu ID do agente seja válido e que sua chave IAM tenha permissões suficientes para acessar a instância.