

Desenvolvido por: Nathalia Trazzi e Guilherme Indicatte
Período: Agosto/2025
Atualizado e revisado por Nathalia Trazzi em Fevereiro/2026

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

- `/list_agents` Retorna uma lista de agentes disponíveis em uma determinada instância de Watsonx Orchestrate
- `/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).

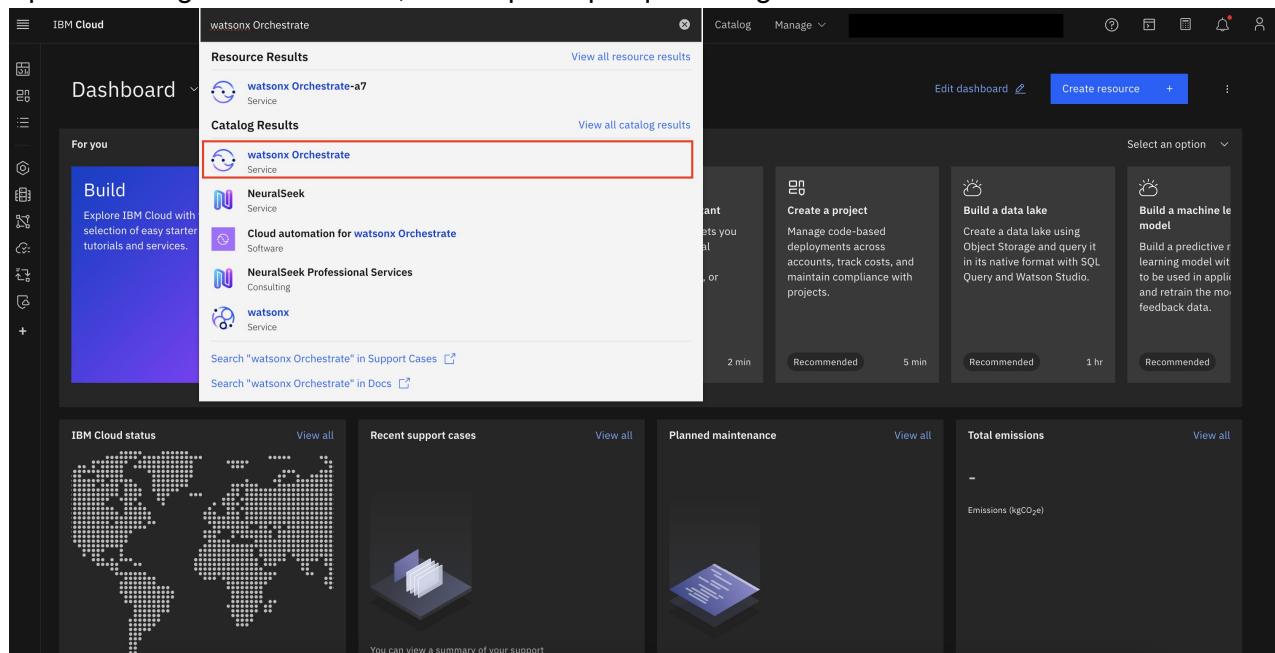
Para acessar o asset, acesse este link: https://github.com/nanatrazzi/watsonx-orchestrate-chat-with-agents/tree/main/source/rest_api

Endpoints implementados em: https://github.com/nanatrazzi/watsonx-orchestrate-chat-with-agents/blob/main/source/rest_api/routes

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. Após fazer login na IBM Cloud, no campo de pesquisa e digite **watsonx Orchestrate**.



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

Watsonx Orchestrate

Lighten your workload with IBM Watsonx Orchestrate, powered by LLMs. Streamline workflows and automate tasks with a catalog of skills. No specialized training needed.

Create **About**

Type: Service Provider: IBM Last updated: 08/26/2025 Category: AI / Machine Learning

Location: Sydney (au-syd), Toronto (ca-tor), Frankfurt (eu-de), London (eu-gb), Tokyo (jp-tok), Dallas (us-south), Washington DC (us-east)

Related links: Docs, Terms

Select a location: Sydney (au-syd)

Select a pricing plan: Prices shown are for country or location:

Plan	Features and capabilities	Pricing
Trial	Experience everything in the Standard plan *free for 30 days* Agentic AI builder and Assistant builder Unified chat experience Conversational Search (RAG) AI-guided actions Multi-channel experience Generative AI Import automations effortlessly with OpenAPI specifications	Free
30-day free access to the Standard plan		
Lite plan services are deleted after 30 days of inactivity.		

Essentials Plan Optimized for customers to go live quickly using Agentic AI to build Domain \$500.00 USD/Essentials Instance

I have read and agree to the following license agreements: [Terms](#)

Create **Add to estimate**

3. Por fim, digite o nome do serviço, opcionalmente tags e tags de dicionamento, por fim selecione um **resource group**

4. Aceite os termos da IBM Cloud ao lado direito da tela e clique em **Create**

Watsonx Orchestrate

and Development for Oracle HCM Access
\$100.00 USD/Domain Agent for Invoice Management with Coupa Access
\$100.00 USD/Domain Agent for Sourcing and Contract Management Access
\$100.00 USD/Domain Agent for Purchasing with Coupa Access
\$100.00 USD/Domain Agent for Supplier Management with Coupa Access
\$100.00 USD/Domain Agent for Sales Prospecting Access

Standard Agentic with Monthly Active Users

Optimized for customers seeking enterprise capabilities with domain agents and AI assistants
Includes 40,000 Monthly Active Users (MAUs) and each MAU is limited to 50 messages per month
Channels: (webchat, Voice, Slack, Teams)
Agent Builders
Documents for Agent knowledge are limited to 50 per Agent with a maximum file size of 25MB. The maximum number of files uploaded in one batch is limited to 20 and the total file size of all uploaded files in one batch must not exceed 50MB. In addition, the total number of documents that can be uploaded across all agents is limited to 100GB.
Includes IBM Pre-built Domain Agents

Configure your resource

Service name: watsonx Orchestrate-4n

Select a resource group: Default

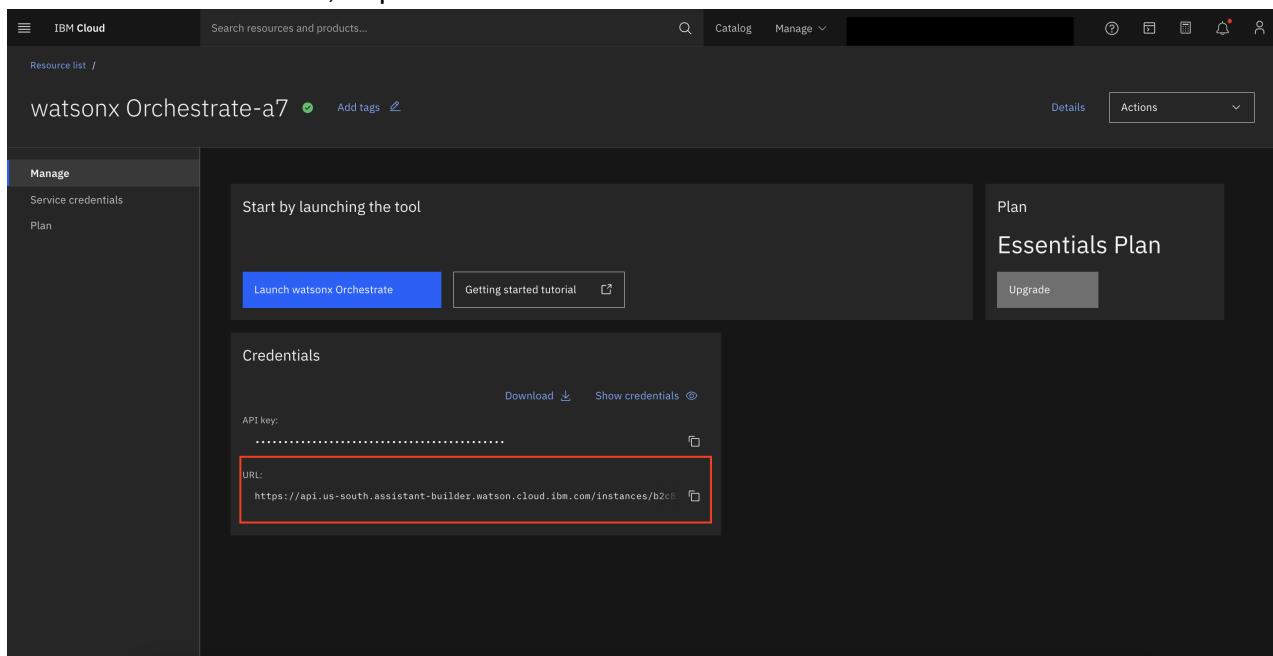
Tags: Examples: env:dev, version-1

Access management tags: Examples: access:dev, proj:version-1

I have read and agree to the following license agreements: [Terms](#)

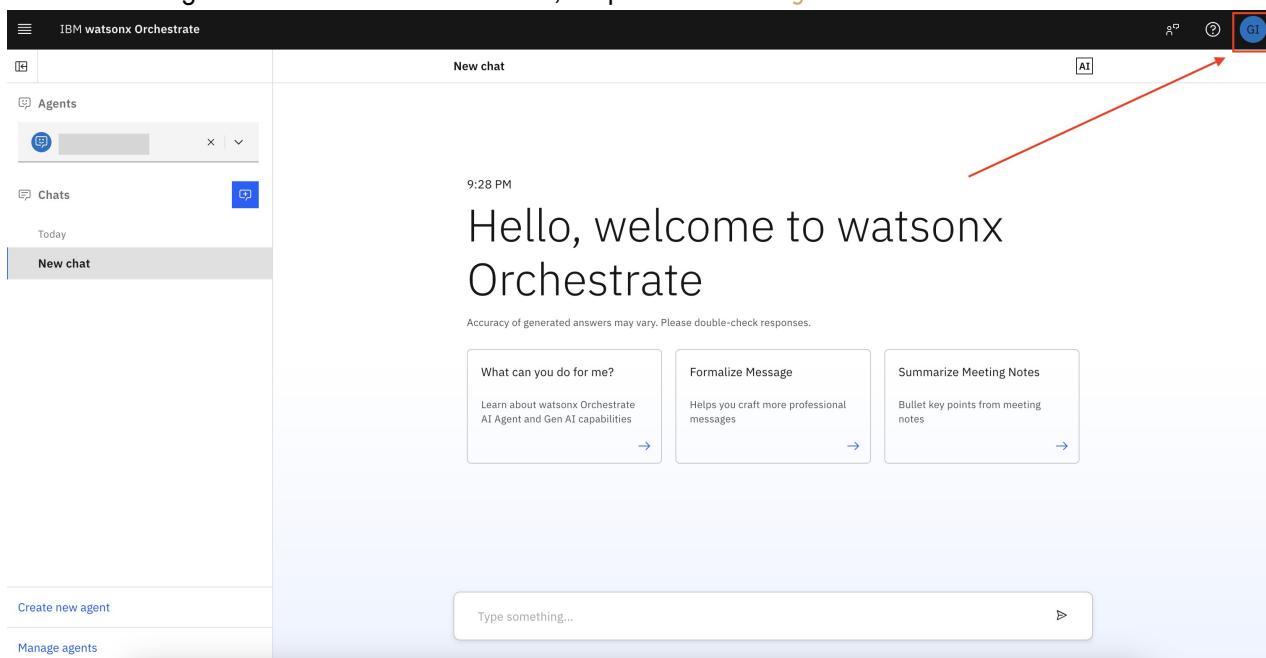
Create **Add to estimate**

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



The screenshot shows the IBM Cloud Resource list interface. A service instance named "watsonx Orchestrate-a7" is selected. On the right, there is a "Plan" section for the "Essentials Plan" with an "Upgrade" button. Below the plan, the "Credentials" section displays an API key and a URL. The URL is highlighted with a red box. On the left, a sidebar shows "Manage", "Service credentials", and "Plan". At the top, there is a search bar and navigation buttons for Catalog and Manage.

6. No ícone da sigla de seu nome e sobrenome, clique em **Settings**



The screenshot shows the IBM watsonx Orchestrate interface. On the left, there is a sidebar with "Agents", "Chats", "Today", and "New chat" buttons. The "New chat" button is highlighted with a red box. The main area shows a "New chat" window with a message from the AI agent: "Hello, welcome to watsonx Orchestrate". Below the message, there are three cards: "What can you do for me?", "Formalize Message", and "Summarize Meeting Notes". At the bottom, there is a text input field with "Type something..." placeholder text and a send button. On the far right, there is a profile picture of an AI agent with a red box and an arrow pointing to it.

7. 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}
```

Generate API keys

API keys are intended to generate the JWT token that you need to authenticate your user before you make requests to the API. Use the service instance URL to request access to the environment's resources.

Learn more about API keys [\[link\]](#)

Generate API key→

Service instance URL

https://api.us-south.watson-orchestrate.cloud.ibm.com/instances/ [REDACTED]

8. Também na mesma página, clique em **Generate API key**.

9. Você será redirecionado para o catálogo onde você cria, visualiza e trabalha com chaves de API da IBM Cloud

Search resources and products...

Catalog Manage 3122813 - IBM Po... ? 🌐 🌐 🌐 🌐 🌐 🌐

IAM

Overview Dashboard

Manage identities

Users Trusted profiles Service IDs

API keys

Identity providers

Manage access

Access groups

Authorizations

Roles

Gain insight

Settings

Documentation

Enterprise IAM Docs

API keys

View My IBM Cloud API keys

Filter by API key name or description

Create +

ID	Name	Description	Date created	Status	Enabled	Expires in
[REDACTED]	[REDACTED]	[REDACTED]	2-12-2026 21:02 GMT	🔓	<input checked="" type="checkbox"/> Yes	-
[REDACTED]	[REDACTED]	[REDACTED]	2-12-2026 21:41 GMT	🔓	<input checked="" type="checkbox"/> Yes	-
[REDACTED]	[REDACTED]	[REDACTED]	12-4-2025 17:58 GMT	🔓	<input checked="" type="checkbox"/> Yes	-

Items per page: 25 1-25 items Page 1

10. Uma **API Key** é um segredo que permite que usuários, aplicações ou serviços acessem a IBM Cloud de forma autenticada. Ela funciona como uma “senha técnica” usada em CLIs, automações, pipelines ou integrações.

Name: Dê um nome descritivo para a chave, como:

- apikey-wxo
- local-development-key

Dica: Use nomes que indiquem o propósito, responsável ou ambiente.

Description (Opcional): Aqui você pode adicionar uma descrição da sua chave, ajudando na identificação dela.

Lead Action

Disable the leaked key (recomendado)

A chave é **automaticamente desativada**. Boa opção para produção.

Delete the leaked key

A chave é deletada permanentemente.

Nothing

Nenhuma ação é tomada. *Não recomendado.*

API Key Expiration

API Key Expiration — Definir data de expiração?

Você pode escolher se a chave deve expirar automaticamente.

Set expiration

Permite escolher uma **data futura**.

Ideal para acessos temporários, integrações de curto prazo, etc.

Quando expirar, a chave é automaticamente desabilitada.

Sem expiração (default)

A chave não expira.

Session Management – Ativar login de sessão no CLI?

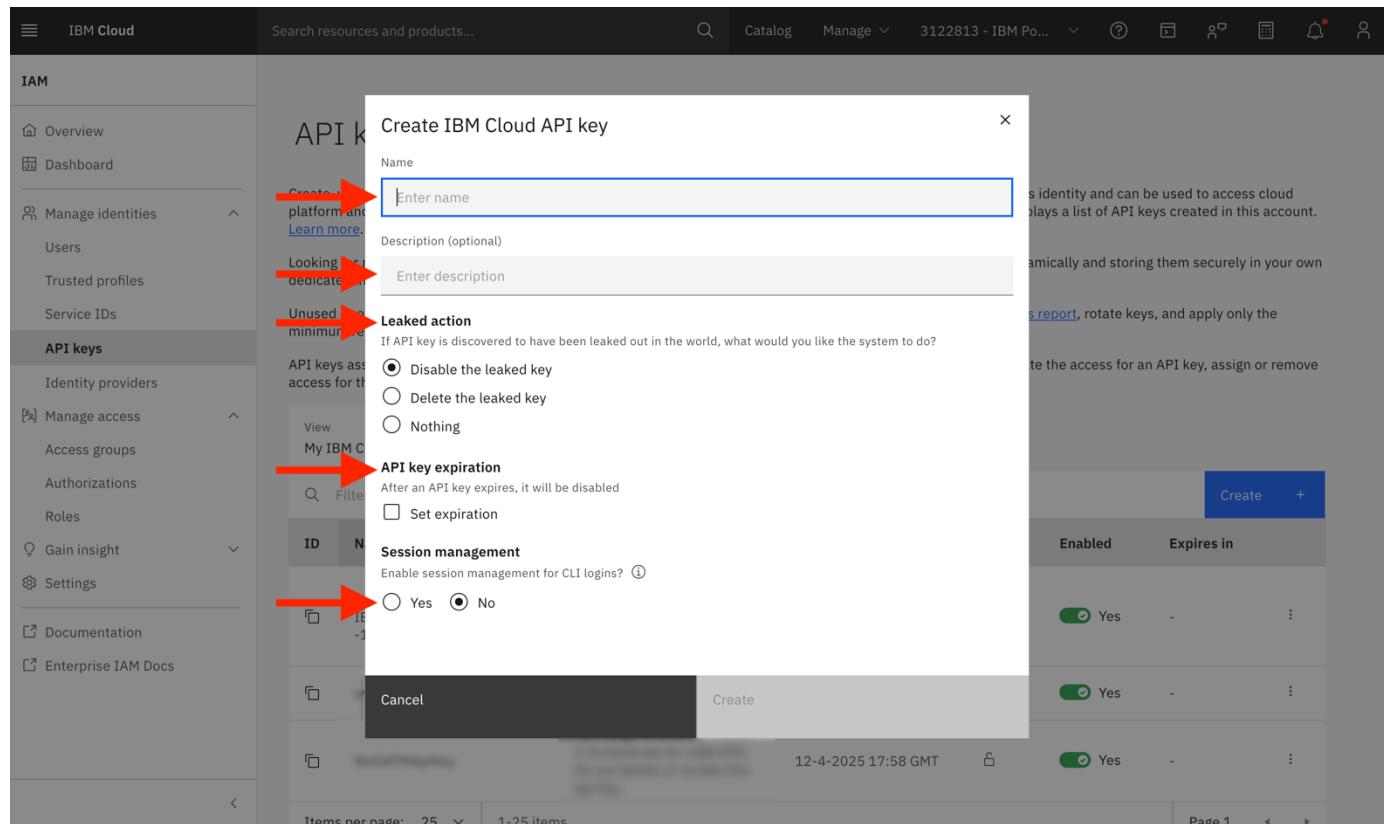
Define se esta chave pode ser usada para **gerenciamento de sessões** no IBM Cloud CLI.

Yes

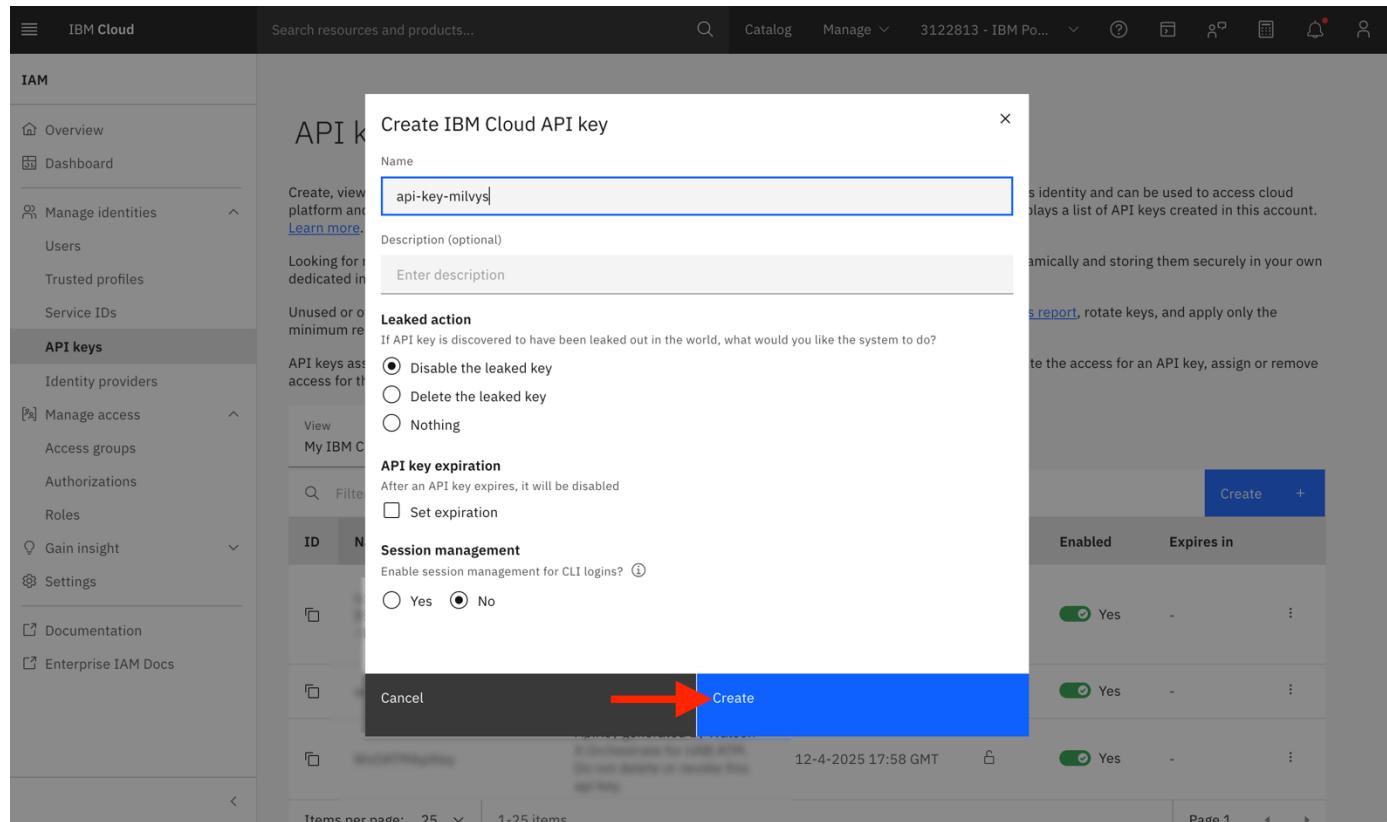
Permite que a chave seja usada para fazer login no CLI com `ibmcloud login --apikey`.

⚠️ No

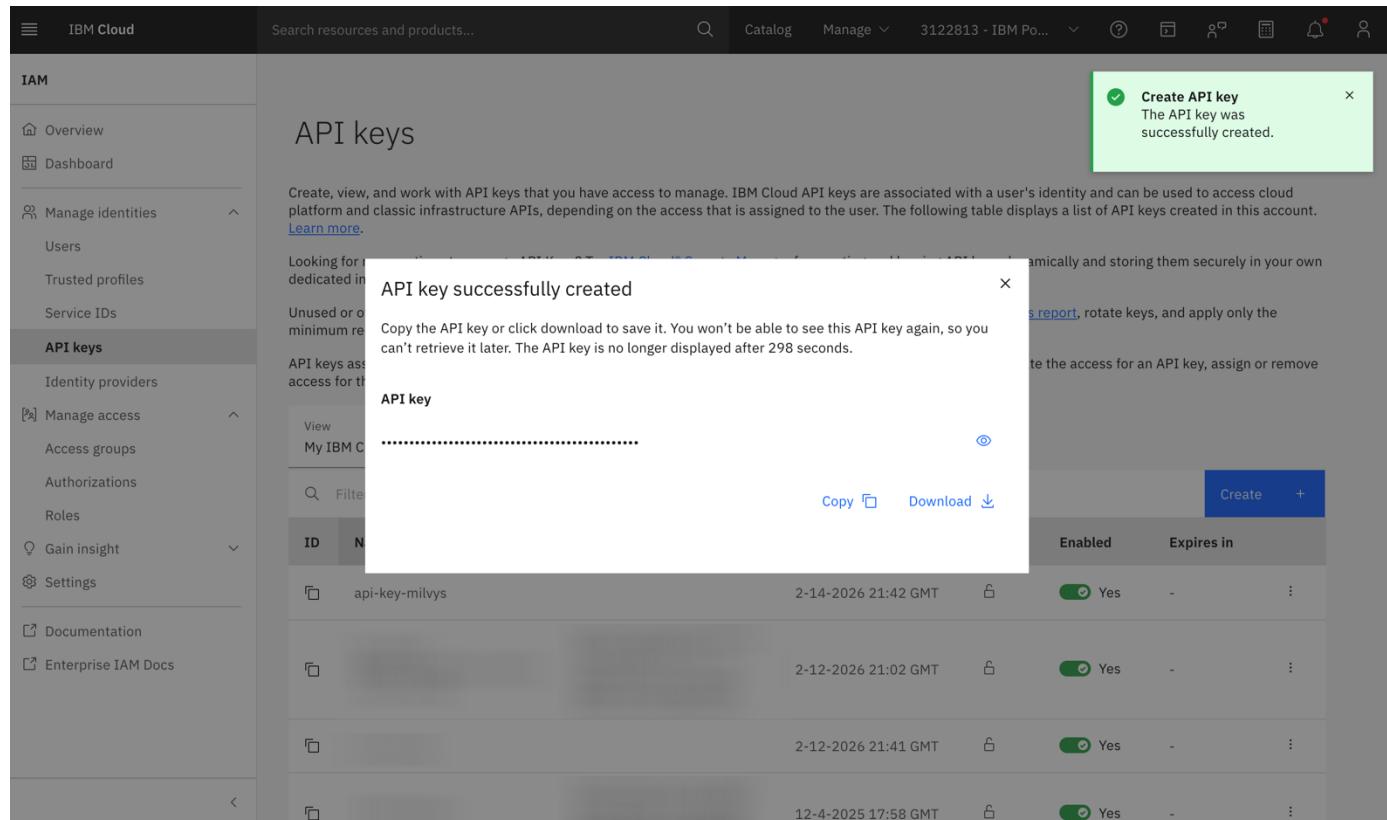
A API Key **não poderá ser usada para login** em sessões CLI. Use quando desejar restringir a chave apenas para chamadas específicas de API.



10. Após preencher tudo, clique em **Create**.



Aproveite esse momento para copiar ou fazer o download sua API Key. A chave ficará disponível por alguns segundos. **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).

Navegue até o diretório do projeto, no caminho ./source/rest_api/

Altere o nome do 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 seguinte comando em seu terminal

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

Sobre a API desenvolvida

Rotas desenvolvidas

GET `/list_agents`

Lista os agentes disponíveis em sua instância

Resposta:

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

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 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

- Certifique-se de que seu ID do agente seja válido e que sua chave IAM tenha permissões suficientes para acessar a instância.
- O endpoint `/stream_response` usa streaming (SSE) para interagir com o agente watsonx Orchestrate processando respostas em tempo real.

