

# Watsonx Orchestrate Chat with Agents

API REST para comunicação com IBM Watson Orchestrate



## Índice

- [Visão Geral](#)
- [Endpoints Disponíveis](#)
- [Como Obter as Credenciais](#)
- [Configuração do Projeto](#)
- [Documentação da API](#)
- [Variáveis de Ambiente](#)
- [Notas Importantes](#)

## Visão Geral

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

| Endpoint                      | Descrição  |
|-------------------------------|--|
| <code>/list_agents</code>     | Retorna uma lista de agentes disponíveis em uma determinada instância de watsonx Orchestrate |
| <code>/get_response</code>    | Retorna uma única resposta de um agente WxO  |
| <code>/stream_response</code> | Transmite a resposta do agente em tempo real usando Server-Sent Events (SSE)                 |

## Links Úteis

- **Código Fonte:** [GitHub Repository](#)
- **Endpoints Implementados:** [Routes Directory](#)

## Criar uma Instância do Watsonx Orchestrate

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

The screenshot shows the IBM Cloud Watsonx Orchestrate dashboard. At the top, there's a navigation bar with 'IBM Cloud' and 'watsonx Orchestrate'. Below it is a 'Dashboard' section with a 'For you' sidebar containing 'Build' tutorials and services like 'NeuralSeek' and 'Cloud automation for watsonx Orchestrate'. The main area shows 'Resource Results' and 'Catalog Results' sections, with 'watsonx Orchestrate-a7' and 'watsonx Orchestrate' highlighted. To the right, there are cards for 'Create a project', 'Build a data lake', and 'Build a machine learning model'. Below these are sections for 'IBM Cloud status', 'Recent support cases', 'Planned maintenance', and 'Total emissions'.

Escolha a **location** de sua instância e o **plano** que lhe atende melhor

The screenshot shows the 'Catalog' page for 'watsonx Orchestrate'. On the left, there are filters for 'Type', 'Provider', 'Last updated', and 'Category'. A red arrow points from the 'Last updated' field to the 'Select a pricing plan' dropdown, which is currently set to 'Sydney (au-syd)'. Another red arrow points to the 'Create' button. The right side shows a summary of the selected instance, including location 'Sydney (au-syd)', plan 'Trial', and price '\$500.00 USD/Essential Instance'. There are also terms and conditions checkboxes and a 'Create' button.

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

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

The screenshot shows the IBM Cloud Catalog interface. On the left, there's a sidebar with 'IBM Cloud' and a search bar. The main area displays a summary of the 'watsonx Orchestrate' service, which is free. To the right, there's a detailed configuration section with tabs for 'Summary', 'Watsonx Orchestrate', and 'Free'. The 'Watsonx Orchestrate' tab is active, showing details like location (Sydney), plan (Trial), and service name (watsonx Orchestrate-4n). Below this, there's a 'Configure your resource' section with fields for 'Service name' (set to 'watsonx Orchestrate-4n'), 'Select a resource group' (set to 'Default'), 'Tags' (examples: env:dev, version-1), and 'Access management tags' (examples: access:dev, projversion-1). A red box highlights this configuration area.

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

The screenshot shows the details page for the 'watsonx Orchestrate-a7' instance. On the left, there's a sidebar with 'Manage', 'Service credentials', and 'Plan'. The main area has a 'Start by launching the tool' section with a blue 'Launch watsonx Orchestrate' button (which has a red arrow pointing to it), a 'Getting started tutorial' link, and a 'Plan' section showing the 'Essentials Plan' with an 'Upgrade' button. Below this, there's a 'Credentials' section with an API key and URL.

Para se comunicar com o Watsonx Orchestrate via API, você precisará obter duas credenciais essenciais:

- IAM\_API\_KEY
- WXO\_INSTANCE\_ID

## Como obter as credenciais do watsonx Orchestrate

No ícone da sigla de seu nome e sobrenome ao lado direito da interface

The screenshot shows the IBM Watsonx Orchestrate web interface. At the top right, there is a blue circular icon with the letters 'AI' and a question mark icon. A red arrow points from the bottom right towards this icon. The main area displays a welcome message: "Hello, welcome to watsonx Orchestrate". Below the message are three cards: "What can you do for me?", "Formalize Message", and "Summarize Meeting Notes". The sidebar on the left includes sections for Agents (with an 'AskOrchestrate' entry), Chats, Today, and New chat (which is selected). At the bottom, there are links for "Need help creating an agent?", "Create new agent", and "Manage agents".

Clique em **Settings**.

The screenshot shows the same IBM Watsonx Orchestrate interface as above, but with a user profile dropdown menu open on the right side. The profile information includes a blue circular icon with 'NT', the name 'Nathalia Trazzi', and the email 'n.trazzi@ibm.com'. The dropdown menu lists several options: Region (br-sao), Plan (Trial), Owner, Instance name (Orchestrate-Fundacao) with a 'Switch' link, and a large red arrow pointing to the 'Settings' option. Other visible options include 'Provide feedback', 'Privacy', 'About', and 'Log out'.

Na página de configuração e entre na aba **API details**

The screenshot shows the 'Data Retention' tab selected in the navigation bar. Below it, a section titled 'Manage the data retention policy' contains a note about retaining chat history for a minimum of 30 days and a maximum of 365 days. A slider is set to 30 days. At the bottom are 'Cancel' and 'Save' buttons.

## 1. Copie o ID de sua instância

The screenshot shows the 'API details' tab selected. In the 'Generate API keys' section, there is a note about generating a JWT token for authentication. A link 'Learn more about API keys' is provided. Below is a 'Service instance URL' field containing a long URL, with a red arrow pointing to the copy icon next to it. A red circle labeled '1' is near the URL, and another red circle labeled '2' is near the 'Generate API key' button.

## 2. Em seguida, clique em Generate API Key

Na próxima página, clique em **Generate API key**.

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

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

## O que é uma API Key?

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.

## Configuração da API Key

### Name (Nome)

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 (Ação em Caso de Vazamento)

| Opção   | Descrição  |
|---|--|
| <input checked="" type="radio"/> <b>Disable the leaked key</b><br>(recomendado) | A chave é automaticamente desativada. Boa opção para produção. |
| <input checked="" type="radio"/> <b>Delete the leaked key</b>                   | A chave é deletada permanentemente.                            |
| <input checked="" type="radio"/> <b>Nothing</b>                                 | Nenhuma ação é tomada. Não recomendado.                        |

### API Key Expiration (Expiração)

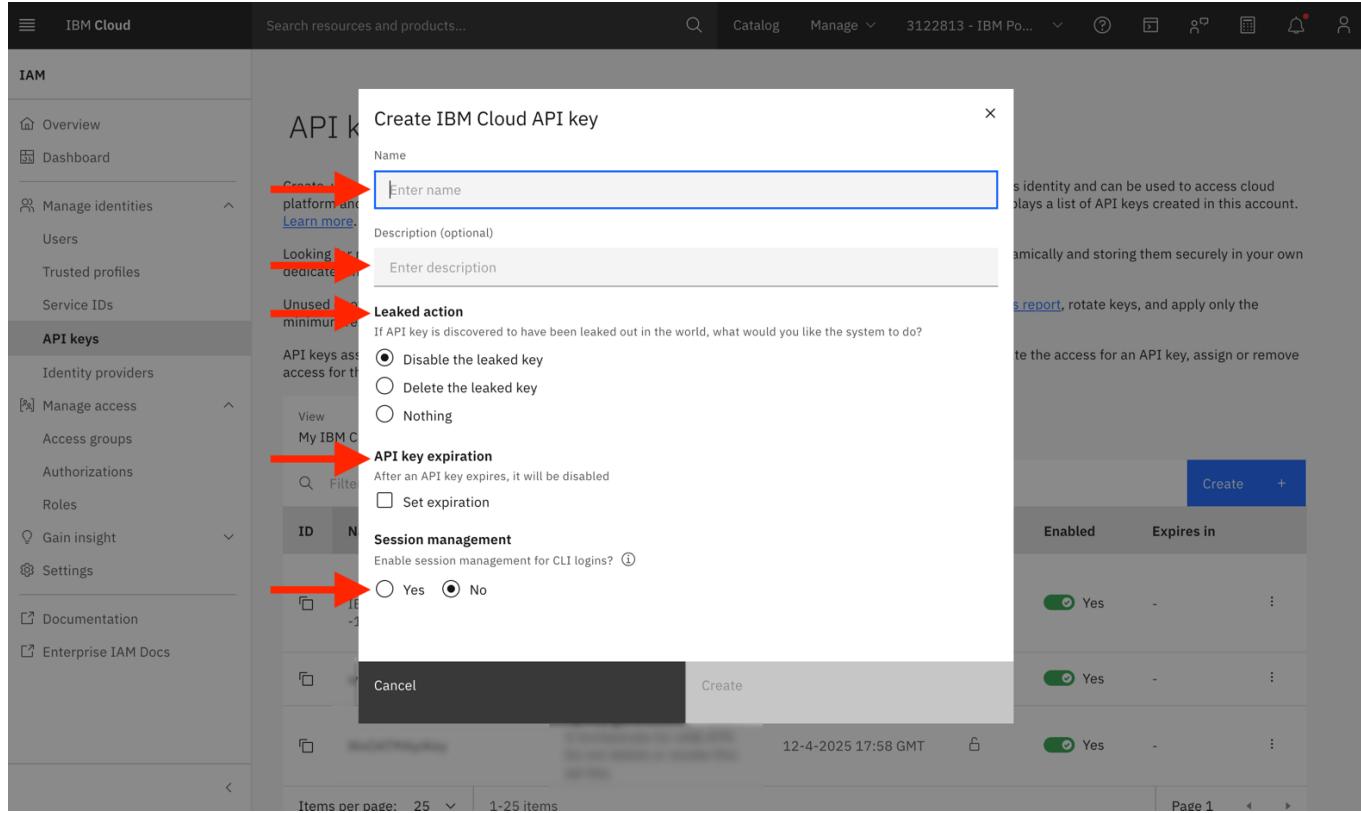
Você pode escolher se a chave deve expirar automaticamente.

| Opção   | Descrição   |
|---|---|
| <input checked="" type="radio"/> <b>Set expiration</b>          | Permite escolher uma data futura.<br>-> Ideal para acessos temporários, integrações de curto prazo, etc.<br>Quando expirar, a chave é automaticamente desabilitada. |
| <input checked="" type="radio"/> <b>Sem expiração</b> (default) | A chave não expira.   |

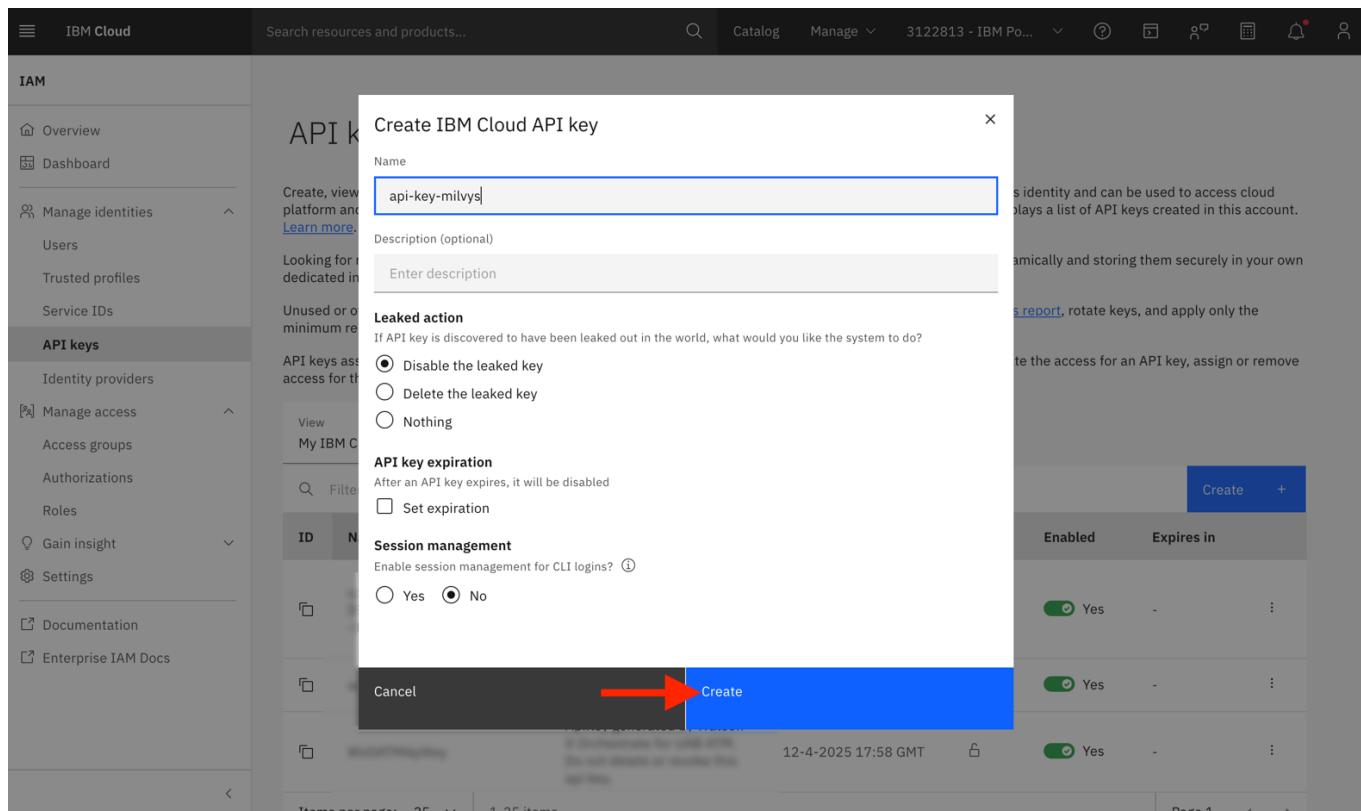
### Session Management (Gerenciamento de Sessão)

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

| Opção                                       | Descrição  |
|---|--|
| <input checked="" type="radio"/> <b>Yes</b> | Permite que a chave seja usada para fazer login no CLI com <code>ibmcloud login --apikey</code>  |
| <input checked="" type="radio"/> <b>No</b>  | A API Key não poderá ser usada para login em sessões CLI. Use quando desejar restringir a chave apenas para chamadas específicas de API. |



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



**IMPORTANTE:** 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.

The screenshot shows the IBM Cloud IAM API keys interface. On the left, there's a sidebar with options like Overview, Dashboard, Manage identities, API keys (which is selected), Identity providers, Manage access, Documentation, and Enterprise IAM Docs. The main area displays a table of API keys. A modal window titled "API key successfully created" is open, containing instructions to copy or download the API key, noting it will no longer be displayed after 298 seconds. The modal also has "Copy" and "Download" buttons. In the top right corner, a green notification box says "Create API key" and "The API key was successfully created." The table lists four API keys, each with columns for ID, Name, Created at, Enabled status (all are Yes), and Expires in (all are blank). The first API key is named "api-key-milvys".

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

### 1. Navegue até o diretório do projeto:

```
cd ./source/rest_api/
```

### 2. Configure as variáveis de ambiente:

- Altere o nome do arquivo `.env_example` para `.env`
- Preencha com suas credenciais:

```
IAM_API_KEY=sua-api-key-aqui
WX0_INSTANCE_ID=seu-instance-id-aqui
```

### 3. Instale as dependências:

```
pip install -r requirements.txt
```

### 4. Execute o servidor:

```
python app.py
```

## Gerenciamento de Conversas (Thread ID)

**Importante:** Para manter o contexto da conversa entre múltiplas mensagens:

- **Primeira mensagem:** 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`.
- **Mensagens subsequentes:** 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]
```

**Parâmetros:**

- `message` (obrigatório): A mensagem a ser enviada
- `agent_id` (obrigatório): ID do agente
- `thread_id` (opcional): ID da thread para continuar uma conversa

**Observações:**

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

**Resposta:**

```
{  
  "response": "Agent response here",  
  "thread_id": "thread-id-value"  
}
```

### 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]
```

## Parâmetros:

- **message** (obrigatório): A mensagem a ser enviada
- **agent\_id** (obrigatório): ID do agente
- **thread\_id** (opcional): ID da thread para continuar uma conversa

## 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

data: {"chunk": "Response", "thread_id": "thread-id-value"}

data: {"chunk": " chunk", "thread_id": "thread-id-value"}

data: {"chunk": " here", "thread_id": "thread-id-value"}
```

Cada bloco de dados do evento é enviado em tempo real conforme o agente processa a resposta.

---

## Variáveis de Ambiente

| Variável        | Descrição                              |
|-----------------|--|
| IAM_API_KEY     | Chave de API do IBM Cloud IAM          |
| WX0_INSTANCE_ID | ID da instância do Watsonx Orchestrate |

---

## Notas Importantes

### Atenção:

- Certifique-se de que seu **ID do agente** seja válido
- Sua **chave IAM** deve ter 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

---

Agora você está pronto para integrar seu aplicativo com o IBM Watson Orchestrate.

**Desenvolvido por Nathalia Trazzi**