

<epam>

Work with /chat/completion DIAL API

There are no magic! It is simple REST API!

Before we start:

- **Raise your hand and ask questions if you have any**
- **It is better to ask questions when you have**
- **Also, type them in chat**
- **We will need DIAL API key for this session**

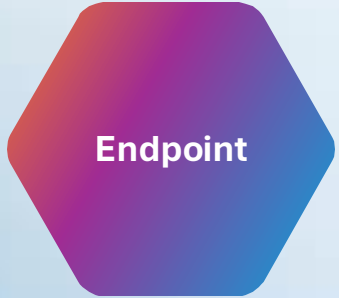
Agenda:

- **Presentation:**
 - **Request structure**
 - **Structure of the regular response**
 - **Structure of the streaming response**
- **Workshop:**
 - **Implement console chat**
 - **Implement DialClient with regular REST response handling**
 - **Implement DialClient with streaming response handling**

Request

<Request and its params>

Request:



<https://ai-proxy.lab.epam.com/openai/deployments/{model}/chat/completions>

Request:



Endpoint

Authorization

API-KEY: {YOUR_API_KEY}

**[https://support.epam.com/ess?id=sc_cat_item&table=sc_cat_item
&sys_id=910603f1c3789e907509583bb001310c](https://support.epam.com/ess?id=sc_cat_item&table=sc_cat_item&sys_id=910603f1c3789e907509583bb001310c)**

Request:

Endpoint

Authorization

Request body



Ensure chat history is included in each request, as context significantly affects the LLM's output.

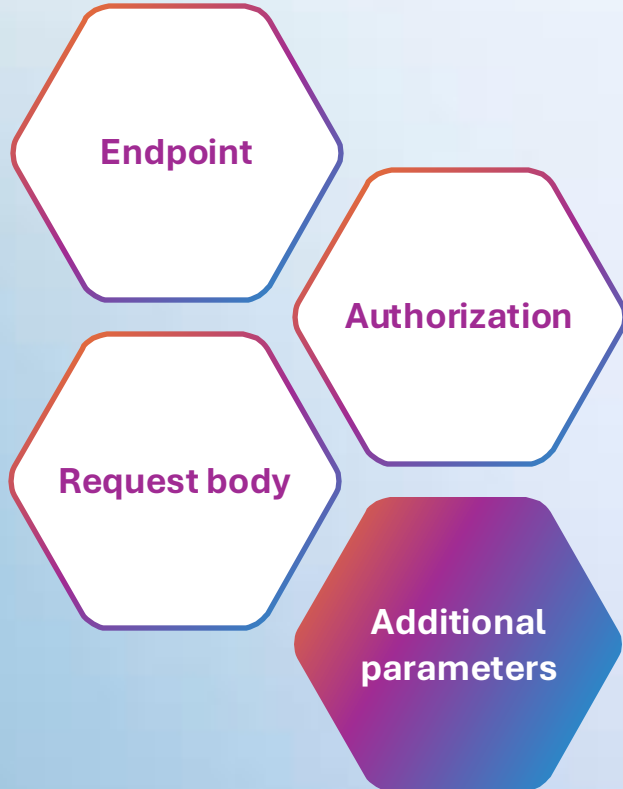
```
{  
  "messages": [  
    {  
      "role": "system",  
      "content": "You are a haiku writer."  
    },  
    {  
      "role": "user",  
      "content": "Write a haiku about recursion in  
programming."  
    }  
  ],  
  "stream": false  
}
```

System prompt

User prompt

To stream response or not
By default *false*

Request:



Enables response streaming, where the server sends back partial responses as they are generated. This is useful for applications needing real-time updates.

Default value: false

Response

<Response and its handling>

Response:

Default

Array with choices. Here will be presented messages and additional metadata.

Array size from 1 up to 10.
Default 1.

Choices are configured with parameter `n` in request

```
{
  "id": "chatcmpl-7vHVghsEnUygFKP9sBxX3TEBh4L3q",
  "object": "chat.completion",
  "created": 1678928387,
  "model": "gpt-4",
  "usage": {
    "prompt_tokens": 12,
    "completion_tokens": 24,
    "total_tokens": 36
  },
  "choices": [
    {
      "index": 0,
      "message": {
        "role": "assistant",
        "content": "Why don't scientists trust atoms? Because they make up everything!"
      },
      "finish_reason": "stop"
    }
  ]
}
```

Provides the information about used tokens. This info can be used to calculate the request cost

Message with role

Finish reason represent the information about why LLM finished to generate content. **stop** means that LLM is fully finished content generation

Response:

Default

Streaming
stream: true

Chunk 1

```
data: {
```

data: is the key word (6 chars) that
each streamed message started from

```
"choices": [
```

```
{
```

```
"delta": {
```

```
"role": "assistant",
```

```
"content": "Why don't scientists"
```

```
}
```

```
}
```

```
]
```

```
}
```

Response:

Default

Streaming
stream: true

Chunk 1

Chunk 2

```
data: {  
  "choices": [  
    {  
      "delta": {  
        "content": " trust atoms?"  
      }  
    }  
  ]  
}
```

Response:

Default

Streaming
stream: true

Chunk 1

Chunk 2

Chunk 3

```
data: {  
  "choices": [  
    {  
      "delta": {  
        "content": " Because they make up  
everything!"  
      }  
    }  
  ]  
}
```

Response:

Default

Streaming
stream: true

Final chunk

Chunk 1

Chunk 2

Chunk 3

data: [DONE]



Join us:



Subscribe to WeAreCommuntiy

<https://wearecommunity.io/communities/dialx>

Keep in touch with our latest updates.
Here you find webinars, workshops and
articles about DIALX features and products.

Subscribe to YouTube

<https://www.youtube.com/@TeamDIALX>

Here we publish videos about our newest
products and features.

Join our Discord community

<https://discord.gg/jvTCQv4E4q>

🌟 AI DIALX Community 🌟 is the place where
you can find help with your questions about
DIALX, direct communication with DIALX team
and contributors.

A satellite view of Earth at night, showing the curvature of the planet and glowing city lights against the dark background of space.

DIALX
COMMUNITY
POWERED BY <epam>

Thank you!