interweb

Gerado a 28/11/2026 20:38

BASE DE CONHECIMENTO · DOCUMENTO PÚBLICO

Enviar ferramentas (function calling) num pedido de chat

Usar a API de IA

CÓDIGO	VERSÃO	PAÍS · IDIOMA
RLX-AO-KB-3006	v2026.11.28	Angola · Português

Interweb Angola
RLX Angola, Lda. · Grupo RLX Holdings
NIF 500 260 5724 · Matrícula 26464-25/250724

https://www.meu.interweb.ao/knowledgebase/3006

Enviar ferramentas (function calling) num pedido de chat

Usar a API de IA · Criado em **28 de novembro de 2026** · 3 min de leitura



A resposta curta: as ferramentas (function calling) deixam o modelo pedir ao seu programa que corra uma função, por exemplo consultar uma encomenda, e responder com o resultado. O modelo não corre nada: só diz que função quer e com que valores. Quem a corre é o seu programa. No chat, `POST /v1/chat/completions`, as ferramentas vão no campo `tools`, no formato da OpenAI.

Como funciona, em três passos

1. O seu programa envia as mensagens e a lista de ferramentas, cada uma com o nome, a descrição e os parâmetros (em JSON Schema).

2. Se o modelo quiser usar uma, a resposta vem com `finish_reason` `tool_calls` e, em `message.tool_calls`, o `id` da chamada, o nome da função e os argumentos em `function.arguments` (um texto em JSON).

3. O seu programa corre a função e faz um segundo pedido com a conversa, a mensagem do modelo e uma mensagem com `role` `tool`, o `tool_call_id` e o resultado em `content`. O modelo responde então em texto.

O primeiro pedido

```
{
  "model": "deepseek-ai/deepseek-v4-flash",
  "messages": [{"role": "user", "content": "Em que estado está a encomenda 1042?"}],
  "tools": [{
    "type": "function",
    "function": {
      "name": "get_order_status",
      "description": "Returns the status of an order by its number.",
      "parameters": {
```

```
        "type": "object",
            "properties": {"order_id": {"type": "string", "description": "The order
number"}},
        "required": ["order_id"]
      }
    }
  }],
  "max_tokens": 300
}
```

A resposta e o segundo pedido

O modelo pede a função (excerto da resposta; o `id` é sempre o que vier na sua resposta):

```
"message": {
  "role": "assistant",
  "content": null,
  "tool_calls": [{"id": "call_1", "type": "function",
    "function": {"name": "get_order_status", "arguments": "{\"order_id\": \"1042\"}"}}]
},
"finish_reason": "tool_calls"
```

O segundo pedido leva o mesmo `model` e as mesmas `tools`, e estas mensagens:

```
"messages": [
  {"role": "user", "content": "Em que estado está a encomenda 1042?"},
    {"role": "assistant", "content": null, "tool_calls": [{"id": "call_1", "type":
"function",
          "function": {"name": "get_order_status", "arguments": "{\"order_id\":
\"1042\"}"}}]},
  {"role": "tool", "tool_call_id": "call_1", "content": "{\"status\": \"shipped\"}"}
]
```

Os outros campos da OpenAI, como `tool_choice`, passam tal como vêm.

Que modelo usar

Use um modelo que aceite ferramentas. Os modelos recomendados nas **ferramentas de programação** aceitam-nas todos, entre eles `moonshotai/kimi-k2.7-code`, `zai-org/glm-5.2` e `deepseek-ai/deepseek-v4-flash`. Antes de pôr no ar, teste o ciclo completo com o modelo que escolheu.

> **Confirme os argumentos antes de correr a função.**
> O modelo nem sempre gera JSON válido, e pode pedir valores que não fazem sentido. Valide-os, e nunca deixe o modelo decidir sozinho coisas como pagamentos ou apagar dados.

Custo, Sandbox e os outros formatos

Cada pedido paga os seus tokens: uma resposta com uma ferramenta faz dois pedidos ao modelo. Com uma chave `sandbox_`, o chat responde sempre «Resposta de teste (sandbox).», por isso não aparecem chamadas de ferramentas. Para testar o ciclo, use a chave `prod_` com um modelo barato e um `max_tokens` baixo. No formato da Anthropic, as chamadas vêm em blocos `tool_use` e respondem-se com `tool_result`; no Responses, em itens `function_call` e `function_call_output`.

RECOMENDADO

Quer escolher um modelo que aceite ferramentas? Veja o catálogo no All in One AI.

Ver o All in One AI →

VEJA TAMBÉM

- Aplicações feitas para o Claude: o formato Anthropic Messages em /v1/messages
- O formato OpenAI Responses: /v1/responses e respostas guardadas durante 30 dias
- Escolher um modelo de texto: preço, contexto e para que serve cada um
- Quanto custa uma conversa pela API: a reserva, o max_tokens e o acerto no fim
- All in One AI

interweb

Interweb Angola

RLX Angola, Lda. · Grupo RLX Holdings

NIF	500 260 5724
MATRÍCULA	26464-25/250724
SEDE	Rua Direita do Patriota, Bairro Dangereux Belas, Luanda, Angola
CONTACTOS	info@interweb.ao · +244 942 011 191 · www.interweb.ao
HORÁRIO	2ª. à 6ª. Feira - 08 à 16 Horas Sábado & Domingo - Fechado

Versão online deste documento

RLX-AO-KB-3006 · v2026.11.28 · Gerado a 28/11/2026 20:38