interweb

Gerado a 28/11/2026 22:10

BASE DE CONHECIMENTO · DOCUMENTO PÚBLICO

As opções do aio chat: --system, --max-tokens, o stdin e a conversa no terminal

O comando aio

CÓDIGO
RLX-MZ-KB-3084

VERSÃO
v2026.11.28

PAÍS · IDIOMA
Moçambique · Português

Interweb Moçambique
RLX Holdings, S.A. · Grupo RLX Holdings
NUIT 401 191 607 · NUEL 105 045 974 · CAE 62022

https://www.meu.interweb.co.mz/knowledgebase/3084

As opções do aio chat: --system, --max-tokens, o stdin e a conversa no terminal

O comando aio · Criado em **28 de novembro de 2026** · 2 min de leitura



O `aio chat` faz uma pergunta a um modelo de texto e mostra a resposta à medida que o modelo a escreve. Os primeiros passos estão em **a IA no terminal com o comando aio**. Aqui vemos as opções, uma a uma.

Três formas de perguntar

```
aio chat "Explica o que é um registo DNS do tipo MX"
aio chat "Escreve um slogan para uma padaria" --model deepseek-ai/deepseek-v4-flash --max-tokens 300
cat pedido.txt | aio chat
aio chat
```

Forma	O que acontece
Pedido entre aspas	O modelo responde e o terminal fica livre.
Texto pelo `|` (stdin)	Sem pedido entre aspas, o texto que chega pela entrada padrão é o pedido. Serve para um ficheiro ou para o resultado de outro comando.
Sem pedido nem stdin	Abre uma conversa no terminal. `/reset` começa de novo, e `/exit` ou `Ctrl+C` sai.

As opções

Opção	O que faz
`--model <id>`	Escolhe o modelo. Sem ela, vale a variável `AIO_MODEL` ou o modelo por omissão, que aparece em `aio help chat`.

Opção	O que faz
`--system "..."`	Dá instruções ao modelo, por exemplo `--system "Responde em português"`.
`--max-tokens`	Limita o tamanho da resposta.
`--no-stream`	Recebe a resposta toda de uma vez, em vez de aos poucos.
`--json`	A resposta em JSON, para usar noutro programa.

Exemplos do dia-a-dia

```
cat erro.log | aio chat --system "Explica este erro em duas frases"
cat contrato.txt | aio chat --system "Resume em cinco pontos" --max-tokens 400
```

O `--system` fixa o tom e o formato. O texto do ficheiro chega pelo `|` e é o pedido.

O --max-tokens e o saldo

Antes de chamar o modelo, reservamos créditos para a entrada e para o máximo de saída pedido. No fim, cobramos os tokens reais e devolvemos o resto. Com pouco saldo, um `--max-tokens` mais baixo ajuda, porque a reserva fica menor. No fim de cada resposta, o `aio` mostra os tokens de entrada e de saída.

Com um Coding Plan, os modelos do plano gastam pontos em vez de créditos. A conta toda está em **quanto custa um pedido de chat pela API**.

Com uma chave `sandbox_`, o `aio chat` responde sempre «Resposta de teste (sandbox).» e não gasta nada. Serve para testar um script antes de usar a chave `prod_`.

RECOMENDADO

Quer conversar com os modelos de IA no terminal? Veja o All in One AI.

Ver o All in One AI →

VEJA TAMBÉM

A IA no terminal com o comando aio: instalar, entrar e conversar →

Quanto custa um pedido de chat pela API: a reserva, o max_tokens e o acerto final →

Ver modelos e preços no terminal: aio models e aio model

Escolher um modelo de texto: preço, contexto e para que serve cada um

All in One AI

interweb

Interweb Moçambique

RLX Holdings, S.A. · Grupo RLX Holdings

NUIT	401 191 607
NUEL	105 045 974
CAE	62022
SEDE	Bairro 7 de Setembro Vilankulo 1304 · CEP 0411-02, Inhambane, Moçambique
CONTACTOS	info@interweb.co.mz · +258 84 635 1111 · +258 86 635 1111 · www.interweb.co.mz
HORÁRIO	2ª. à 6ª. Feira - 08 à 16 Horas Sábado & Domingo - Fechado

Versão online deste documento

RLX-MZ-KB-3084 · v2026.11.28 · Gerado a 28/11/2026 22:10