interweb

Generated on 28/11/2026 21:59

KNOWLEDGE BASE · PUBLIC DOCUMENT

Seeing what a message will cost before you send it in the Playground

Playground & Text Models

CODE	VERSION	COUNTRY · LANGUAGE
RLX-MZ-KB-2992	v2026.11.28	Mozambique · English

Interweb Mozambique
RLX Holdings, S.A. · Group RLX Holdings
Tax ID (NUIT) 401 191 607 · Legal entity no. (NUEL) 105 045 974 · Economic activity code (CAE) 62022

https://www.meu.interweb.co.mz/knowledgebase/2992

Seeing what a message will cost before you send it in the Playground

Playground & Text Models Created on **November 28, 2026** 2 min read



In the Playground, under your message, the cost shows before you click **Send**. It changes as you type, and when you change the model or the maximum tokens. After the answer, it shows the real cost and your balance.

How it is worked out

Part	How it is counted
Input	About one token per 3 characters, from your message and the conversation so far.
Output	The "Maximum tokens in the reply", or 4,096 if you leave it empty (No limit). Never more than the model's maximum output.
Price	The model's, shown under it: "Price: X input and Y output, per 1 million tokens."

The messages you see

Message	When
"Maximum cost of this request: ..."	You typed a maximum number of tokens. It is the ceiling of what you can pay.
"Estimated cost of this request: ... no limit"	You left the maximum empty. The sum uses 4,096 output tokens.
"Real cost: X. Balance: Y."	After the answer: what you actually paid and the balance left.

In the first two cases, at the end you only pay for what the model used. The real cost is usually below the estimate.

To bring the cost down

1. Type a "Maximum tokens in the reply" that fits what you need.
2. Start a **New conversation** when the subject changes: the old conversation no longer weighs on the input.
3. Try a cheaper model.
4. To see the screen working without spending, choose **Sandbox (free)**: the answer is a test one and the model is not called.

Compare, images, video and voice

In **Compare models**, the cost is added up: "Estimated cost of the N requests, added up". In Image, Video and Voice you see the "Expected cost", which is reserved when generating: at the end you pay the real cost and get the difference back.

! If your balance is not enough, you see "This request holds X credits and your account has Y. Buy credits and try again." Nothing is charged. See **the 402 insufficient_credits error**.

RECOMMENDED

See the cost before you send in the Playground.

See All in One AI →

SEE ALSO

- **Chatting with a text model in the Playground before you use the API**
- **What a chat request costs through the API: the reservation, max_tokens and the final adjustment**
- **Cheap or strong text model: comparing the price per million tokens**
- **The 402 insufficient_credits error: what required and balance mean, and how to fix it**
- **All in One AI**

interweb

Interweb Mozambique

RLX Holdings, S.A. · Group RLX Holdings

TAX ID (NUIT)	401 191 607
LEGAL ENTITY NO. (NUEL)	105 045 974
ECONOMIC ACTIVITY CODE (CAE)	62022
REGISTERED OFFICE	Bairro 7 de Setembro Vilankulo 1304 · CEP 0411-02, Inhambane, Mozambique
CONTACT	info@interweb.co.mz · +258 84 635 1111 · +258 86 635 1111 · www.interweb.co.mz
OPENING HOURS	Monday to Friday - 08 AM to 04 PM Saturday & Sunday - Closed

Online version of this document

RLX-MZ-KB-2992 · v2026.11.28 · Generated on 28/11/2026 21:59