Mozout

Generated on 29/11/2026 04:04

KNOWLEDGE BASE · PUBLIC DOCUMENT

AI API errors: 402, 403, 404, 422, 429 and 502, and what to do in each case

Using the AI API

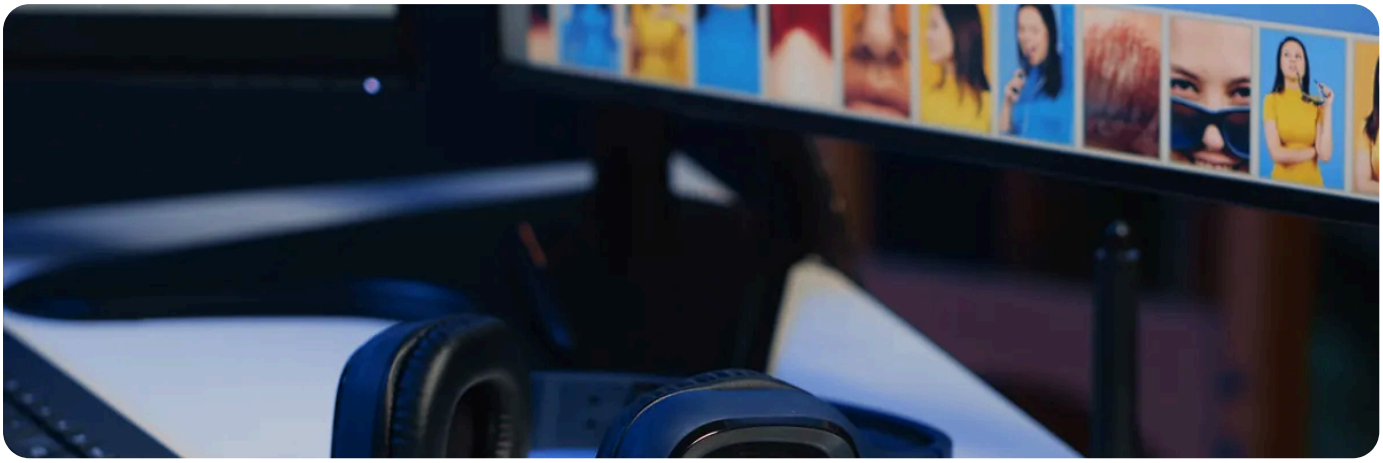
CODE	VERSION	COUNTRY · LANGUAGE
RLX-INT-KB-9108	v2026.11.28	International · English

MozOut International
RLX Holdings (Pty) Ltd. · Group RLX Holdings
Tax no. (SARS) 943 907 6234 · Registration no. (CIPC) 2025/667953/07

https://www.my.mozout.com/knowledgebase/9108

AI API errors: 402, 403, 404, 422, 429 and 502, and what to do in each case

Using the AI API Created on **November 28, 2026** 3 min read



In short: when an AI request fails, the answer has an HTTP status and, inside `error`, a stable code that says what happened. That code is what your program should read. Some errors go away if you wait; the others only go away when you fix something.

The error format

```
{
  "error": {
    "code": "insufficient_credits",
    "message": "Not enough credits for this request. Nothing was charged.",
    "correlation_id": "0f8fad5b-d9cb-469f-a165-70867728950e",
    "details": {"required": 46.12, "balance": 12.5}
  }
}
```

Field	What it is
`code`	A stable code, in English. Use this field in your programs.
`message`	Text for people, in the language of the `Accept-Language` header. It can change: do not compare against it.
`correlation_id`	The request identifier. It is the same as the `X-Correlation-ID` header.
`details`	Extra data about the error. It may be empty.

The errors, one by one

Status and code	When it happens	What to do
`402 insufficient_credits`	The credits do not cover the reservation. `details.required` says how much is needed and `details.balance` how much you have. Nothing was charged.	**Buy credits**, lower `max_tokens` or choose a cheaper model.
`403 insufficient_scope`	The key does not have `ai:read` or `ai:write`.	Create a key with both permissions.
`403 account_suspended` and other `account_*`	The All in One AI account is not active.	Open All in One AI and check the account status.
`403 ip_not_allowed`	The request came from an IP that is not on the key's list.	Use an IP on the list or create another key.
`404 model_not_found`	The model does not exist, is not of the right type (an image model in chat, for example) or is an agent of another account.	Check the `id` in `GET /v1/models`.
`404 not_found`	The generation does not exist or belongs to another account.	Check the `id` returned by the request.
`422 validation_failed`	A field is missing or a value is not one of the options.	See `details` and the fields in `GET /v1/models/{id}`.
`429 rate_limited`	You went over the limit of requests per minute.	Wait the seconds in the `Retry-After` header.
`429 plan_limit_reached`	The Coding Plan points ran out and the request cannot be paid with credits.	Wait for `details.resets_at` or turn on «Use credits when the plan runs out» on the **Plan** tab.
`502 upstream_error`	The AI service did not respond as expected.	Try again in a little while, or choose another model.

There are two more you may see: `401 unauthorized`, when the key is missing, wrong, revoked or expired, and `409 agent_paused`, when the agent you put in `model` is paused.

Retry or not

Worth retrying, after a wait: `429`, `500`, `502` and `503`. The others do not fix themselves: repeating the same request gives the same error. On a `POST`, send the `Idempotency-Key` header from the first attempt and retry with the same value. See **how not to pay twice for a generation**.

In the Anthropic and Responses formats

In `POST /v1/messages`, errors come in the Anthropic format: `error.type` follows the HTTP status (for example `billing_error`, `permission_error`, `not_found_error` or `rate_limit_error`), the message is the same and the identifier comes in `request_id`. In `POST /v1/responses`, they come in the OpenAI format, with the same `code`.

Asking for help

If you do not understand an error, **open a support ticket** with the `correlation_id`, the time of the request and the operation. Never send the key: support does not need it.

The full list of codes for the whole API is in **our API's error codes**. The ones that only apply to AI are in the **AI docs**.

RECOMMENDED

Not enough credits for the reservation? Buy them in All in One AI and send the request again.

See All in One AI →

SEE ALSO

- Our API's error codes: what each one means, and which ones are worth retrying →
- The request limit per key: 429 rate_limited and the Retry-After header →
- Errors in the OpenAI SDK: reading error.message and the HTTP status →
- An AI image failed: reading the error and getting the reserved credits back →
- All in One AI →

Mozout

MozOut International

RLX Holdings (Pty) Ltd. · Group RLX Holdings

TAX NO. (SARS)	943 907 6234
REGISTRATION NO. (CIPC)	2025/667953/07
REGISTERED OFFICE	1111 Olive Street Saint Louis 63101, Missouri, United States
CONTACT	info@mozout.com · +27 78 823 3325 · www.mozout.com
OPENING HOURS	Monday to Friday - 08 AM to 04 PM Saturday & Sunday - Closed

Online version of this document

RLX-INT-KB-9108 · v2026.11.28 · Generated on 29/11/2026 04:04