interweb

Generated on 27/11/2026 09:56

KNOWLEDGE BASE · PUBLIC DOCUMENT

Cursor, Codex CLI, Cline and other coding tools: the right address and model in each

All in One AI

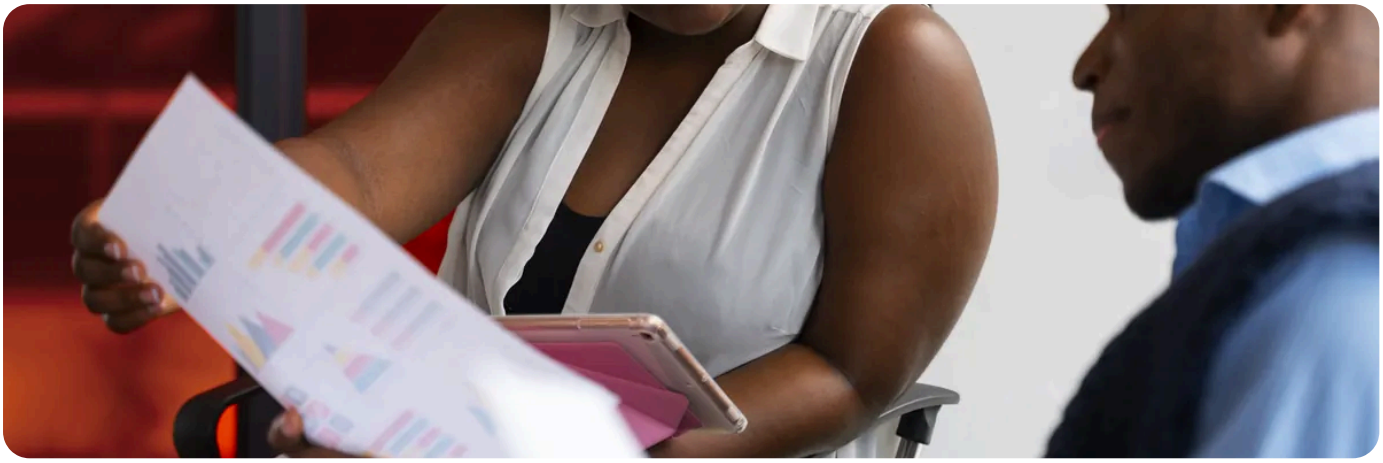
CODE	VERSION	COUNTRY · LANGUAGE
RLX-MZ-KB-2896	v2026.11.26	Mozambique · English

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/2896

Cursor, Codex CLI, Cline and other coding tools: the right address and model in each

All in One AI Created on **November 26, 2026** 3 min read



Each coding tool speaks one format, and the address changes with it. The most common mistake is an extra or missing `/v1`. The key is always the same, with `ai:read` and `ai:write`, and you pay with AI credits or the Coding Plan. Claude Code has an article of its own: **using Claude Code with All in One AI**.

Format	Base address	Tools
Anthropic (Messages)	`https://api.interweb.co.mz`	Claude Code
OpenAI (Responses)	`https://api.interweb.co.mz/v1`	Codex CLI
OpenAI (Chat Completions)	`https://api.interweb.co.mz/v1`	Cursor, Cline, Roo Code, OpenCode and Kilo Code

Recommended models

They all support tools (function calling) and are part of the Coding Plan. Use the id exactly as it is, with the slash:

Model	Context	What for
`moonshotai/kimi-k2.7-code`	262,144 tokens	Built for coding: the best starting point for everyday work.
`deepseek-ai/deepseek-v4-pro`	1,048,576 tokens	Strong reasoning, for big changes and hard bugs.
`zai-org/glm-5.2`	1,048,576 tokens	Made for agents: handles tools well in long tasks.

Model	Context	What for
`qwen/qwen3.8-max`	1,000,000 tokens	A large model, for reasoning and whole-project code.
`deepseek-ai/deepseek-v4-flash`	1,048,576 tokens	Fast and cheap, for background tasks and short questions.
`minimaxai/minimax-m2.5`	196,608 tokens	Light and cheap, good for simple edits.

Codex CLI

Create a provider in `~/.codex/config.toml`:

```
model = "moonshotai/kimi-k2.7-code"
model_provider = "api"

[model_providers.api]
name = "Interweb"
base_url = "https://api.interweb.co.mz/v1"
env_key = "INTERWEB_API_KEY"
wire_api = "responses"
```

`env_key` is the name of the variable that holds the key. `wire_api = "responses"` is required. The provider id can have another name, but not `openai`, `ollama` or `lmstudio`. The provider is only set in the user file: a project file cannot change it. For another model in a single session: `codex -m zai-org/glm-5.2`.

Cursor

1. Open **Cursor Settings** and then **Models**.
2. Under **API Keys**, turn on **OpenAI API Key** and paste your `prod_` key.
3. Turn on **Override OpenAI Base URL** and type `https://api.interweb.co.mz/v1`.
4. Under **Add Custom Model**, add the model id, for example `moonshotai/kimi-k2.7-code`, and pick it in the selector.

Requests go through Cursor's servers before they reach our API. According to Cursor's documentation, your own keys only work for chat models: Tab keeps using Cursor's models. While the address is overridden, requests to OpenAI models in Cursor also come to us. Turn off **Override OpenAI Base URL** to go back to them.

Cline and Roo Code

In both, choose **OpenAI Compatible** under **API Provider**, put `https://api.interweb.co.mz/v1` in **Base URL**, the key in **API Key** and the model id in **Model ID** (**Model** in Roo Code). Under **Model Configuration**, enter the model's context and maximum output. Roo Code only works with models that support tools.

OpenCode and Kilo Code

Both add an *OpenAI compatible* provider in a configuration file, with the address `https://api.interweb.co.mz/v1` and the key in an environment variable. In the Kilo Code extension for VS Code, the model list comes from `GET /v1/models`. The full files are on the **coding tools page of the docs**.

Never put the key in a file that goes into the repository.
Use the environment variable. If a key ends up in Git, revoke it and see **what to do, in the right order**.

Coding agents send many requests in a row. If you get `429 rate_limited`, wait the seconds in `Retry-After`: the tools usually retry on their own. To test the setup without spending, use a `sandbox_` key.

RECOMMENDED

Want to use your favourite tool with our models? Visit the Coding with AI page.

See Coding with AI →

SEE ALSO

- **Using Claude Code with All in One AI: the variables, the model and common errors**
- **Coding Plan: which plan to choose, what it costs and how to subscribe**
- **The All in One AI MCP server: images, video and voice inside Claude Code, Codex and ChatGPT**
- **An API key or token has leaked: what to do, in the right order**
- **Coding with AI in 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-2896 · v2026.11.26 · Generated on 27/11/2026 09:56