

Один API-ключ к 700+ моделям

OpenAI-совместимый формат для текста, задания для медиа, бюджет в рублях на каждый ключ.

[/v1/chat/completions](#)[Idempotency-Key](#)[max_cost_microrub](#)

```
import OpenAI from "openai";

const client = new OpenAI({
  baseUrl: "https://api.kvantora.ai/v1",
  apiKey: process.env.KVANTORA_API_KEY,
});

const reply = await client.chat.completions.create({
  model: "anthropic/claude-sonnet-5.5",
  messages: [{ role: "user", content: "Привет!" }],
});
```

Тот же запрос из любого языка

Модель задаётся публичным ID из каталога. Он одинаков в карточке модели, API и Нейростудии.

Текст: /v1/chat/completions,
/v1/responses, /v1/messages

Векторы: /v1/embeddings

Изображения, видео, аудио: задание, статус, файл

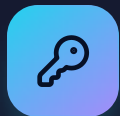
```
curl https://api.kvantora.ai/v1/chat/completions \
  -H "Authorization: Bearer $KVANTORA_API_KEY" \
  -H "Idempotency-Key: $REQUEST_ID" \
  -H "Content-Type: application/json" \
  -d '{"model": "xai/grok-4.7",
      "messages": [{"role": "user", "content": "Привет!"}]}'
```

Что совместимо, а что нет

Клиент и операция	Путь	Возможности и границы
OpenAI SDK · Chat	/v1/chat/completions	Текст, function tools, text и json_object, SSE. Без vision, hosted tools и json_schema.
OpenAI SDK · Responses	/v1/responses	Только текст, stream=false и store=false. Без серверной истории.
Anthropic SDK · Messages	/v1/messages	Текст, обязательный max_tokens. Без потока, tool_use и cache_control.
OpenAI SDK · Embeddings	/v1/embeddings	Строки, float и base64. Число входов и dimensions ограничены.
Медиа · JSON Kvantora	задания API	Оценка, задание, статус, файл. Полная совместимость медиа с OpenAI SDK не заявлена.

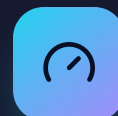
Полная матрица с версиями клиентов: docs.kvantora.ai/compatibility

Ключи, права и бюджет



Права на операции

Ключу выдаётся только нужное: `models:read` для проверки, `text` для генерации.



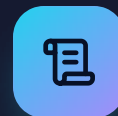
Дневной бюджет

Лимит в рублях учитывает открытые резервы. Исчерпан лимит, и платные запросы по ключу не проходят.



Секрет на сервере

Полный ключ показывается один раз. В браузер и репозиторий он не попадает.



Журнал

Каждый запрос со статусом и суммой виден в истории и через `GET /v1/usage`.

Стоимость известна до запроса

Оценка не вызывает модель и ничего не резервирует

max_cost_microrub отсекает запрос дороже предела:
HTTP 409

В ответе генерации есть usage и billing.status

```
curl https://api.kvantora.ai/v1/chat/quote \
  -H "Authorization: Bearer $KVANTORA_API_KEY" \
  -H "Content-Type: application/json" \
  -d '{"model": "deepseek/deepseek-v4-flash",
      "messages": [{"role": "user", "content": "..."}],
      "max_tokens": 800}'

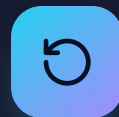
# estimated_cost_microrub, expires_at,
# reservation_created: false
```

Оборвалась сеть? Второго платного запроса не будет



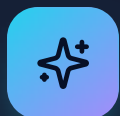
Idempotency-Key

Один ID на логический запрос. Новый ID означает новый платный запрос.



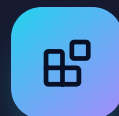
Автовозврат

Модель не вернула результат? Резерв вернётся на баланс через 5 минут, для медиа до 3 часов.



Статус вместо догадок

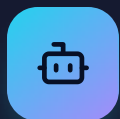
Таймаут не доказывает отмену. Сначала проверьте статус, потом решайте про повтор.



Понятные ошибки

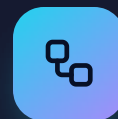
code, message и request_id в каждом ответе с ошибкой.

Кроме REST



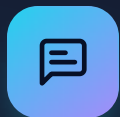
МСР-сервер

mcp.kvantora.ai/mcp: модели, автоматизации и ИИ-сотрудники для ваших агентов.



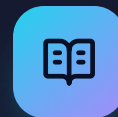
Kvantora Flow

Вебхуки, HTTP-запросы и шаги с моделями в визуальном редакторе.



Нейростудия

Отладьте промпт в браузере и перенесите в код с тем же ID модели.



Документация

docs.kvantora.ai: справочник API, совместимость, ошибки и повторы.

Первый ответ без платной генерации

1

Создайте ключ

Для проверки хватит права `models:read`. Секрет показывается один раз.

2

Сохраните секрет

`KVANTORA_API_KEY` в окружении сервера, не в браузере и не в репозитории.

3

Запросите `/v1/models`

HTTP 200 с `data` и `pagination` подтверждает доступ. Деньги не списываются.

Документация открыта без регистрации

Посмотрите примеры на Python, JavaScript и cURL, затем выпустите ключ с лимитом.

Документация

docs.kvantora.ai

Разработчикам

kvantora.ai/developers

API-ключ

kvantora.ai/register