

» PROMT

• AI Gateway Platform

Один кабинет — все AI-модели мира.

Прозрачный рабочий РФ-шлюз к 13 провайдерам и 425+ моделям. Общий баланс в рублях, биллинг по токенам, аналитика и история запросов — в одном интерфейсе.

• GPT-5

• Claude Opus 4.5

• Gemini 2.5 Pro

• DeepSeek R1

+ 421 модель

Каждый провайдер — это отдельная интеграция.

Команды подключают модели по одной — и тонут в ключах, биллингах и SDK.

13

разных API, форматов запроса и схем ответа на каждый сервис

× Ключи разбросаны по .env и головам разработчиков

∞

биллингов, карт и валют — расход невозможно свести в одном месте

× Нет fallback — упал провайдер, встала фича

0

общей аналитики, истории и контроля лимитов по команде

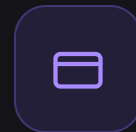
× Перерасход замечают только в конце месяца

PROMT — один шлюз между вами и любой моделью.



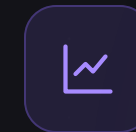
Единый API

Один формат запроса для 425+ моделей. Адаптер сам нормализует ответ, токены и стоимость.



Общий баланс в ₽

Один счёт в рублях на всех провайдеров. Прозрачное списание по токенам с hold и возвратом.

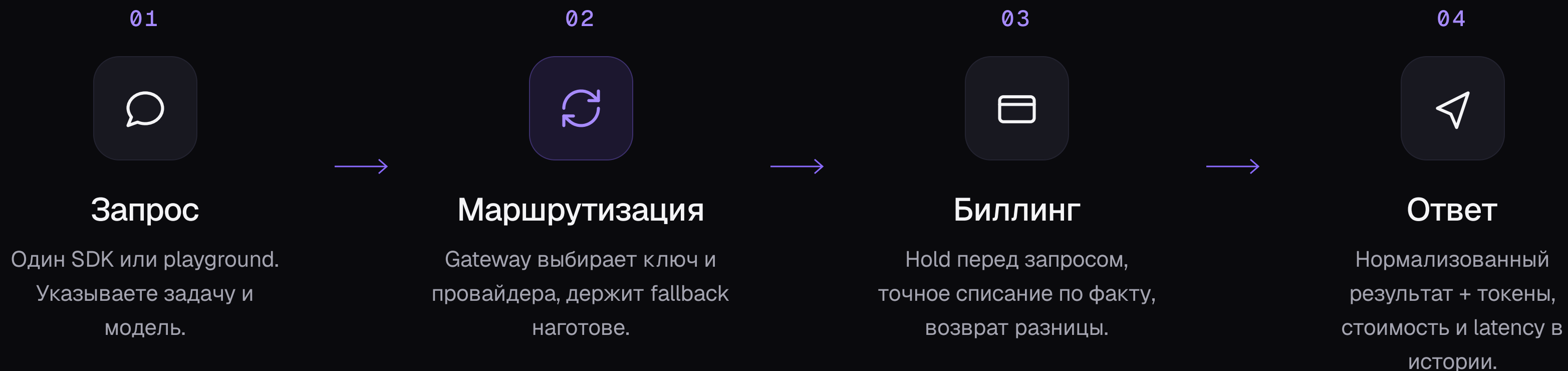


Контроль и аналитика

История каждого запроса, расход по моделям и командам, лимиты и auto-fallback.

90 секунд от регистрации до первого запроса.

Путь одного запроса





Auto-fallback


Если основной провайдер упал — запрос переключается на резервный ключ, модель или другого провайдера. В прошлом месяце fallback спас 78.9% сбойных запросов.

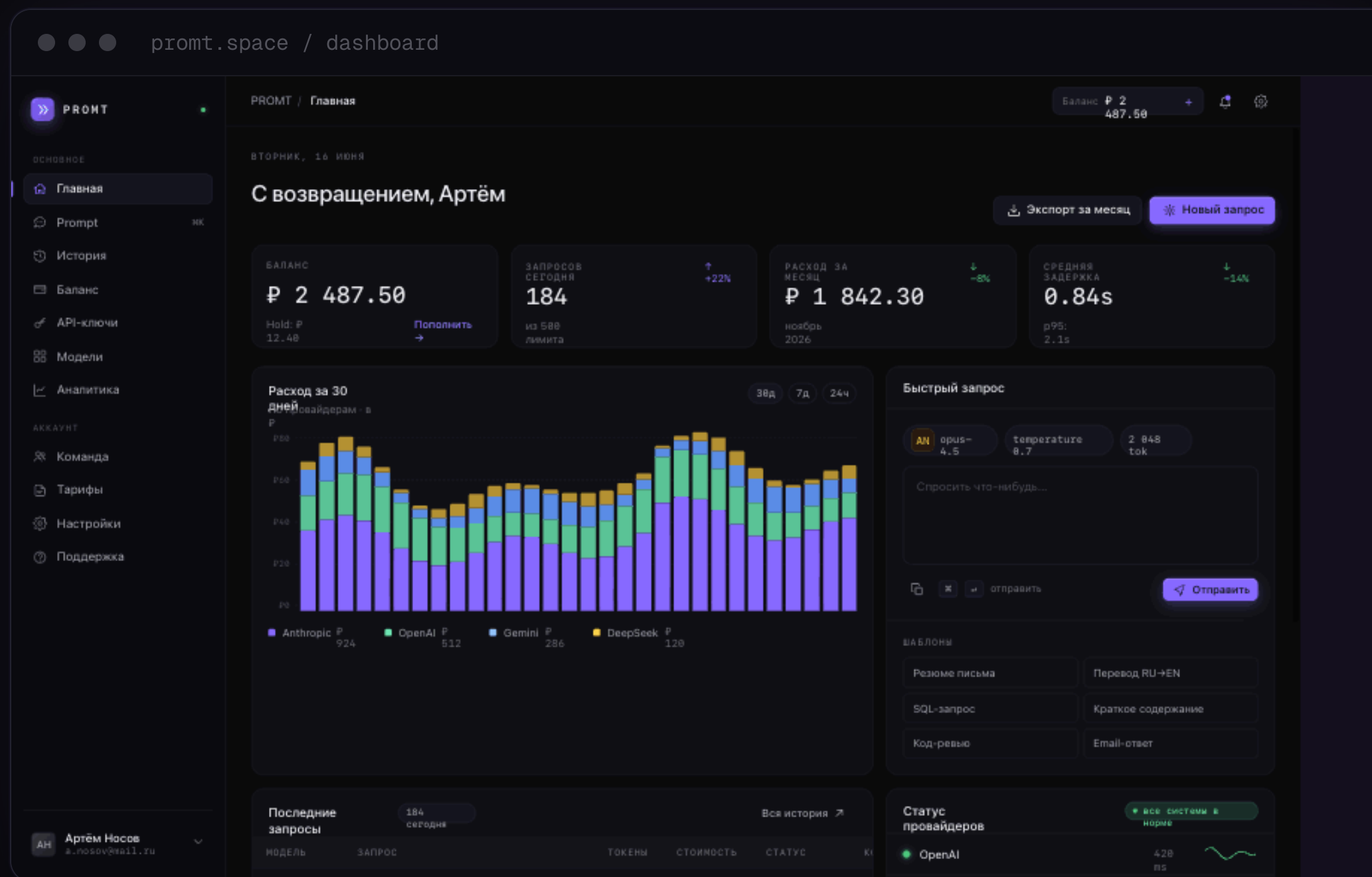
ПРОДУКТ · ГЛАВНАЯ

Весь расход — на одном экране

 **График расхода 30 дней**
Стек по провайдерам в реальном времени

 **Быстрый запрос**
Промпт прямо с главной + шаблоны

 **Статус провайдеров**
Latency и доступность каждого сервиса

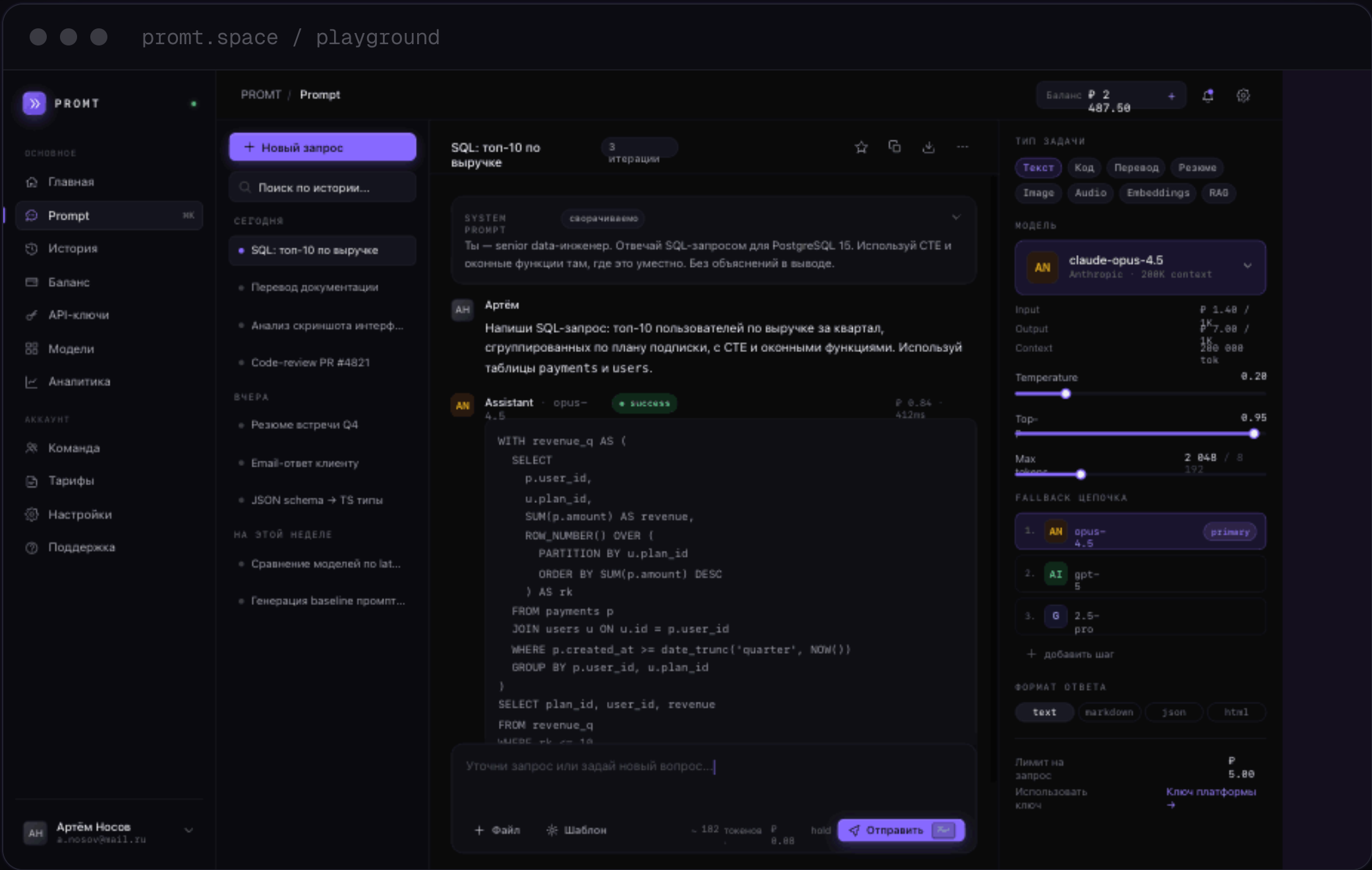


Prompt-окно с полным контролем

Сессии слева, диалог в центре, параметры справа — как в IDE для запросов.

- выбор модели
- temperature · top-p
- max tokens
- fallback-цепочка
- формат ответа
- system prompt
- streaming
- оценка стоимости

Перед отправкой видно прогноз: ~182 токена · ₹ 0.08 hold. Никаких сюрпризов в счёте.



ПРОДУКТ · ИСТОРИЯ

Каждый запрос — под учётом

14 829 запросов с фильтрами по провайдеру, модели, статусу, типу и стоимости.

96.7%

успешных запросов

0.84s

средняя задержка

Экспорт в CSV/JSON, повтор любого запроса в один клик.

promt.space / history

ПРОМТ / История запросов

Баланс: Р 2 487.50

История запросов

Всего 14 829 запросов · последние 30 дней

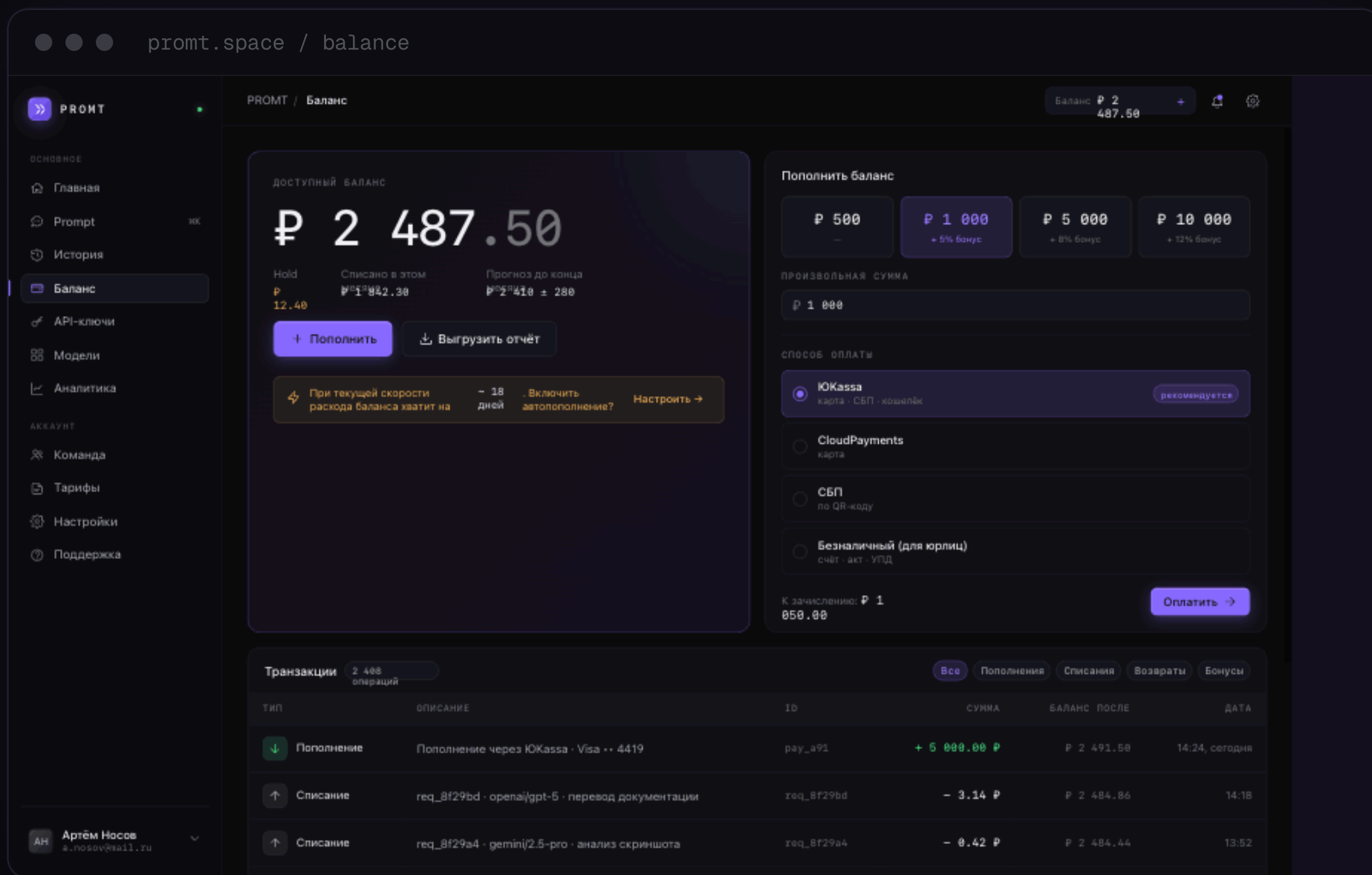
CSV JSON Повторить выбранное

ВСЕГО	УСПЕШНО	ОШИБОК	СРЕДНЯЯ LATENCY	ПОТРАЧЕНО
184 сегодня	178 96.7%	6 3.3%	0.84s p95 2.7s	Р 124.50 средн. Р 0.68

Период: 30 дней | Провайдер: Все | Модель: Все | Статус: Все | Тип: Все | Стоимость: любая | Поиск по prompt / id...

ID	МОДЕЛЬ · ПРОВАЙДЕР	ТИП	PROMPT	TOKENS (IN/OUT)	СТОИМОСТЬ	LATENCY	СТАТУС	СОЗДАН
req_8f2a1c	AN opus-4.5	text	SQL: топ-10 пользователей по выручке за ка...	148 / 312	Р 0.84	412 ms	success	14:24
req_8f29bd	AI gpt-5	text	Перевод документации (10 страниц)	2 418 / 2 388	Р 3.14	2 384 ms	success	14:18
req_8f29a4	G 2.5-pro	image	Анализ скриншота интерфейса	328 / 548	Р 0.42	612 ms	success	13:52
req_8f293f	DS r1	code	Code-review pull request #4821	3 928 / 1 248	Р 0.18	3 412 ms	success	13:31
req_8f2922	AN opus-4.5	summary	Резюме встречи Q4 planning	5 248 / 828	Р 2.18	1 984 ms	success	12:48
req_8f298a	OR qwen-3-max	rag	Embed knowledge base / RAG	1 188 / 0	Р 0.12	—	failed	12:22
req_8f28e6	HF bge-r3	embed	Batch embeddings · 240 docs	48 888 / 0	Р 0.96	6 248 ms	success	11:38
req_8f28c3	AN haiku-4	text	Email-ответ клиенту по жалобе	248 / 188	Р 0.84	284 ms	success	11:14

АН Артём Носов a.nosov@mail.ru



ПРОДУКТ · БИЛЛИНГ

Прозрачный биллинг до токена

HOLD Резерв по прогнозу перед запросом

CHARGE Точное списание по факту токенов

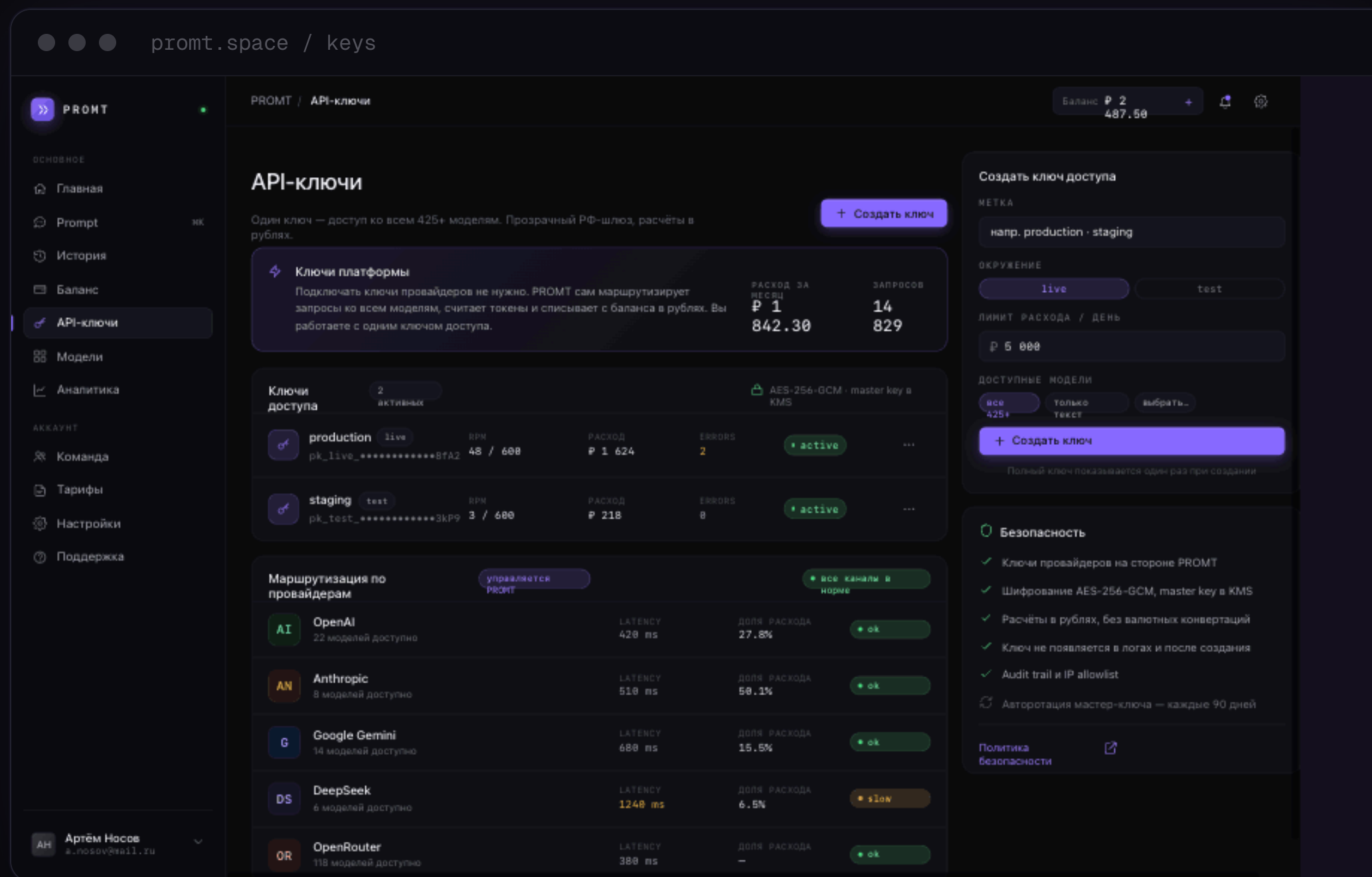
REFUND Возврат разницы, если вышло дешевле

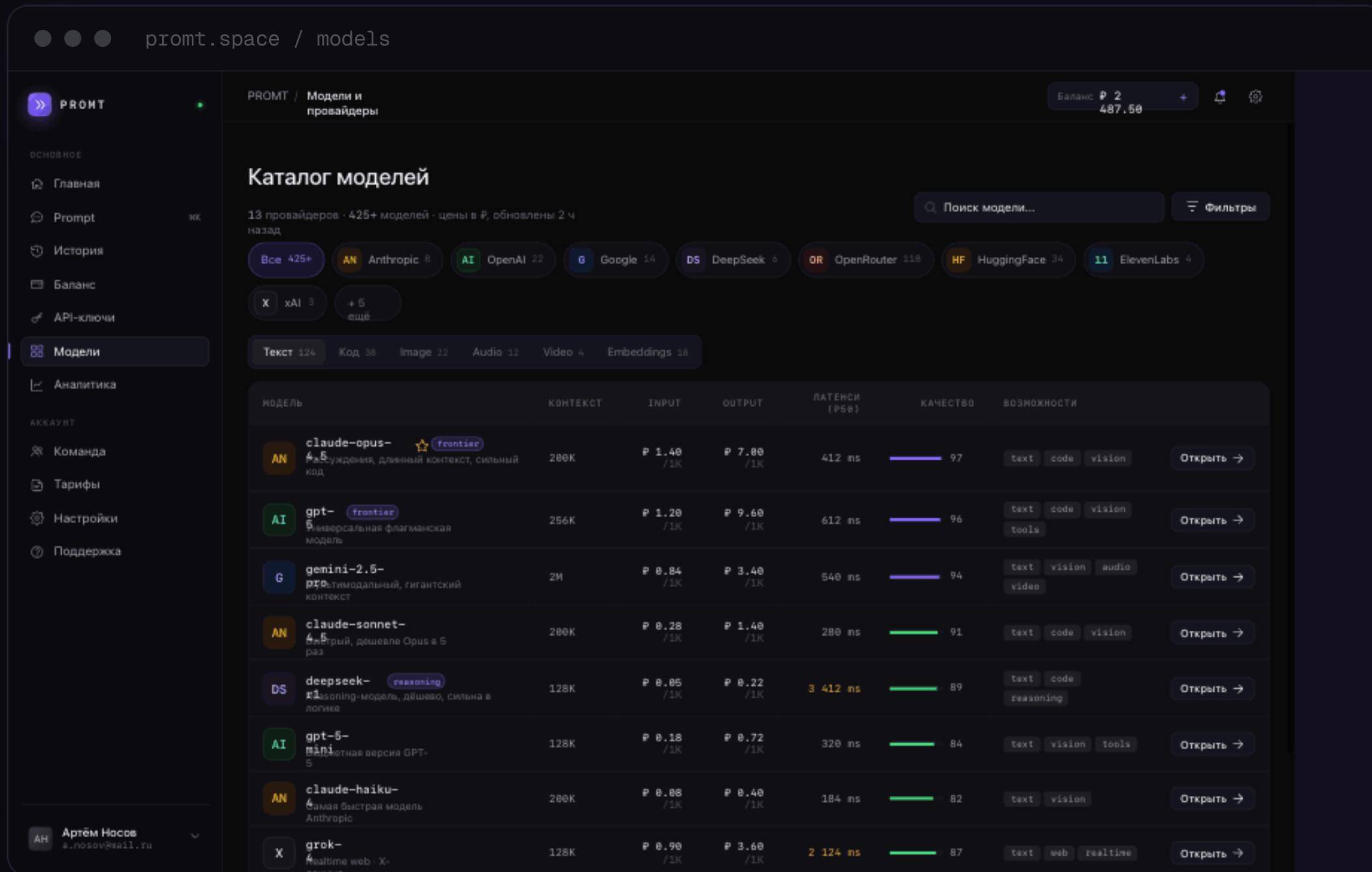
ЮKassa, СБП, безналичный для юрлиц + закрывающие документы.

Один ключ — все модели

Ключи провайдеров подключать не нужно. PROMT маршрутизирует запросы ко всем 425+ моделям и считает токены за вас.

- ✓ Один ключ доступа — все 425+ моделей
- ✓ Расчёты в рублях, без валютных конвертаций
- ✓ Шифрование AES-256-GCM, master key в KMS
- ✓ Лимиты, RPM и счётчик ошибок на каждый ключ





ПРОДУКТ · МОДЕЛИ

425+ моделей в одном каталоге

Цены, контекст, latency и оценка качества — сравнивайте и выбирайте под задачу.

13 425+ 6

провайдеров моделей категорий

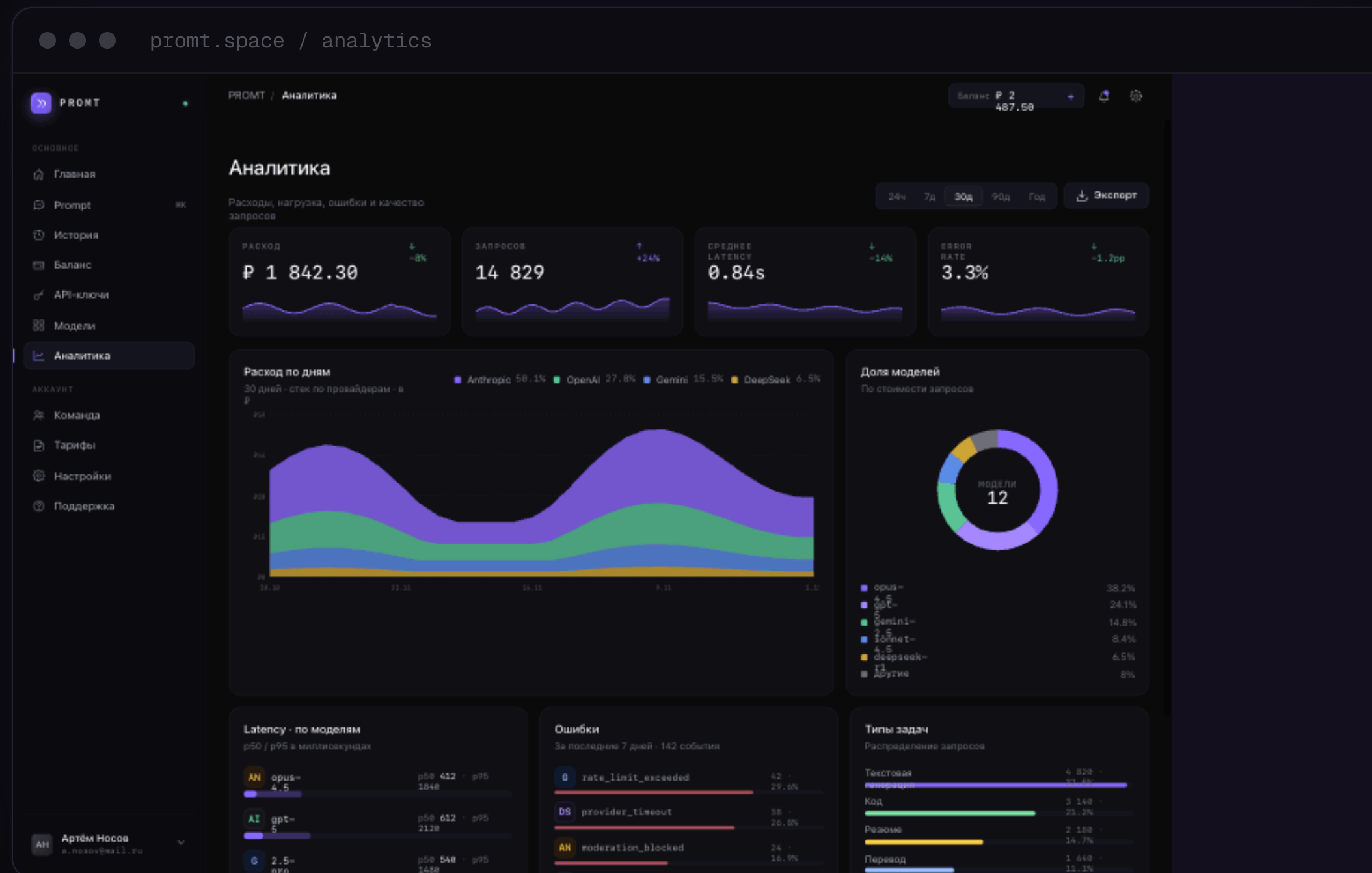
Текст, код, image, audio, video, embeddings — цены обновляются автоматически.

ПРОДУКТ · АНАЛИТИКА

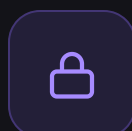
Видно, куда уходит каждый ₹

Расход по дням и провайдерам, доли моделей, latency, ошибки и типы задач.

- ▶ Стек расхода и доля каждой модели
- ▶ p50 / p95 latency по моделям
- ▶ Топ ошибок и эффективность fallback
- ▶ Экспорт в любой период



Ключи и данные — под защитой



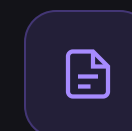
AES-256-GCM

Все ключи шифруются, master key хранится в KMS / Vault.



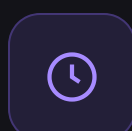
Без утечек

Ключи не появляются в логах и не отображаются после сохранения.



Audit trail

Журнал всех действий с аккаунтом, ключами и лимитами.



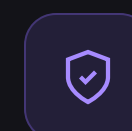
2FA и passkey

Двухфакторная аутентификация, биометрия, контроль сессий.



Rate limiting

Лимиты RPM и расхода, защита от перерасхода через API.

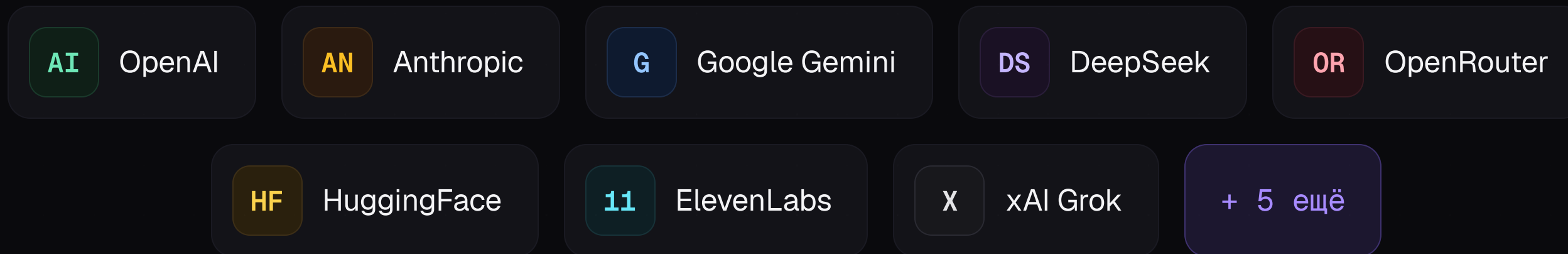


Защита промптов

Модерация и защита от prompt injection на входе.

ЭКОСИСТЕМА

13+ провайдеров. 425+ моделей. Один интерфейс.



Новый провайдер появляется в каталоге — без изменений в вашем коде.

для команд

Один баланс — вся команда

Роли, персональные лимиты и аналитика по каждому сотруднику и отделу.

- ▶ 5 ролей: от владельца до «только чтение»
- ▶ Лимит расхода на каждого сотрудника
- ▶ SSO (Google, Yandex) на тарифах Business+
- ▶ Закрывающие документы и счета для юрлиц

The screenshot shows the 'ПROMT | Команда' (Team) management interface. At the top, it displays the overall team balance: 'Баланс P 2 487.58'. Below this, a summary for the 'Workspace «Носов и Ко»' (Tariff Business, up to 25 seats, 7 active) shows 7 seats used out of 25, a team expense of P 18 240, and 84 requests, with a 120% increase in requests for the month.

The main section is a table of team participants (Участники) with 7 members and 2 pending invitations. The table columns are: СОТРУДНИК (Employee), РОЛЬ (Role), ЛИМИТ / МЕС (Limit / Month), ИСПОЛЬЗОВАНО (Used), ЗАПРОСОВ (Requests), СТАТУС (Status), and АКТИВНОСТЬ (Activity). The roles listed are: Владелец (Owner), Администратор (Administrator), Разработчик (Developer), Аналитик (Analyst), and Только чтение (Read-only).

СОТРУДНИК	РОЛЬ	ЛИМИТ / МЕС	ИСПОЛЬЗОВАНО	ЗАПРОСОВ	СТАТУС	АКТИВНОСТЬ
АН Артём Носов a.nosov@mail.ru	Владелец	Без лимита	—	P 4 820 / 18 420	active	сейчас
МЛ Мария Лебедева m.lebedeva@mail.ru	Администратор	P 5 880	—	P 3 240 / 14 280	active	12 мин
ИС Игорь Соколов i.sokolov@mail.ru	Разработчик	P 5 880	—	P 4 910 / 22 140	active	1 ч
ДК Дарья Кузнецова d.kuznetsova@mail.ru	Разработчик	P 3 880	—	P 1 820 / 8 640	active	3 ч
ПО Павел Орлов p.orlov@mail.ru	Аналитик	P 2 880	—	P 920 / 4 210	active	вчера
ЕЖ Елена Жукова e.zhukova@mail.ru	Только чтение	P 580	—	P 140 / 380	active	2 дня
k.belov@mail.ru приглашение отправлено	Разработчик	P 3 880	—	—	pending	—
s.frolov@mail.ru приглашение отправлено	Аналитик	P 2 880	—	—	pending	—

At the bottom, a role selection bar shows the current user 'Артём Носов' (Role: Владелец) and the counts for each role: Владелец (1), Администратор (1), Разработчик (3), Аналитик (2), and Только чтение (1). A legend below explains the permissions for each role.

Платите только за использованное

START	PRO популярный	BUSINESS	ENTERPRISE
0 ₪ pay-as-you-go	990 ₪ /мес + расход по токенам	4 990 ₪ /мес + расход по токенам	от 100 000 ₪ /мес · SLA
<ul style="list-style-type: none">✓ Все провайдеры✓ 100 запросов / день✓ Email-поддержка	<ul style="list-style-type: none">✓ Streaming responses✓ Ключи платформы✓ История 90 дней✓ Webhooks, экспорт	<ul style="list-style-type: none">✓ До 25 сотрудников✓ Роли и лимиты✓ SSO (Google, Yandex)✓ Закрывающие документы	<ul style="list-style-type: none">✓ Private deployment✓ Выделенные каналы✓ SLA 99.95%✓ Аудит и compliance

Прозрачный рабочий РФ-шлюз. Все расчёты в рублях — без валютных конвертаций и зарубежных карт.

Без шлюза и с PROMT

БЕЗ ШЛЮЗА

- ✗ Отдельная интеграция на каждый провайдер
- ✗ Ключи в .env, нет ротации и контроля
- ✗ Несколько биллингов и валют
- ✗ Упал провайдер — встала фича
- ✗ Расход виден только постфактум

С PROMT

- ✓ Один SDK на 425+ моделей
- ✓ Шифрованные ключи, лимиты, audit
- ✓ Общий баланс и закрывающие
- ✓ Auto-fallback между провайдерами
- ✓ Аналитика и лимиты в реальном времени

Первый запрос за 90 секунд

01

Регистрация

Без карты. Сразу 200 ₽ на баланс для тестов.

02

Выбор модели

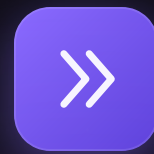
Один ключ доступа, баланс в рублях. Без ключей провайдеров.

03

Запрос

Из playground или одним вызовом SDK. Готово.

```
$ npm install @prompt/sdk  
> const r = await prompt.chat({ model: 'claude-opus-4.5', prompt })
```



Перестаньте жонглировать API-ключами.

Один кабинет. Один баланс. Один SDK.

Начать бесплатно

Демо для команды

`prompt.space · sales@prompt.space · @prompt`