ПРОДУКТ · LLM-ИНФРАСТРУКТУРА

AgentHub — один шлюз для всех ИИ-моделей компании

Ниже счета за API, полный контроль расходов, никакой зависимости от вендоров. Уже работает в производстве.

29 моделей · экономия API-трафика **20–40%** · окупаемость **8,0 мес**

ПРОБЛЕМА

Почему вы переплачиваете за ИИ

Счета без контроля

Каждый сотрудник — своя подписка. Никто не знает, сколько тратит компания и на что именно.

Зависимость от одного вендора

Смена модели или рост цен — боль всей команды: миграция ключей, кода, промптов.

Нет квот и аудита

Нельзя ограничить сотрудника, увидеть историю запросов, выставить биллинг отделов.

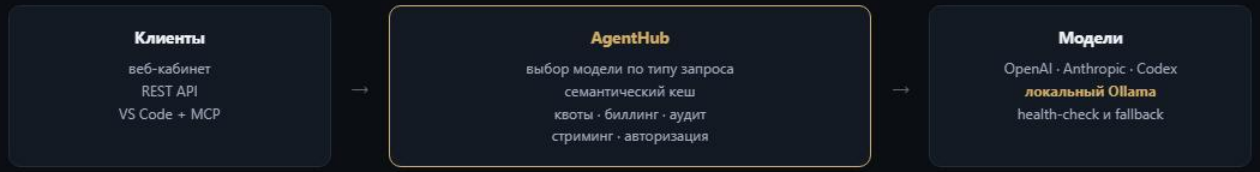
Данные уходят наружу

Конфиденциальные запросы летят в чужие облака вместо локальной модели.

РЕШЕНИЕ

Всё — через одну точку

Шлюз сам решает, какая модель ответит на каждый запрос и сколько это стоит. Вы видите каждый рубль.



Работает в производстве: api.logeionai.ru · PostgreSQL + Redis + Caddy · 239+ автотестов

ВОЗМОЖНОСТИ

Что умеет AgentHub

Ротация 29 моделей
Health-чеки и автоматический fallback на локальную модель — сервис работает, даже если упал вендор.

Умная маршрутизация
Запрос классифицируется по типу и уходит самой дешёвой подходящей модели.

Семантический кеш
Повторяющиеся вопросы не оплачиваются дважды.

Квоты и биллинг
Лимиты по пользователям и отделам, статистика расходов, админ-панель.

Потоковая генерация
SSE-стриминг ответов и генерация изображений.

MCP-экосистема
Агент с инструментами и расширение VS Code — фундамент для корпоративных агентов.

ЭКОНОМИКА

Сколько стоит и как быстро окупается

500 000 ₽

эффект в месяц

6 000 000 ₽ в год

4 000 000 ₽

вложения в разработку

разово

8,0 мес

срок окупаемости

консервативная оценка

За счёт чего эффект

Маршрутизация на дешёвые модели, кеш повторных запросов, локальные модели для внутренних задач.

Чем это лучше подписок

Замена Chat GPT / Claude Code - Основнео ядро для разработки.

Что входит

Шлюз, админ-панель, биллинг, MCP-расширение, документация, развёртывание в контуре заказчика.

ПОЧЕМУ МЫ

Почему AgentHub, а не ещё одна подписка

Работает сегодня

Не прототип: в производстве, 239+ автотестов, документация. Внедрение — дни, а не кварталы.

Данные остаются вашими

Локальные модели в вашем контуре для конфиденциальных сценариев.

Независимость от вендоров

Смена модели — изменение конфигурации, а не переписывание интеграций.

Фундамент для агентов

MCP, инструменты, стриминг, квоты — готовая база для корпоративных ИИ-агентов.

ПРЕДЛОЖЕНИЕ

Три формата сотрудничества

ПИЛОТ

1 неделя

— развёртывание на ваших данных
— демо маршрутизации и экономии
— отчёт с расчётом эффекта

РЕКОМЕНДУЕМ

ВНЕДРЕНИЕ ПОД КЛЮЧ

4 000 000 ₽ разово

— шлюз, биллинг, админка в вашем контуре
— подключение ваших моделей и провайдеров
— обучение команды, документация
— окупаемость 8,0 мес

РАЗВИТИЕ

SLA

— поддержка и обновления
— новые провайдеры и модели
— кастомные агенты на базе MCP

Запросить пилот

AgentHub · production LLM-шлюз · api.logeionai.ru