

ИИ, который работает **в бою**, а не на демо.

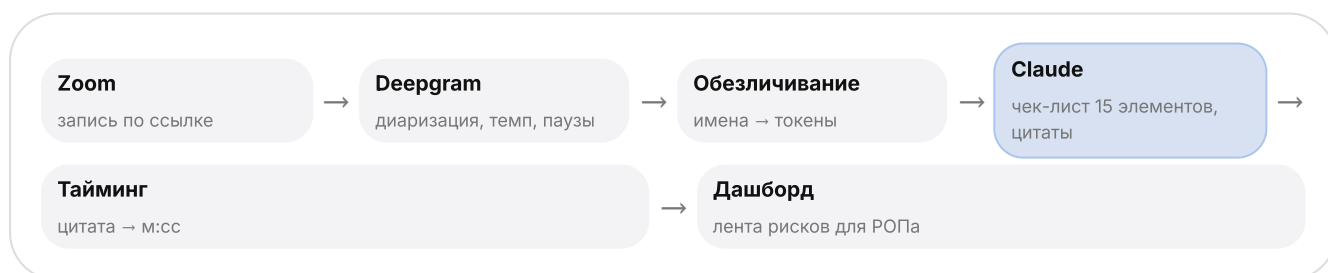
Три года единственный инженер в компании с отделом продаж на 10 человек и 2 000 лидов в месяц. Четыре системы ниже я спроектировал, собрал и держу в эксплуатации один.

12 проектов в параллели за месяц	177 коммитов за сентябрь 2026	10 ночных конвейеров без платной платформы	26 встреч в месяц с руководителями и подрядчиками	18 подрядчиков и сервисов на мне
--	---	--	---	--

КЕЙС 01 · ИИ-КОНТРОЛЬ КАЧЕСТВА ПРОДАЖ

Каждая встреча менеджера разобрана моделью. 300 в месяц. 26 ₽ за встречу.

Руководитель продаж мог послушать единицы из 300 Zoom-встреч. Внешний подрядчик стоил дорого и давал оценки без доказательств.



300 встреч в месяц покрыто разбором вместо выборочных	≈26 ₽ один разбор на Claude Sonnet, стоимость пишется из usage API	22/30 совпадение лучшей модели с ручной оценкой на эталонах
---	--	---

Доказательства, не мнения

На каждый из 15 элементов модель отвечает «спрошено / присутствует», объясняет почему и приводит цитату. Код находит цитату в расшифровке и ставит тайминг.

Выбор модели по цифрам

Четыре модели на одних эталонах против человека. Дешёвая на поток, дорогая на спорные. Весь бюджет 7 800 ₽ в месяц при согласованных 10 000.

Как устроено внутри

- Речевые метрики из диаризации: доля речи менеджера (норма 45–60 %), длина монологов, темп, паузы, вопросов в минуту.
- Формула балла: цель встречи 30 + элементы 30 + этапы 25 + речь 15; красная зона перекрывает всё. Вкладка каждого менеджера с его личным скриптом.
- Обезличивание перед отправкой модели с обратной подстановкой в ответе. Содержание встреч не попадает в логи и git.
- Раннер GitHub Actions, четыре сценария: пакет по листу, одна запись, диаризация, пересчёт. Смена провайдера моделей за день, когда прежний ограничил аккаунт.

Python

Deepgram nova-3

Claude Sonnet / Opus

OpenAI-совместимый API

GitHub Actions

Google Sheets API

Cloudflare Pages

natasha

КЕЙС 02 · ИИ-АССИСТЕНТ ПРОДАЖ В ЧАТЕ САЙТА

Чат, который сам квалифицирует лида и ведёт переписку в CRM в обе стороны.

Лиды с сайта падали в «Неразобранное» без источника, менеджеры отвечали с задержкой, ночью не отвечал никто.

Виджет

один сниппет, любой сайт

Бот-квалификатор

команда, оборот, роль, ниша

Дедуп по телефону

активная сделка есть → не плодим

amoCRM

сделка с источником и UTM

Менеджер

отвечает из CRM, клиент видит в чате

100 %

лидов из чата попадают в CRM с источником и UTM, минуя «Неразобранное»

0

нарушений правил в стресс-тесте на 50 диалогов

0 ₺

за хостинг и модель в базовом режиме: Cloudflare + Workers AI

Неотличим от менеджера

Персона и примеры из реального корпуса переписки, дебаунс, «печатает...», тайминги как у человека. Защита от инъекций и провокаций: два подряд и бот замолкает.

Критики на агентах

Живые диалоги разбирают шесть ролей (судья правил, уход клиента, голос, сверка с CRM, парсер) и скептик, который отбросил 4 из 14 находок как ложные.

Как устроено внутри

- Cloudflare Pages Functions + D1, деплой из GitHub Actions, двусторонняя интеграция через API чатов amoCRM (amojo) с подписью запросов.
- Диалог как конечный автомат: приветствие → касания → квалификация → передача. Ночная ветка, лесенка цены, четыре разных захода за номером.
- Нецелевых (команда меньше 5 или оборот ниже порога) бот мягко отсеивает и закрывает сделку автоматически.
- Статистика: воронка по диалогам, медианы ответа по менеджерам, источники с раскрытием до страницы и переписки.
- Ловушка, которую поймал: автоназначение ответственного в CRM система считала перехватом человеком и глушила бота. Закрото списком служебных пользователей и окном 60 секунд.

Дочитали досюда? Значит, вы из тех, кто изучает, а не листает. Для вас тут кое-что есть

Node.js

Cloudflare Pages + D1

amoCRM API

amojo

Workers AI / Claude

Claude Agent SDK

SQL

КЕЙС 03 · ЦЕНТР УПРАВЛЕНИЯ КОМПАНИЕЙ

Один экран у собственника вместо десятка таблиц. Подрядчики просили 150 000 ₺ в месяц.

Цифры жили в Google-таблицах, CRM, АТС и Метрике. Собственник хотел план и факт в одном месте и ответы на вопросы без аналитика.

Источники

CRM · АТС · Метрика ·
таблицы



Ночной сборщик

Workers, крон



D1 + KV

витрина данных



Экраны

воронки, срезы, звонки



Чат с ИИ

вопросы по цифрам, генерация экранов

150 000 ₽

в месяц: оценка аналога
подрядчиками. Сделано
внутри

4 шага

досье клиента перед
встречей собирает
модель с веб-поиском по
событию в CRM

0

pull request: коммит в
master, автопроверка API
и страниц после выкладки

Для руководителей

Воронки план/факт с прогнозом на
месяц, конверсии между ступенями,
срезы по источникам, тегам и UTM,
звонки по менеджерам первой линии.

Для команды

Права по разделам и ролям, вход
через Telegram, генератор ссылок для
блогеров, карта воронок, дашборд
проектов и утренний бриф в Telegram.

Как устроено внутри

- Cloudflare Pages + Workers + D1 + KV. Сборщик ночью тянет CRM, звонки АТС, Яндекс Метрику и таблицы; страницы читают API.
- Подготовка диагноста к встрече: по переходу сделки на этап «встреча подтверждена» модель собирает досье в четыре шага. После пилота стоимость 335 ₽ за подготовку снижена базой знаний «было → стало».
- Разбор встречи после: лояльность клиента, ключ к следующему контакту, контекст для бота пингов.

Node.js

Cloudflare Workers / Pages / D1 / KV

amoCRM API

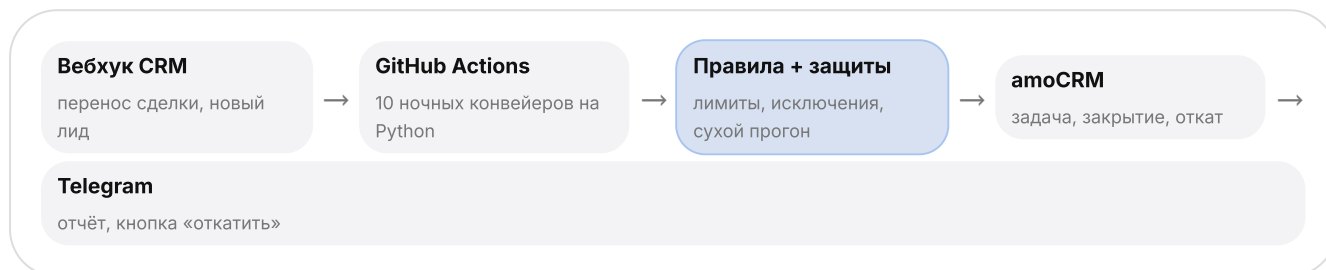
Яндекс Метрика API

Claude Sonnet + web search

Telegram Login

Вся автоматизация переехала с n8n на GitHub Actions. Робот впервые получил право писать в CRM.

Компания платила за n8n, сценарии ломались молча, править их мог один человек. Нужно было убрать платформу, не потеряв ни одного отчёта.



0 ₽

за платформу автоматизации вместо подписки n8n

1 386

групп дублей объединено, около 2 400 контактов, 0 ошибок

1 621

лид восстановлен после инцидента с разметкой источников

Автозакрытие с тормозами

27 дней без движения → задача ответственному, 30 дней → закрытие с причиной. Лимит на запуск, список исключений, тег «не автозакрывать», сухой прогон при ручном старте.

Сторож за один день

Вебхук ловит перенос сделки в чужую воронку, пишет в Telegram, откат по кнопке. Пять инцидентов за месяц закрыты в день обнаружения.

Как устроено внутри

- Конвейеры: сделки и контакты, звонки АТС, телефония по отделам, события по менеджерам, активность базы, транскрибация Zoom. 21 workflow и 27 скриптов в одном репозитории, отчёты и ошибки в Telegram.
- Прокладка между чат-ботом и CRM решает, создавать ли сделку, по правилам воронок и поиску по последним 10 цифрам телефона.
- Резервный Telegram-бот за день, когда внешняя платформа чат-ботов легла: тексты, дожимы, сделка в CRM, переписка в обе стороны.
- Секреты в Doppler, хостинг Cloudflare и Timeweb App Platform.

Python

GitHub Actions

amoCRM API + webhooks

Google Sheets API

Telegram Bot API

Cloudflare Workers

Doppler

Timeweb

КАК РАБОТАЮ

От разговора с бизнесом до боевой эксплуатации. Один.



Сначала процесс, потом код

Шесть лет руководил отделами продаж. Решение начинается с интервью с заказчиком и заканчивается экраном, которым реально пользуется менеджер.



Claude Code и агентные конвейеры

Работаю через агентов, поэтому один закрываю объём небольшой команды: 177 коммитов и 12 проектов в параллели за месяц.



Цифры и защиты

Стоимость каждого запроса в отчёте, сверка с человеком на эталонах, сухие прогоны, лимиты и откат по кнопке до того, как робот трогает данные.



2018–2019

Менеджер по продажам



2019–2023

Руководитель отдела продаж, команды до 15 человек



2023–2024

Project Manager: запуски, метрики, подрядчики



2024–сейчас

AI-инженер, всё техническое направление компании

Покажу системы вживую.

20 минут созвона: открою ассистента, центр метрик и конвейер контроля качества. Ищу роль AI-инженера от 250 000 ₽, удалённо или гибрид в Москве. Готов к проектам.

Telegram: @kaganovichnt · <https://t.me/kaganovichnt> · Страница: <https://cases-cc2.pages.dev/>

Хабр Карьера · <https://career.habr.com/nisrock>

LinkedIn · <https://www.linkedin.com/in/никита-каганович-49a011444>

GitHub · <https://github.com/Nisrock>

Никита Каганович · 2026 · данные компании и клиентов не раскрываются, цифры агрегированы