

Андрей Катков

Инженер LLM-систем и автоматизации

Собираю системы, которые работают в проде без дежурного: данные, роботы, ИИ-агенты и инфраструктура под ними. Проектирую, пишу и эксплуатирую сам — от схемы базы до firewall и отката после неудачного релиза.

55	51	5	87	148	1 человек
сервисов в реестре	живых прямо сейчас	серверов	контейнеров	задач по расписанию	проектирует и эксплуатирует

- Аналитика данных
- Маркетплейсы
- Цены и репрайсинг
- Кастомные CRM
- ИИ-агенты
- Интеграции
- Документооборот и ЭДО
- Telegram-боты
- Парсинг и роботы
- Сайты и витрины
- Инфраструктура
- Бэкапы
- Сети и доступы
- Безопасность
- Юнит-экономика

+7 985 251 7615 Telegram @sales_wb_ozon a-katkov@yandex.ru aiklerk.ru

01

НАПРАВЛЕНИЯ

Что я делаю

Пятнадцать направлений, и каждое — не описание компетенции, а система, которая работает в проде и которую я обслуживаю. По любому пункту открываю код, конфигурацию и журнал эксплуатации.

Данные, маркетплейсы, цены, учётные системы

01

Аналитика данных и дата-платформы

Единый fact-store вместо ручных выгрузок: слои raw → stg → dim/fact → витрины, миграции с проверкой DDL, API поверх данных, Lineage Explorer «откуда взялась эта цифра». Дашборды продаж, P&L с фактической себестоимостью, воронка, CPO и DPP, когорты.

Отличие. Отдельный контур достоверности: ревизоры инвариантов ловят искажения в отчётности площадок до того, как цифра попадёт клиенту.

02

Маркетплейсы: операционка под ключ

Ozon, Wildberries, Яндекс.Маркет. Планировщик поставок по складам и федеральным округам, кратность коробов, грузоместа и штрихкоды, файл распределения и письмо на склад. Отзывы, возвраты и ПВЗ, свод артикулов, ротация ШК, аудит рич-контента, защита бренда от фейковых продавцов.

Чем подтверждено. Три клиентских кабинета на одной платформе, ежедневная операционка без ручных выгрузок.

03

Ценообразование и репрайсинг

Отдельная услуга, а не подпункт маркетплейсов. Мониторинг полок трёх площадок собственным парсером, эталон целевых цен, авто-корректор в двух репрайсерах, разведение повышений и понижений по фазам вокруг календаря акций.

Отличие. Факт установки цены проверяется чтением полки, а не ответом API: площадка отвечает «ок» и не меняет цену чаще, чем принято думать.

04

Кастомные CRM и учётные системы

Система под процесс собственника, а не «карточки сделок». В CRM для розницы техники Apple собственник видит поводы позвонить: у кого техника устарела и пора предложить апгрейд с зачётом старой, кому подойдёт пришедшая партия, у кого день рождения.

Чем подтверждено. Рабочее демо + три телеграм-бота вокруг; отдельно — торговая площадка с аукционом закрытых ставок и отчётом по VIN.

ИИ-агенты, интеграции, документооборот, боты

05

ИИ-агенты и LLM-инфраструктура

Не «прикрутить чат», а собрать контур: роутер запросов между движками с failover, ярусы моделей под цену задачи, сборка контекста, учёт расхода токенов по потребителям. 22 специализированных агента, 81 автоматизированная процедура.

Отличие. Правила держатся кодом: агент без явно выбранной модели не стартует, отправка почты мимо общей библиотеки бросает исключение.

06

Интеграции

Ozon Seller API, Wildberries API, Яндекс.Маркет API, Google API, Telegram Bot API, IMAP/SMTP, Контур.Диадок, MOEX ISS, NocoDB, MCP-серверы для агентов.

Там, где API нет вообще — headless-браузер вместо интеграции: роботизация на управляемом Chrome, включая распознавание капчи зрением модели.

07

Документооборот и ЭДО

Мультиарендный кабинет с изоляцией на уровне строк базы, разбор входящего потока (реквизиты, суммы, OCR сканов), задачи по отделам, аудит, календарь налоговых дедлайнов. Робот сам ходит в Контур.Диадок и готовит черновики.

Генератор документов: договор, акт, счёт из одного файла данных сразу в HTML, PDF, DOCX и XLSX; битый документ физически не уходит клиенту.

08

Telegram-боты и цифровые персоны

От утилитарных ботов (вход по одноразовому коду, алерты, приём файлов, ответ готовым документом) до роя из 14 персон в одном процессе: имя, характер, память о людях, общий ростер отношений, разметка настроения, цепочки обращений друг к другу.

Чем подтверждено. Четыре независимых гейта изоляции чатов — данные одного клиента не утекают в разговор другого.

Роботы, витрины, инфраструктура, бэкапы

09

Парсинг, мониторинг, разовая аналитика

Парсеры полок и каталогов, разбор биржевых PDF, сбор вакансий из потока Next.js, проверка контрагента по ИНН из открытых реестров вместо платной подписки на агрегатор.

Разовые отчёты по инцидентам: товар клиентов на складах, попавших под удары, с компенсациями и детализацией по артикулам — клиентский PDF на кабинет.

10

Сайты, витрины и презентации

Лендинги и сайты-визитки, каталоги на WordPress, где контент и карточки правятся из рабочего места сотрудника через собственные REST-эндпоинты — без доступа к серверу, с логом версий и откатом одним запросом.

Движок деков: слайд хранится как данные — версии, снапшоты, замок релиза; слайд не сохранится без указания источника цифры.

11

Инфраструктура, DevOps, эксплуатация

Docker, Proxmox/LXC, systemd, обратный прокси на 113 маршрутов и 23 домена, автоматический TLS, деплой с бэкапом до правки и командой отката в отчёте, мониторинг доступности API и сессий кабинетов.

Автономный сторож ежечасно обходит хосты и чинит сломавшееся сам — с бэкапом, планом отката и журналом, не трогая необратимое без человека.

12

Бэкапы и катастрофоустойчивость

Бэкап, который переживает взлом источника: боевой сервер пишет в репозиторий «только на дозапись», ключ на его стороне физически не может стереть историю. Версионирование, дедупликация, глубина 14 дней / 8 недель / 6 месяцев.

Перед тем как резать старое — побайтовая сверка с источником; диск подключается только на ночное окно, молчание узла считается тревогой.

Сети и доступы, безопасность, экономика, голос

13

Сети, доступы, рабочие места команды

Отказоустойчивая транспортная сеть между площадками с балансировкой и fail-closed, выдача и отзыв ключей через веб-панель, kill-switch на рабочих машинах.

Виртуальные рабочие места: изолированные окружения, единая раскатка правил, разграничение прав, журнал, блокировка уволенного одной процедурой.

14

Прикладная безопасность

Отдельная модель читает каждое входящее письмо раньше человека и ищет мошеннические реквизиты и скрытые команды, которыми пытаются перехватить управление агентами. Сработало — письмо уходит в карантин, а не в конвейер.

В основании: разделение арендаторов средствами базы, аудит действий, запрет необратимых операций коду, а не инструкции.

15

Управленческий учёт и юнит-экономика

Система, отвечающая на вопрос «сколько на самом деле стоит час работы»: расход токенов моделей по сессиям и потребителям, рабочее время в разрезе клиент × операция, доходы, расходы, ФОТ и прибыль по проектам.

Данные собираются сами: агент снимает активность, демон читает транскрипты сессий и льёт в базу; у каждого сотрудника свой кабинет с расчётами.

+

Голос, встречи и медиа

Автоматическая запись Zoom, Teams и Телемоста — включая бота-участника, который сам заходит по ссылке и пишет звук; расшифровка собственным whisper на выделенном сервере, саммари в трекер.

Плюс: голосовой ассистент hands-free поверх рабочей сессии и генерация изображений и видео для карточек по референсам.

02

ДОКАЗАТЕЛЬСТВА

Что уже работает

Не пересказ и не концепты: системы в моей собственной эксплуатации.

Абсолютные показатели клиентов — под NDA; методику расчёта и код разбираю на встрече.

Платформа агентства маркетплейсов

Три клиентских кабинета живут на одной дата-платформе — заказы, реклама, возвраты и прибыль считаются из одного факта, а не из выгрузок вручную.

Задача. Рост штата под рутину оказался дороже, чем автоматизация: каждый новый клиент означал ещё одного человека на выгрузки, сверки и переносы цифр в отчёты.

Что сделано. Дата-платформа на PostgreSQL со слоями raw → stg → dim/fact → витрины и миграциями с проверкой DDL. Поверх — клиентские и командные кабинеты с ролевым доступом и аудитом, P&L по заказам с фактической себестоимостью, рекламные дашборды с воронкой и CPO/ДРР, планировщик поставок, рабочее место по отзывам, дашборд возвратов и ПВЗ.

Результат. Около пятидесяти сервисов и 148 задач по расписанию держит один человек. Ежедневная операционка идёт без ручных выгрузок, а расхождения в отчётности площадок ловятся ревизорами инвариантов до того, как попадут в отчёт клиенту.

СОСТАВ

3 клиентских кабинета

87 контейнеров на боевом хосте

148 задач по расписанию

113 маршрутов и 23 домена в прокси

Стек: Python, PostgreSQL, FastAPI, Docker, Playwright, собственный LLM-роутер с failover. Абсолютные показатели клиентов — под NDA.

AI Klerk — продукт с нуля до рабочего пилота

Мультиарендный кабинет документооборота с ИИ-ролями по отделам и headless-интеграцией с «Контур.Диадоком».

Идея. Продукт устроен как наём роли, а не как «автоматизация»: пятнадцать цифровых сотрудников со специализацией — бухгалтер, финансист, кадры, делопроизводство, юрист, закупки, казначейство, комплаенс — плюс секретариат.

Что внутри. Организации и участники, клиенты и контакты, документы с версиями и извлечениями, задачи по отделам с ролевым доступом, аудит действий, календарь налоговых дедлайнов на собственном парсере правил повторения. Робот забирает входящие из ЭДО, проверяет комплект и готовит черновики — подпись и отправка заблокированы кодом.

Как защищены данные. Приложение ходит в базу ролью, которая не владеет таблицами, и на каждый запрос выставляет организацию: забытое условие в запросе не даёт утечки чужих данных — их не отдаст сама база.

ЧЕМ ПОМЕРЕНО

200+ документов разобрано

15 контрагентов через Диадок

15 ролей + секретариат

5 контейнеров, антивирус и OCR вложений

Стек: Node 22 / Express 5, PostgreSQL с изоляцией арендаторов на уровне строк, Apache Tika + OCR, контейнер разбора файлов без выхода в сеть.

Заказные системы и внутренние продукты

Рой из 14 цифровых сотрудников

Четырнадцать ботов-персон в одном процессе: у каждого свой аккаунт, характер, двухслойная память о людях и общий ростер отношений в коллективе.

Полезная работа: журнал клиентских групп с дедупликацией, разбор фактов прямо из переписки в таблицу, ночная рефлексия разговора в факты о клиенте.

CRM для розницы техники Apple

Кастомная система вместо коробочной: оценка trade-in по модели и состоянию, подбор под пришедшую партию, поводы для контакта, сертификаты, конструктор кампаний.

Стек: Flask и SQLite в Docker на сервере заказчика, три телеграм-бота вокруг — магазин, продажи, ремонт.

Сайт-каталог с правкой через REST

Тема отдаёт собственные эндпоинты: контент, карточки товаров и медиа меняются из рабочего места сотрудника — без доступа к серверу и без правки файлов темы.

Результат: лог версий и откат одним запросом; сайт передан клиенту на его хостинг.

Дашборд топлива

Фьючерсы MOEX FORTS и спот СПбМТСБ за квартал: расчётные цены, объёмы, открытый интерес — включая разбор биржевых PDF-бюллетеней.

Зачем: отраслевая аналитика, которую иначе покупают подпиской.

Поиск по клиентскому архиву

9,4 ГБ материалов: полнотекстовый поиск по содержимому документов (текст, xlsx, pdf, docx) и транслит-поиск по именам файлов.

Доступ: опубликован команде через служебный туннель с боевого хоста, наружу не смотрит.

Автобиржа

Торговая площадка автомобилей: витрина с галереей, аукцион с закрытыми ставками и сбросом таймера, гараж владельца, переписка по объявлению, отчёт по VIN, админка.

Связка Telegram-бота и веб-кабинета на общей базе; демо сохранено и работает.

03

МЕТОД

Как я работаю

Шесть правил, из-за которых системы не разваливаются, когда меня нет рядом.
Это и есть разница между «сделали проект» и «система живёт».

Шесть правил эксплуатации

1

Правило держится кодом, а не инструкцией

Договорённость, которую можно забыть, не считается выполненной. Запрет вшивается в библиотеку и падает исключением — а не живёт в памяти исполнителя.

3

Необратимое остаётся за человеком

Подпись, отправка, удаление боевых данных — робот готовит, решает человек. В продукте это заблокировано кодом, а не настройкой, которую можно переключить.

5

Источник правды один и машиночитаемый

Реестр служб, доменов и серверов с валидатором связей. Любая цифра на витрине прослеживается до строки в базе и до выгрузки, из которой она посчитана.

2

Бэкап до правки, команда отката — в отчёте

Каждый выкат сопровождается точной строкой возврата в прежнее состояние. Не «сможем восстановить», а готовая команда, которую можно выполнить прямо сейчас.

4

Проверка исполнением, а не отчётом

«Сервис поднялся» не равно «данные живые». Цена проверяется чтением полки, страница — исполнением её скриптов, бэкап — побайтовой сверкой с источником.

6

Данные арендаторов разделяет база, а не код

Изоляция на уровне строк: забытое условие в запросе не даёт утечки чужих данных. Безопасность, которая не зависит от внимательности разработчика.

Форматы работы

От разового аудита до полной эксплуатации. Любой формат начинается с разбора процесса, а не с выбора технологий.

ФОРМАТ	ДЛЯ ЧЕГО	ЧТО НА ВЫХОДЕ
Аудит и дорожная карта	Понять, что автоматизировать первым, что переделать, а что просто выключить.	Разбор процессов, список задач с оценкой трудоёмкости, приоритеты и порядок внедрения.
Проект под ключ	Конкретная система: дашборд, интеграция, CRM, робот, документооборот, сайт.	ТЗ → прод → передача с документацией, бэкапом и проверенной командой отката.
Эксплуатация и развитие	Система должна жить и меняться, а не «сдаться и забыться».	Мониторинг, бэкапы, доработки по потоку задач, отчётность о состоянии.
AI-контур в компании	Агенты, которые реально работают в процессах, а не демонстрируются на встрече.	Роутер моделей, гейты и проверки выводов, учёт расхода, раздача процедур команде.
Спасение проекта	Код есть, прода нет; подрядчик ушёл; система работает, но её никто не понимает.	Инвентаризация, доведение до прода, документация и дальнейшая эксплуатация.

Стек

ЯЗЫКИ И ДАННЫЕ

Python

PostgreSQL

SQL

SQLite / FTS5

JavaScript

Node

Bash

FastAPI

Express

LLM И АГЕНТЫ

Оркестрация агентов

MCP-серверы

Ярусы моделей под цену задачи

Гейты выводов

Собственный инференс

whisper

Локальные модели

Учёт токенов

ИНФРАСТРУКТУРА

Docker

Proxmox / LXC

systemd

Caddy

nftables

Proxmox Backup Server

borg

headless Chrome

Playwright

Syncthing

ИНТЕГРАЦИИ

Ozon Seller API

Wildberries API

Яндекс.Маркет API

Google API

Telegram Bot API

IMAP / SMTP

Контур.Диадок

MOEX ISS

Объём кода

2 285 файлов на Python, около 570 тысяч строк — рабочая кодовая база, которая обслуживает всё перечисленное в этой презентации.

Пишется вместе с агентами, которых я построил, — но проектирую и отвечаю за результат я.

Чего я не делаю

Обучение моделей с нуля: претрейн, архитектуры и статьи — не моё. Пишу об этом сразу, чтобы не тратить ваше время на несовпадение.

Шестнадцать лет управления и всё это время — код

С 2008 года — аналитические и управленческие позиции; с 2016-го продажи и проекты в нефтехимии и ТЭК: **Роснефть, СИБУР, Солид**. Формульные контракты, экспортные направления, тендеры; объём по своему направлению поднял с одной тысячи тонн в месяц до десяти.

Всё это время параллельно писал себе инструменты, потому что готовых под задачу не было: сначала расчёты и парсеры, потом целые сервисы. В 2023-м ушёл в свой проект, сделал агентство — и быстро упёрся в простую вещь: нанимать людей под рутину дороже, чем её автоматизировать.

С конца 2024 года инженерия — основная работа. Управленческий опыт не заменился инженерным, а сложился с ним: я одинаково понимаю, почему у процесса такая себестоимость и почему у запроса такой план выполнения.

ОБРАЗОВАНИЕ

MBA Finance and Accounting,
University of Liverpool

МГУ им. М. В. Ломоносова

Языки: английский C2,
французский B2

ГДЕ ЖИВУТ СИСТЕМЫ

2 боевых сервера — Москва, Амстердам

2 домашних узла: бэкап и хранилище

4 страны в транспортной сети

Опишите задачу в двух предложениях

Отвечу, реально ли это автоматизировать, сколько времени займёт и во что обойдётся. Если задача не моя — скажу сразу и не буду продавать.

Любой пункт этой презентации готов открыть кодом, конфигурацией и журналом эксплуатации.

Написать в Telegram

ТЕЛЕФОН

+7 985 251 7615

TELEGRAM

@sales_wb_ozon

ПОЧТА

a-katkov@yandex.ru

ПРОДУКТ

aiklerk.ru

ГОРОД

Москва