

# Александр Голубев

Промпт-инженер и разработчик AI-ассистентов на Python

RAG и базы знаний · Telegram-боты · агенты под документооборот ·  
автоматизация · детерминированные расчёты



## Операторы тонут в одинаковых вопросах?

Ассистент забирает рутину и отвечает круглосуточно, а живые сотрудники занимаются сложными случаями и продажами.

## Сотрудники не могут быстро найти нужный регламент?

Ассистент по базе знаний отвечает по вашим документам и инструкциям, не отрывая экспертов от работы.

## Смету, заявку или отчёт собирают руками из десятка файлов?

Агент делает это за секунды и объясняет каждую строку: откуда взята цифра, из какого файла и с какой строки.

### 01

**Расчёты делает код, а не нейросеть.** Модели нельзя доверять арифметику: она уверенно ошибается. Расчёт и консультация у меня разведены, суммы воспроизводимы до копейки.

### 02

**Качество измеряется, а не обещается.** Перед сдачей прогоняю набор реальных вопросов и показываю цифру: на сколько система ответила верно.

### 03

**Каждое решение объяснимо.** Там, где цена или вывод попадают в документ, есть протокол: почему именно так и на основании чего.

## Агент тендерной сметы

Коммерческий заказ: участие в тендерах на поставку стройматериалов. Агент читает прайсы поставщиков — крепёж, стройматериалы, электрика, сантехника, кабель, — сопоставляет позиции с ведомостью потребности заказчика и заполняет два документа по его же шаблонам: смету в Excel и заявку на тендер в Word. Ручной перенос цен не нужен.

**88 845**

позиций в сводном каталоге

**~18**сек

 полный прогон: прайсы  
→ документы

**30 / 37**

 позиций подобрано без  
человека

**4**

 документа на выходе,  
включая протокол

### // КОНВЕЙЕР

**01**

#### Прайсы

 11 файлов от 4 поставщиков:  
крепёж,  
стройматериалы,  
кабель

**02**

#### Разбор структуры

шапка на 1-й, 7-й, 26-й строке

**03**

#### Единый каталог

 фасовка  
приведена к  
единице

**04**

#### Сопоставление

 артикул → имя  
→ нечёткое

**05**

#### Выбор цены

 правило из  
конфига,  
арифметика

**06**

#### Документы

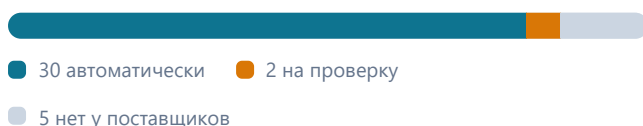
 смета · заявка ·  
протокол

Прогон сделан на настоящих прайсах, скачанных с сайтов поставщиков, — без синтетических данных. Один файл из одиннадцати общий парсер не берёт, и агент не делает вид, что справился: пишет причину отказа в протокол.

### // ЧТО ВЫЯСНИЛОСЬ НА РЕАЛЬНЫХ ДАННЫХ

- **Главная ловушка — фасовка, а не цены.**  
«Дюбель 6×25» стоит 0,10 ₽ у одного поставщика и 672,57 ₽ у другого: это штука и упаковка 500 штук. Цена приводится к единице.
- **Ошибка в одном символе дороже непонятого слова.** M10x160 и M10x16 нечётким сравнением почти неразличимы, поэтому кандидаты берутся только внутри того же типоразмера.
- **Спорное не попадает в смету молча.** Ниже порога уверенности позиция получает статус «на проверку», а не подставленную цену.

### // ЧТО СТАЛО С 37 ПОЗИЦИЯМИ ВЕДОМОСТИ



### // ЦЕНА ПРАВИЛА ВЫБОРА — ВИДНА ДО ПОДАЧИ

Та же ведомость, те же прайсы, разные правила в конфигурации. Прогон занимает секунды, поэтому сравнить сценарии можно до подачи заявки, а не после торга.

 Правило «минимальная цена» **353 850,13 ₽**

 Правило «приоритет поставщика» **400 940,68 ₽**

**+47 090,55 ₽** — столько стоит лояльность к одному поставщику на одном торге. Раньше этой цифры просто не существовало: сравнить два сценария вручную дороже, чем согласиться на первый.

# AI-консультант для магазина стройматериалов

Менеджеры отвечали на одни и те же вопросы клиентов — доставка, сроки, возврат, характеристики, опт — и вручную считали, сколько кирпича, блоков или утеплителя нужно на объект. Я собрал ассистента, который отвечает строго по базе знаний компании и сам считает материалы. Работает в Telegram и виджетом на сайте из одного ядра.

**83 %**

hit-rate@5 петривера на golden-set

**2**

интерфейса из одного ядра

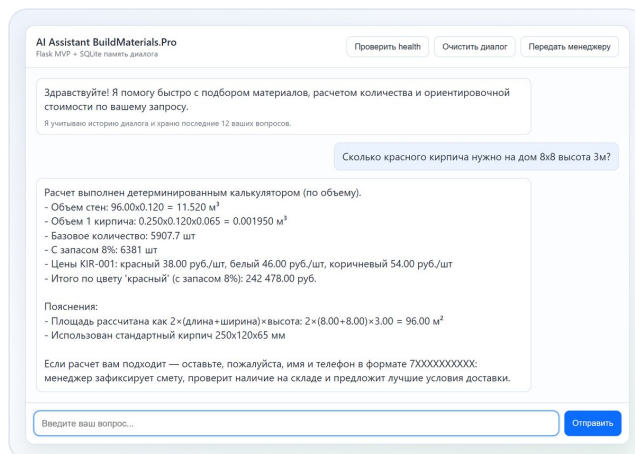
**20–40 %**

экономии на вызовах модели

## // РЕШЕНИЕ

- **Детерминированный калькулятор** — материалы считает код, а не модель, поэтому арифметических ошибок нет
- **Ответы только из документов компании** — на вопрос вне базы бот честно говорит, что не знает, и предлагает менеджера
- **Гибридный поиск** — BM25 и вектор на эмбедингах GigaChat, слияние по RRF, опциональный реранкер
- **Роутинг моделей** — простые вопросы уходят на базовую модель, сложные на Pro
- **Лид-воронка** — мягкий сбор контакта и передача сложных случаев человеку
- **Администрирование без кода** — цены в JSON с горячей перезагрузкой, база знаний меняется файлами

## // СНИМОК ИНТЕРФЕЙСА



Расчёт кирпича на дом 8×8 м: объём стен, количество с запасом, цена по выбранному цвету. Каждый шаг расчёта показан пользователю — эту часть выполняет программный код, а не нейросеть, поэтому сумма всегда одна и та же.

## // ЧТО ПОЛУЧИЛ ЗАКАЗЧИК

### РАБОТА

Бот в Telegram и виджет на сайте, поднимаемые одной командой через Docker Compose — индексация, бот и веб вместе.

### УПРАВЛЕНИЕ

Цены и параметры расчёта — в одном файле, применяются сразу. База знаний пополняется заменой документов, без программиста.

### ЗАЯВКИ

Контакты заинтересованных клиентов выгружаются менеджеру таблицей: дата, имя, телефон, канал и история диалога.

## Ещё три работающих инструмента

### Мониторинг цен и остатков конкурентов

ШАБЛОН · 24 ПРОВЕРКИ ЗЕЛЁНЫЕ

Разовый парсинг даёт снимок одного дня и через неделю бесполезен. Система обходит каталоги конкурентов по расписанию, сравнивает с прошлым разом и присылает в Telegram именно **изменения** — что подешевело, что подорожало, что пропало из наличия, — а не выгружает заново весь список из пяти тысяч позиций, в котором ничего не видно. История складывается в Excel или Google Таблицу.

**1–7**

сайтов конкурентов

**24**

проверки в смоук-тесте

**Telegram**

канал доставки отчёта

### Починка кода, собранного через нейросеть

СВОЙ ИНСТРУМЕНТ ДИАГНОСТИКИ

Свежая волна заказов: человек собрал бота или скрипт через нейросеть, и он почти работает. Написал `audit_project.py` — экспресс-диагностика за 10–15 минут: находит настоящие секреты в коде, обрывы обработки ошибок, скрытые падения. Проверен на демо-проекте и на двух реальных: не срабатывает вхолостую на путях URL и на except `ImportError`.

**10–15 мин**

до вердикта, бесплатно

**Отчёт**

что сломано и что чиню

### LeadRadar — конвейер поиска заказов

РАБОТАЕТ НА МНЕ САМОМ

Собственный инструмент, и поэтому честный показатель инженерной дисциплины. Следит за тринадцатью источниками, двухшаговым скорингом отделяет настоящие заявки от рекламы фрилансеров и чужих резюме, пишет их в трекер и присылает в Telegram с кнопками подтверждения. Состояние сохраняется атомарно: прогон, упавший на середине, ничего не теряет и ничего не оценивает дважды.

**2 194**

позиции оценено за сутки

**13**

источников одновременно

**0**

потерь при сбое прогона

# Как я работаю

Разница между «ботом на нейросети» и системой, которой можно доверить деньги и документы, — не в модели. Она в том, что происходит вокруг модели.

## Расчёты выполняет код

Всё, что превращается в сумму, считает программа, а не языковая модель. Модель может помочь распознать и сопоставить наименование, но её решение проходит через порог уверенности и попадает в протокол.

## Качество измеряется до сдачи

Прогоняю набор реальных вопросов заказчика и показываю результат цифрой. Вы получаете измерение, а не обещание.

## Секреты не попадают в репозиторий

Ключи, доступы к кабинетам, базы с перепиской и данными клиентов исключаются до того, как что-либо выкладывается наружу. Проверка автоматическая — тем же инструментом диагностики.

## Повторный прогон даёт тот же результат

Для тендерного документа это не вкусовщина: те же исходные файлы обязаны дать те же цифры до копейки, иначе документ невозможно проверить.

## На каждую цифру есть ответ «почему»

Протокол сопоставления построчно: с чем сопоставлено, по какому правилу, из какого файла и с какой строки взята цена.

## Проект передаётся, а не привязывается ко мне

Исходный код, инструкция обычным языком и возможность самому пополнять базу знаний. Сопровождение — по желанию, а не потому, что иначе всё встанет.

### Чего я не обещаю

- Что ассистент полностью заменит менеджера. Он снимает типовые вопросы, а сложные передаёт человеку — так надёжнее.
- Что система закроет 100% позиций или вопросов. Она обязана честно показать, что закрыла сама, а что отдала специалисту.
- Что я соберу за вас базу знаний или напишу тексты. Материалы даёте вы, структурирую и подключаю уже я.
- Что обойду защиту от ботов или сделаю веерную рассылку. Это путь к бану аккаунта заказчика, а не к результату.

# С чем можно прийти

## услуга

## цена от

ИИ-бот, отвечающий по вашим документам

8 000 ₽

База знаний из ваших материалов, поиск по смыслу, Telegram или виджет на сайт, инструкция

Мониторинг цен и остатков конкурентов

5 000 ₽

Обход по расписанию, отчёт об изменениях в Telegram, выгрузка в Excel или Google Таблицу

Починка кода, собранного через нейросеть

3 000 ₽

Диагностика бесплатно: за 10–15 минут скажу, что сломано и укладывается ли это в работу

Мультимодальный ассистент для документов и сканов

6 000 ₽

Извлечение нужных полей из фото и PDF, структурированный вывод под CRM или учёт

Интеграция и автоматизация процессов

3 000 ₽

n8n, Make, Albato, Google Таблицы, связь существующих сервисов между собой

Агент под ваш документооборот или расчёты

расчёт

Проектная работа с разбором процесса, замером «до», планом по неделям и приёмкой по критериям

после разбора

Цены стартовые, за базовый объём. Точная сумма называется после того, как я посмотрю задачу, — и до этого момента ничего не стоит. По проектной разработке ориентир ставки — около 1 400 ₽ в час, а оценка выдаётся в часах с расшифровкой по этапам, чтобы было видно, за что платите.

## // КАК ПРОХОДИТ РАБОТА

01

**Разбор задачи**

смотрю процесс и говорю, решается ли он и за сколько — бесплатно

&gt;

02

**Замер «до»**

сколько времени процесс занимает сейчас, чтобы потом было с чем сравнить

&gt;

03

**Сборка и сверка**

работа по неделям, промежуточные показы на ваших реальных данных

&gt;

04

**Приёмка**

по заранее записанным критериям, а не «на глаз»

&gt;

05

**Передача**

код, инструкция обычным языком и поддержка на срок тарифа

**Напишите — скажу честно, решается ли ваша задача**

Если задача не решается или решается дешевле другим способом, я так и скажу: братья за то, что не даст результата, невыгодно нам обоим. Разбор задачи и оценка — бесплатно.

ПОЧТА

1513@ro.ru

TELEGRAM

@vogelaa

САЙТ И ПРОЕКТЫ

aigaa.ru

МАРКЕТПЛЕЙС ЗЕРОКОДЕРА

marketplace.zerocoder.ru/zerocoder/21162