

ФОРУМ ИТ И ИННОВАЦИЙ В РИТЕЙЛЕ · МОСКВА, ЦЕНТР СОБЫТИЙ РБК · 3 СЕНТЯБРЯ 2026

Если ваш ИИ «врёт» и «забывает»...

Диагноз, схема лечения, профилактика — на опыте Аптечной сети «36,6»

9 методик, которые можно забрать с собой.

Алексей Еферов · ИТ-директор ГК «36,6»



Галлюцинации: источник у каждого ответа

Уверенная выдумка — свойство технологии; лечится дисциплиной, а не «моделью получше»

ПРОБЛЕМА

Выдумывает уверенно — так устроена:

училась без разбора, где правда



пробелы закрывает догадкой



«не знаю» = провал: обучена помогать



МЕТОД

- Ответ — только со ссылкой на источник; нет источника — честное «не знаю»
- Каждое утверждение — с маркером: ФАКТ · ГИПОТЕЗА · ВОПРОС
- Нет цитаты, документа или замера — утверждение вычёркивается
- Ответственные решения — всегда за человеком

КАК ЭТО У НАС — ДОСЛОВНО ИЗ ПРАВИЛ, КОТОРЫЕ ИИ ЧИТАЕТ КАЖДУЮ СЕССИЮ

из конституции проекта:

Маркировка уверенности в документах: **[ФАКТ]** (есть ссылка) · **[ГИПОТЕЗА]** (вывод без прямого подтверждения) · **[ВОПРОС]** (требуется проверки). Не выдумывать источники, цифры, практики без подтверждения.

из глобального свода правил:

Каждое фактическое утверждение несёт доказательство: дословная цитата (+пусть) / timestamp / замеренное число. Нет — пометь **[ИНТЕРПРЕТАЦИЯ]** или не утверждай.

ОТРИЦАНИЕ – ТОЖЕ УТВЕРЖДЕНИЕ, доказывается СИЛЬНЕЕ. «Не нашёл» $\neq$ «не существует».

Перед отправкой – перечитай каждое утверждение: «артефакт открыт? число замерено?». Нет – вычеркни.

ЗАБРАТЬ С СОБОЙ

Введите маркировку ФАКТ / ГИПОТЕЗА / ВОПРОС в промпты и регламенты.

ИИ работает не с данными, а с кодом

...который работает с данными: в диалог возвращается только результат — сами данные в чат не попадают



ПОЧЕМУ НЕ ЧАТ

- дорого: данные жгут токены
- невоспроизводимо: второй раз тот же ответ не получить
- модель в данных тонет — и начинает выдумывать
- окно диалога конечно: миллионы строк не влезут никогда

ЧТО ДАЁТ КОД

- **Воспроизводимо:** код перезапускается без ИИ — хоть каждую ночь по расписанию
- **Проверяемо:** код можно прочитать и отревьюить — у результата есть путь до источника
- **Масштаб не ограничен:** коду всё равно, тысяча строк или миллиард
- **Дёшево:** модель тратится на решение, а не на перекачку данных через себя

КАК У НАС

316 млн строк ценовой аналитики — глубокий анализ на локальной машине
Выгрузка из базы SQL с незнакомой структурой — от формата до полноценного обмена за 1 день

ЗАБРАТЬ С СОБОЙ

Чат — для решений; данные — коду.

Промпт устаревает и деградирует — живёт контекст

Выигрывает не тот, кто лучше просит, а тот, у кого лучше память



Внешняя память проекта

Длинный чат сжимается — модель молча теряет детали и ограничения, сказанные один раз.

- Конституция проекта — единый источник правды; перечитывается каждой сессией
- В проекте: реестры требований, гипотез, фактов, решений. Единая формализованная обвязка для проектов.
- В конце сессии — срез «что сделано, что дальше»: новая сессия стартует без потерь.
- Если знание шире, чем проект, — добавляется в Метазнание Компании.

ЗАБРАТЬ С СОБОЙ

Конституция, дословный реестр требований, срез в конце сессии.



Коллективная работа: git + общая база

Шаринг чатов не создаёт общего знания, а командные ИИ-тарифы дороги или недоступны.

- Итог — артефакт в git: документ, скрипт, презентация. Ответ в чате умирает вместе с чатом
- Общая база знаний — простые текстовые файлы под git: читают и пишут и люди, и ИИ
- Модели у каждого свои — правила, инструменты и база знаний одни на всех
- Личные подписки + общая файловая база = коллективный эффект без коллективного бюджета

ЗАБРАТЬ С СОБОЙ

Результат — артефакт в git, а не ответ в чьём-то чате.

Профилактика — это люди

Выдумку ИИ останавливает не фильтр, а постановка на входе и экспертиза на выходе



Постановщики и приёмщики

Код подешевел — дефицит сместился к людям, которые умеют спросить и проверить. Это и есть лучший фильтр от выдумки ИИ.

- Постановка: задача словами + критерий приёма — половина результата
- Приёмка: потребовать доказательство и проверить его — правдоподобное ещё не верное
- Экспертиза до, в процессе и после реализации — дефицит, который моделью не закрывается
- ИИ готовит — человек подписывает

ЗАБРАТЬ С СОБОЙ

Растите постановщиков и приёмщиков — лучший фильтр от выдумки.



Ошибка исправляется один раз

Правила, написанные один раз, мертвы: ИИ и люди повторяют одни и те же ошибки.

- Каждый сбой → правило: что случилось и как не повторить
- Правило — в свод, который ИИ перечитывает каждую сессию
- У правила — точка исполнения: без автопроверки оно остаётся пожеланием
- Догадка записывается как «гипотеза → тест → следствие» — иначе список тонет в «а вдруг»

ЗАБРАТЬ С СОБОЙ

Один свод правил: сб_{ой} → правило → свод.

Базовые принципы успешной работы с ИИ

Четыре установки, которые не меняются при смене модели

1

ИИ работает не с данными, а с кодом, который работает с данными!

Данные в чате — дорого и невоспроизводимо: модель в них тонет и начинает выдумывать. Агент пишет код, код обрабатывает данные — в диалог возвращается только результат.

2

Модель не так важна, как обвязка!

Данные, интеграции, контроль и дисциплина вокруг модели решают больше, чем выбор LLM. Модели меняются каждый квартал — обвязка остаётся вашей.

3

Не просто данные — а цифровой двойник Компании!

Любая информация ценна! Каждое действие несёт цифровой след. Встречи, решения, реестр фактов — складываются в цифрового двойника, с которым ИИ обязан сверяться. Ответ без опоры на двойника — мнение.

4

Профилактика — это люди: постановка на входе, экспертиза на выходе.

ИИ не отличает правдоподобное от верного — это делает человек, знающий предмет. Код подешевел, экспертиза подорожала: растите постановщиков и приёмщиков.

Код дешёв. Ценность — в экспертизе и Time-To-Market.

РАБОЧИХ ПРОЕКТА ЗА 11 НЕДЕЛЬ

23

от рыночной аналитики до интеграции при покупке сети — внутренней командой с ИИ

ЦЕНОВАЯ АНАЛИТИКА ПО РЫНКУ

316 млн строк

массив обрабатывает скрипт, который написал ИИ; в чат не ушла ни одна строка

ПРОЕКТОВ ПОД
АВТОПРОВЕРКАМИ

30 из 30

скан секретов и автопроверки: непроверенное изменение просто не проходит

ПРИЛОЖЕНИЕ ДЛЯ АПТЕК ПО
КОНТРОЛЮ ВНЕШНЕГО СЕРВИСА

2 недели

собрал один человек с ИИ

ПОЛНЫЙ РАЗБОР НЕЗНАКОМОЙ
УЧЁТНОЙ СИСТЕМЫ

1 день

карта метаданных 1С приобретённой аптечной сети — без конфигуратора и ресурсов продавца

ПРИЛОЖЕНИЕ ПОИСКА
ПОМЕЩЕНИЙ ПОД АПТЕКИ

**63 версии за
7 недель**

полевые менеджеры считают экономику точки на выезде; модель откалибрована на всей сети; сделал один человек с ИИ.

КРИТИЧНЫХ ОШИБОК ПОСЛЕ
ВЫПУСКА — ТРИ РЕЛИЗА
КАССОВОГО ПО ПОДРЯД

0

~3 000 аптек, июль—август: ни одного Highest-бага после выпуска и ни одного внепланового релиза-фикса (ИИ-ревью с апреля).

КОРПОРАТИВНЫЙ ИИ-ПОРТАЛ

2,5 месяца

Совместная работа над проектами, маскирование данных, локальные модели — в изолированном контуре - 52 требования ИБ реализованы. С помощью ИИ)

Аптечка: заберите с собой

Три рецепта — по одному на каждую часть диагноза



РЕЦЕПТ ОТ «ВРЁТ»

- Ответ — только с источником; маркировка ФАКТ · ГИПОТЕЗА · ВОПРОС
- Один файл фактов: число — строкой, у строки источник и статус
- Вывод «стало лучше» без метрики не принимается



РЕЦЕПТ ОТ «ЗАБЫВАЕТ»

- Конституция проекта + дословный реестр требований
- Срез в конце сессии — следующая стартует без потерь
- Итог действия — артефакт в git; у документа — одна роль



ПРОФИЛАКТИКА

- Шлагбаумы вместо просьб: отказ получает агент, а не вы
- Сбой → правило → свод; каждому правилу — точка исполнения: хук, линтер, гейт
- Растите постановщиков и приёмщиков

«Доверие к ИИ строится не вычитыванием каждой строчки, а правильно выстроенной проверкой»



eferov.pro

материалы доклада и контакты