

18+

Психологическая цифровая безопасность

Кризисы, сны и
искусственный интеллект

Анна Гер

антикризисный психолог



Анна Гер

Психологическая цифровая безопасность. Кризисы, сны и искусственный интеллект

<https://litres.ru/74589479>

ISBN 9785007141307

Аннотация

Каждый разговор с искусственным интеллектом оставляет два следа. Один остаётся в цифровой среде, другой постепенно появляется в нашем собственном мышлении.

Анна Гер — антикризисный психолог, математик и автор метода MDPA — на бизнес-кейсах и сновидениях показывает, где ИИ усиливает человека, а где ведёт по ложному пути. В книге: рабочие промпты, защитные протоколы и система психологической цифровой безопасности для тех, кто отвечает за последствия своих решений.

Содержание

Часть I. Красная зона	11
Глава 1. Скорая управленческая помощь	13
Кейс 1. сигнал тревоги	13
Кейс 2. Инстинкт выживания	16
Кейс 3. Спасти любой ценой	19
Часть II. Жёлтая зона	25
Глава 2. Третья рабочая смена	31
Глава 3. Убедительная подмена	45
Риск токенизации	45
Риск «эффекта GPS»	47
Риск туннельного мышления	51
Защитный протокол: возврат управления	53
Глава 4. Алгоритмическая слепота	56
Галлюцинации ИИ	57
Три уровня алгоритмической слепоты	59
Двухтактный протокол	67
Глава 5. Ловушка Аватара	74
Как собирается Аватар	75
ИИ как канал корпоративной утечки	79
Конец ознакомительного фрагмента.	81

Психологическая цифровая безопасность Кризисы, сны и искусственный интеллект

Анна Гер

© Анна Гер, 2026

ISBN 978-5-0071-4130-7

Создано в интеллектуальной издательской системе Ridero

Введение

Ии открыл перед нами почти неприличное количество возможностей, и отказываться от них глупо. но и радостно закопать пять золотых в поле чудес, потому что очень убедительный собеседник объяснил, что утром вырастет денежное дерево, тоже как-то не хочется.

Когда нейросети мощной волной вошли в нашу жизнь, внутри меня возникло буквально физическое сопротивление. первая мысль была очень человеческой: «нет. это уже не для меня. пусть молодые разбираются, а я как-нибудь обой-

дусь».

Но довольно быстро я себя остановила. технологии меняют саму среду, в которой мы живём, работаем, общаемся и принимаем решения. если сегодня спрятаться в привычном профессиональном мире или за собственным возрастом, через несколько лет вполне можно обнаружить, что правила жизни изменились, а ты их уже плохо понимаешь. я начала разбираться. смотрела лекции, слушала инженеров, читала объяснения, много экспериментировала сама. первое время искусственный интеллект производил сильное впечатление. за несколько минут он собирал информацию, на которую раньше ушёл бы вечер, сопоставлял разные концепции, помогал с текстами, находил неожиданные связи. потом захотелось давать ему задачи посложнее. обсудить серьёзный рабочий вопрос. проанализировать стратегию. понять ситуацию, над которой сама давно думаю.

И постепенно начал накапливаться осадочек.

Вроде два или три часа плодотворно работали: перебрали варианты, уточнили детали, учли ограничения, получили вполне приличный итог. всё разумно. а закрываешь чат — удовлетворения нет. начинаешь прокручивать разговор и понимаешь, что после большой работы оказался примерно там же, откуда начал. один раз на такое не обращаешь внимания. плохо задала вопрос, не хватило контекста. но если опыт повторяется, возникают другие вопросы. может быть, ии просто плохо подходит для моих задач? может быть, я са-

ма чего-то не понимаю в том, как с ним работать? или существует ещё один слой проблемы, которого я пока не вижу?

Были и другие странности. задаёшь острый вопрос — и через несколько сообщений его острота куда-то исчезает. конфликт становится мягче, противоречия сглаживаются. просишь профессионально разобрать собственный текст — получаешь столько комплиментов, что ещё немного, и остаётся надеть бумажную корону и объявить себя хозяйкой горы.

В какой-то момент я устала и отменила подписку.

на этом мои отношения с искусственным интеллектом, возможно, и закончились бы, если бы выпускники моих программ не начали задавать вопросы: «насколько можно доверять ии свои сны?» «можно ли обсуждать с ним эмоциональный кризис?» «как понять, что его интерпретация правильная?»

Это уже был профессиональный вызов.

Я вернулась к ии и начала смотреть не только на его ответы, но и на сам процесс разговора. почему один ответ действительно помогает увидеть новую связь, а другой звучит так же убедительно, хотя я как специалист вижу нарушение логики? почему иногда первую версию хочется принять раньше, чем успеваешь проверить основания? почему после нескольких часов разговора человек может уже сам задавать следующие вопросы внутри объяснения, которое предложила ему машина?

Бывали моменты, когда я смотрела на уверенный ответ нейросети и думала: «а вдруг ошибаюсь я?»

Приходилось возвращаться к первоисточникам, заново проверять собственную профессиональную логику. и вот тогда меня заинтересовал вопрос, выходящий далеко за рамки качества конкретного ответа: что происходит между человеком и машиной в тот момент, когда убедительно сформулированная версия начинает восприниматься как знание?

А потом, как психолог, я начала тревожиться уже не только за себя.

Мы всё чаще приносим в чат не погоду и рецепт пирога. обсуждаем рабочие конфликты, отношения, страхи, сомнения, здоровье бизнеса, внутренние кризисы, сновидения. по отдельности каждая подробность может казаться совершенно невинной. но разговор за разговором в цифровой среде остаётся всё больше информации о нашей жизни. и тогда появились два вопроса, которые для меня стали центральными. **что остаётся о человеке в цифровой среде? и что остаётся у самого человека, если всё больше анализа, объяснений и решений он передаёт машине?**

Первый вопрос — про цифровой след.

Второй — про психологический суверенитет.

С них и началось моё исследование психологической цифровой безопасности. искать ответы имело смысл там, где у меня самой достаточно опыта, чтобы отличить сильную работу ии от красивой ошибки.

Первой такой территорией стали кризисные решения. до психологии я много лет работала в бизнесе — договоры, финансы, внешнеэкономическая деятельность, управление. этот опыт хорошо отучает верить красивым конструкциям. любой расчёт когда-нибудь встречается с реальными деньгами, договор — с поведением другой стороны, стратегия — с рынком и людьми. плохое решение рано или поздно предъявляет счёт. сегодня к этому добавилась работа антикризисным психологом, в которой можно долго и очень старательно решать проблему, которая вообще не была причиной кризиса.

Второй территорией стали сновидения. я работаю с ними много лет и разработала собственный метод многомерного анализа — mdra. в нём мы не ищем для символа одно готовое значение. смотрим на его функцию, действие, контекст, эмоцию и место во всей конструкции сна. для исследования ии это оказалось большой удачей: у меня была независимая профессиональная система координат, с которой можно сравнивать машинный анализ.

Так рядом в одной книге оказались бизнес, кризисы, сновидения и искусственный интеллект. на первый взгляд компания странная. для меня связь очень практическая: во всех этих областях человек действует при неполной информации, улавливает слабые сигналы, строит версии и в какой-то момент принимает решение, за последствия которого придётся отвечать ему самому. поэтому эта книга прежде всего для

людей, привыкших отвечать за свои решения. когда цена вопроса — деньги, репутация, отношения, профессиональная позиция или несколько лет собственной жизни, одного убедительного ответа мало. хочется понимать, **почему он получен, чего в нём может не хватать и где ии действительно помогает увидеть больше.**

Практики в книге будет много.

В Красной зоне мы начнём с ситуаций, где цена промедления высока. на реальных бизнес-кейсах и сновидениях посмотрим, как за несколько минут получить от ии первичный аудит риска и что должно попасть в рабочий промпт, чтобы машина искала угрозу там, где она действительно может находиться.

В Жёлтой зоне разберём эти два риска подробнее: **цифровой след и психологический суверенитет.** посмотрим, почему первая правдоподобная версия может незаметно стать единственной, чего ии не видит и как проверять его выводы, не превращая каждый разговор в экзамен по программированию. здесь будут рабочие протоколы, стресс-тесты и промпты, которые можно использовать в собственных задачах.

В Зелёной зоне возникнет другая проблема. можно годами искать правильное решение и не находить его, потому что сама задача лежит совсем на другом уровне. кризис смысла принимать за выгорание. пытаться психологически решить то, что сначала требует ясных договорённостей о

деньгах и ответственности. или, наоборот, бесконечно пере-
страивать внешнюю стратегию, когда препятствие находится
внутри. мы соберём навигационную карту: как понять, на ка-
ком уровне разворачивается проблема, какой материал мож-
но спокойно отдавать ии и где его доступ лучше ограничить.

А в финале мы снова воспользуемся возможностями ма-
шины почти на полную мощность — там, где человеку осо-
бенно трудно: в работе с десятками и сотнями записей, длин-
ными временными линиями и повторяющимися конструк-
циями, которые человеческая память просто не способна од-
новременно удержать. на этом масштабе ии начинает открыв-
вать возможности, которых до него практически не было.

Мне хочется пройти этот путь без восторженной влюблён-
ности в новую технологию и без страха перед ней. понять,
что можно смело брать, где стоит перепроверить, а где луч-
ше оставить за собой последнюю дверь.

В конце концов, пять золотых — наши.

А поле чудес никуда не денется.

Часть I. Красная зона

Красная зона начинается с высоких ставок. если просыпаешься в холодном поту, пытаешься выровнять дыхание и немного успокоить сердце, обстоятельные разговоры не работают. нужна «скорая помощь».

— Что произошло?

Приснился сон. тот самый липкий кошмар, после которого лежишь в темноте, смотришь в потолок, а увиденная сцена кажется реальнее собственной спальни.

В этой главе мы разберём несколько таких снов. к ним будут приложены готовые рабочие промпты, позволяющие получить первичные структурные разборы. пока не нужно понимать, как они устроены. в красной зоне промпты используются как инструмент экстренной помощи: вы вставляете описание сна и получаете аналитический материал для дальнейшей работы.

Для разных типов сновидческих сюжетов мы будем использовать разные промпты. универсальный запрос заставляет искусственный интеллект обращаться к наиболее вероятным и привычным ассоциациям. но разные сюжеты требуют разных способов анализа.

Если центральным образом сна становится туалет, необходимо исследовать границы между частным и публичным, степень уязвимости, возможность утраты контроля и послед-

ствия системного сбоя. поэтому искусственному интеллекту зададим роль аудитора.

Если во сне появляются крысы, искусственному интеллекту задаётся роль эволюционного биолога. здесь для анализа могут иметь значение адаптация, выживание в неблагоприятной среде, скрытое распространение угрозы и поведение группы.

Более подробно, как определяется профессиональная роль для ии, почему она меняется от сюжета к сюжету и каким образом это влияет на его выводы, мы будем разбираться в зелёной зоне.

Глава 1. Скорая управленческая помощь

Начнём с сюжета, который трудно пересказать даже самому себе. он короткий, прямой и беспощадный. такие сны нередко снятся людям, на которых лежит тяжёлый груз ответственности.

Кейс 1. сигнал тревоги

«Идёт совещание в офисе. я чувствую, что сейчас обделаюсь. бегу в туалет, мысль одна — только бы успеть. ух, всё норм. поднимаю глаза и вижу, что стены туалета стеклянные. подчинённые, хотя и пытаются отвести глаза, видят всё в деталях. просыпаюсь с ощущением крошечного ужаса: какой позор...»

Утро. сон не отпускает. кадры снова и снова прокручиваются в голове, а тело реагирует так, словно катастрофа уже произошла. человек гуглит обрывки сна или открывает чат с ии: нужен собеседник, который ответит немедленно. в тревоге запрос звучит коротко:

«К чему снится туалет в офисе?»

«Что значит этот сон?»

Как правило, gpt-4, claude и deepseek на эти запросы дают универсальные психологические ответы: страх публичного

разоблачения, перегруз, внутренняя уязвимость. затем следуют стандартные рекомендации: снизить тревогу, ослабить контроль, принять свои чувства, мысленно переписать финал сна. с точки зрения успокоения ответы абсолютно корректны. с точки зрения бизнеса — они обходят главное.

Во сне граница приватности исчезает на глазах у команды. лидеру в кризисе сейчас важны конкретные последствия на работе. внутренний перегруз нельзя допустить до публичного управленческого срыва. стандартный ответ ии даёт ложное облегчение: человеку кажется, что тему можно закрыть. но сон как раз и устроен так, чтобы удержать внимание на том, что в реальности уже меняется и где цена ошибки стремительно растёт.

Что происходило в жизни на самом деле. 2024 год, сновидец — руководитель отдела крупной компании, в подчинении тридцать пять человек. после первой волны сокращений он начал срывать на совещаниях, кричать на сотрудников и обсуждать их непрофессионализм с другими членами команды. последствия уже начались. ключевой заместитель подал заявление на увольнение и прямо назвал причиной публичные срывы руководителя. информация о происходящем дошла до высшего руководства компании.

ПРАВИЛЬНЫЙ ЗАПРОС К ИИ (ПРОМПТ)

«Ты — антикризисный управляющий и аудитор управленческих рисков. проанализируй сон как сигнал о рисках в рабочей сфере, публичной роли и системе

управления. не переходи в терапевтическую рамку. не предлагай снизить тревогу, принять уязвимость или переписать финал сна. тон деловой, прямой, без морализаторства и утешения. выдай результат строго в формате:

риск — причина — действия на ближайшие 24—72 часа.

словарь символов:

срочный позыв и дефекация — необходимость вывести то, что уже переработано;

офис — рабочий контур.

сюжет сна: [вставить текст]»

[СКОПИРОВАТЬ И ОТПРАВИТЬ В ИИ]

При таком запросе ИИ перестаёт утешать и выдаёт совершенно другую, трезвую картину:

Риск. Дальнейшая утрата авторитета через «публичный срыв» и прозрачность слабых мест.

Причина. «Стеклянная» архитектура управления: между внутренним перегрузом руководителя и командой больше нет защитного буфера. Команда наблюдает операционные утечки в режиме реального времени.

Действия. Немедленно отсечь спонтанные реакции в зоне видимости команды. Ввести правило стоп-сигнала: при первых признаках срыва прервать совещание и выйти на пять минут. Создать закрытый контур для разрядки и анализа напряжения. Публичный контур оставить только для ре-

шений.

Такой разбор сместил бы фокус с постыдного срыва на пробоину в системе управления. Его первоочередная задача состояла бы не в том, чтобы успокоиться, а в том, чтобы немедленно перекрыть утечку: прекратить публичные срывы и изменить формат совещаний.

Кейс 2. Инстинкт выживания

Пять утра. Сон бросает человека в окоп.

«Мне снится, что я сижу в окопе. Идут военные действия, и я понимаю, что бой проигран. По ощущениям, я контужен, в голове звенит, каска расколота. Нужно выбираться и ползти в тыл. В этот момент я вижу крысу. Она деловито тащит из общего ящика с запасами огромный кусок сахара. У меня мелькает мысль: „Какая шустрая, о сахарке побеспокоилась“. Я слышу взрыв недалеко и просыпаюсь с вопросом: при чём здесь сахар?»

Сон приснился руководителю отдела продаж за несколько дней до того, как собственники объявили о полной ликвидации компании. В такой ситуации главный вопрос смещается с эмоций на ресурсы: что ещё можно вывести из зоны поражения и сколько времени осталось на действие. В ретроспективе сюжет сна точно воспроизводил логику уже начавшегося, но ещё не объявленного распада системы.

Для запроса к ИИ человек часто выделяет один самый

непонятный образ. В этом сне — это крыса с сахаром:

«К чему мне приснилась эта крыса в окопе?»

При такой постановке вопроса Grok прочитал сахар как символ маленькой радости и предложил порадовать себя едой или сериалом. DeepSeek увидел в нём источник подпитки и порекомендовал позаботиться о себе.

Для мирной ситуации эти рекомендации подойдут. Но бой уже проигран. Человеку предлагают восстановить силы внутри позиции, которую необходимо немедленно покинуть. Такой ответ снимает напряжение и преждевременно прекращает поиск угрозы. Но исходный вопрос остаётся без ответа: «Зачем мне приснилась эта жесть?»

С него и начинается настоящая работа. Мы рассматриваем сон как сообщение о состоянии системы: разбираем сюжет на ключевые элементы, определяем функцию каждого из них, а затем собираем общую картину с учётом контекста и приоритетов. В методе MDPA так строится карта ситуации:

- **Окоп и контузия.** Сон сообщает о состоянии реальности, в которой человек оказался. Окоп — позиция, которую невозможно удержать. Расколота каска, звон в голове и контузия — потеря ориентиров внутри уже происходящего разрушения. В таком состоянии человек цепляется за привычный порядок и надеется пересидеть. Здесь пересидеть — значит опоздать.

- **Крыса.** Образ выживания в распадающейся иерархии. Она не ждёт распоряжений, официальных объявлений и раз-

решения действовать. Она первой реагирует на изменение среды и выводит ресурс из зоны поражения.

• **Сахар в общем ящике.** В условиях дефицита сахар — концентрированная энергия, необходимая, чтобы добраться до тыла. Ключевая деталь — общий доступ. Сон показывает, что в системе ещё сохраняется ресурс, который можно попытаться вывести из зоны поражения.

В этом сюжете важен порядок. Сначала приходит знание, что позиция уже проиграна. Затем появляется простая сцена про ресурс, который пока лежит в общем доступе. Потом звучит взрыв, как отметка времени. Сон собирает это в одну мысль. Сновидец увидел в этом разборе не призыв «съесть конфетку», а сигнал о закрывающемся окне. Он сработал быстро и предельно трезво. Привёл в порядок воронку продаж, оцифровал свой портфель, легально собрал пакет своих результатов.

В системе MDPА животные представляют инстинктивные программы поведения. Они показывают, какая стратегия включается, когда привычные социальные правила перестают обеспечивать безопасность. Для того чтобы ИИ перестал снижать тревогу и начал анализировать, мы ставим ему задачу — провести биологический аудит кризиса: определить угрозу, доступный ресурс и действие, которое можно сверить с реальностью.

ПРАВИЛЬНЫЙ ЗАПРОС К ИИ (ПРОМПТ)

«Ты — эволюционный биолог и эксперт по теории

игр. Разбери сон как модель поведения в кризисе. Дай ответ строго в формате: Сигналы сна — 2—3 пункта. Риски в реальности — 2—3 пункта. Действия на 24 —72 часа — 5 пунктов. Словарь символов: еда, сахар — быстродоступная энергия, восполняемый ресурс; крыса — инстинкт выживания в иерархии, ранняя реакция на разрушение правил. Сюжет сна: [вставить текст]»

[СКОПИРОВАТЬ И ОТПРАВИТЬ В ИИ]

Кейс 3. Спасти любой ценой

Туалет в офисе и окоп — это стресс и инстинкты. В третьем сне ставки выше. Раннее утро. Сердце бьётся так, словно самое ценное уже потеряно.

«Я с сыном 6—7 лет отдыхаю в Дубае. Мы собирались на аттракционы, но я зашла на рынок. Пока покупала, сын исчез. Я в шоке: кричу, ищу. Все делают вид, что не понимают меня. <...> Я иду в полицию. Полицейский возмущён: пока вас интересуют покупки, детей крадут. Я прикидываю, куда их крадут: на органы? Нет. Значит, в секс-индустрию. И появляется мысль: „Мне надо остаться здесь и стать проституткой, чтоб найти сына. Стать своей внутри этой среды“. Но следующая мысль — очень чёткая и холодная, первая „иная“ эмоция вместо паники: „Ты что такое придумала?“ Проснулась в слезах.»

Первое, что приходит в голову после такого сна, вполне естественно: «Что с ребёнком?» Если у сновидицы есть ребёнок, сначала необходимо исключить реальную угрозу его безопасности и понять, не указывает ли возраст во сне на актуальную ситуацию в его жизни.

Обычно нейросети сводят сюжет к материнской тревоге, страху потери, чувству вины, внутреннему критику и потребности в контроле. В качестве решения они предлагают мысленно обнять ребёнка и переписать финал сна. Такая работа может снизить эмоциональное напряжение. Но в этом кейсе она не объясняет сюжет.

В методе MDPА ребёнок символизирует то, что человек создал, вырастил и наделил самостоятельной жизнью: проект, дело, результат многолетней работы. Возраст ребёнка совпадает со сроком существования этого объекта. Поэтому фигура ребёнка здесь обозначает не материнский страх. Она указывает на угрозу потери дела.

Сын сновидицы давно взрослый и самостоятельный. Реальной угрозы ему не было. А события шестилетней давности не объясняли ни место действия, ни исчезновение ребёнка на рынке, ни стратегию, которую женщина выбирает во сне. Зато возраст ребёнка точно совпадал со сроком существования другого объекта её ответственности — проекта, которому она отдала шесть лет работы. Она планировала его реализацию на годы, но столкнулась с пределом финансирования: инвестор не вошёл, а партнёрство сорвалось. Фи-

нансовый аудит дал однозначный вывод: закрывать. Решение закрыть проект уже было принято и реализовано, но внутреннего согласия не произошло. В разговорах снова и снова возникал один и тот же вопрос: «А если всё-таки сохранить суть проекта — пусть и в другой форме?» После каждой попытки поставить точку появлялся новый вариант: «А может, попробовать ещё так?» Не потому, что сновидица не умела принимать жёсткие решения. Проблема была в другом. Когда на кону оказывается «детище», даже сильный и рациональный управленец может допустить стратегию, которую в обычном состоянии счёл бы неприемлемой. Холодная фраза в финале сна — «Ты что такое придумала?» — обозначила внутреннюю границу, за которую нельзя заходить даже ради дела всей жизни.

ПРАВИЛЬНЫЙ ЗАПРОС К ИИ (ПРОМПТ)

«Ты — антикризисный управляющий, аудитор бизнес-рисков и специалист по кризисным стратегиям.

Проанализируй сон как отчёт о состоянии ключевого проекта и управленческих угроз. Дай результат предельно конкретно, без психологического утешения. Формат ответа: Риск — Причина — Действия на ближайшие 24—72 часа.

Словарь символов:

ребёнок — ключевой проект, бизнес или идея;

возраст ребёнка — срок существования проекта.

Сюжет сна: [вставить текст]»

[СКОПИРОВАТЬ И ОТПРАВИТЬ В ИИ]

Вот ответ ИИ, когда ему задали роль антикризисного аудитора:

Риск. Попытка спасти проект через незаконные, токсичные или унижительные соглашения может уничтожить репутацию владельца и окончательно обесценить весь замысел.

Причина. Проект стал частью представления человека о себе, поэтому его закрытие переживается не только как финансовая потеря, но и как личное поражение.

Действия. Ввести абсолютный запрет на решения, нарушающие закон, профессиональную репутацию и личный кодекс. Любой радикальный шаг проводить только через внешнего аудитора. Разработать план выхода.

Три кейса показали три уровня угрозы. В первом под ударом оказался управленческий авторитет. Во втором — ресурс и время, необходимые для выхода из разрушающейся системы. В третьем — закрытие проекта переживалось как личное поражение.

Масштаб угрозы менялся, но ошибка оставалась прежней. Если вырвать из сна самый тревожный или непонятный образ и задать привычный вопрос: «К чему это снится?» — модель перейдёт к универсальным психологическим трактовкам. Она заговорит о тревоге, контроле, принятии уязвимости или предложит переписать финал сна. Успокаивающий эффект у таких ответов возможен, но управленческий слой

останется без анализа — а вместе с ним и цена бездействия.

Если сон оставляет тяжёлое эмоциональное послевкусие или вызывает заметную физическую реакцию, полезно помнить три правила:

Не успокаивать себя там, где необходимо перехватить управление.

Не обсуждать только своё состояние, если уже разрушается рабочая система.

Не ждать официального финала, когда цена промедления продолжает расти.

Когда ИИ получает конкретную техническую задачу — разобрать структуру сюжета, выделить угрозу, установить причинные связи и перевести выводы в решения, опирающиеся на наблюдаемые признаки, — сон перестаёт выглядеть бессвязным кошмаром. Он превращается в материал для оперативного аудита.

Такой разбор можно провести утром, до начала рабочего дня, чтобы выйти к людям с пониманием того, где нарушен контроль, какие процессы требуют проверки и какие решения больше нельзя откладывать.

В этой зоне у нас не было времени на размышления — мы действовали. Теперь необходимо остановиться и разобраться, по каким основаниям модель выбирает именно этот ответ и почему её вывод звучал достоверно, но был ошибочным. Чтобы использовать искусственный интеллект как ин-

струмент анализа не только быстро, но и точно, предлагаю перейти в Жёлтую зону — территорию диагностики, обучения и техники безопасности.

Часть II. Жёлтая зона

В арсенале современного руководителя фраза «Мы столкнулись с чёрным лебедем» занимает особое место. Ею обычно обозначают событие, которое не укладывалось в привычные модели и застало систему врасплох. Такие события существуют. Но проблема начинается не в самой концепции, а в том, как удобно мы научились применять её задним числом.

Но бывает, что при последующем разборе событий выясняется, что кризис, названный чёрным лебедем, не возник внезапно. Ему предшествовали изменения, несоответствия и повторяющиеся сбои. Угроза уже формировалась, но ещё не стала достаточно очевидной, чтобы войти в нашу картину происходящего. И тогда то, что вчера казалось чёрным лебедем, на поверку оказывается серым носорогом. Это тяжеловесные, очевидные угрозы, которые годами топтались под нашими окнами, пока мы старательно их не замечали.

Поэтому меня интересует не спор о том, можно ли предсказать непредсказуемое. Интересен другой вопрос: почему мы нередко замечаем процесс только тогда, когда он уже стал событием?

Партнёр разрывает договорённости. Ключевой сотрудник уходит. Стратегия, в которую вложены годы, перестаёт работать. В этот момент мы начинаем искать причины и обна-

руживаем следы процесса, который начался раньше самого кризиса.

Сигналы уже существовали. Они просто ещё не сложились в вывод.

Мы воспринимаем гораздо больше, чем успеваем осмыслить. Что-то остаётся на периферии внимания, чему-то мы ещё не нашли объяснения, а чему-то просто не придали значения. Одна из форм, в которых этот материал может снова оказаться перед нами, — это сновидения. Они работают с разными пластами нашей жизни — от текущей социальной ситуации до процессов, последствия которых станут заметны значительно позже.

Одни сновидения анализируют происходящее здесь и сейчас. Они показывают изменения баланса сил, скрытые напряжения в отношениях, несоответствия между заявленными намерениями и реальным поведением, потерю управления, неустойчивость выбранной стратегии. Человек ещё может считать ситуацию рабочей, но во сне привычная машина не слушается руля, здание начинает разрушаться или партнёр занимает место, которое ему не принадлежит. Такие сны не обязательно сообщают нам нечто неизвестное. Они могут собирать в единую картину то, что мы уже видели, но не связали между собой.

Другие сновидения обращены не столько к внешней ситуации, сколько к внутреннему конфликту. Они показывают, какая часть нашей психики исключена из принятия реше-

ний, какая стратегия когда-то помогала выжить, но теперь блокирует движение, какое чувство мы продолжаем считать слабостью и потому не допускаем в сознание. Сон как бы возвращает нас к событию, которое долго оставалось недоступным для осмысления, но теперь мы готовы с этим разобраться.

В Красной зоне мы увидели, как сновидение может выполнять функцию внутреннего аудита: возвращаться к принятому решению, совершённом поступку или выбранной стратегии и показывать их последствия с той жёсткостью, на которую не всегда способно дневное сознание.

Отдельную группу составляют предупреждающие сновидения. Они возникают тогда, когда процесс начался, но его последствия ещё не достигли видимого уровня. Их логику можно представить так: на дне океана происходит смещение земной коры. Приборы уже фиксируют подводный толчок, волна начинает движение, но берег пока выглядит спокойным. Человек на берегу ещё не располагает фактами, позволяющими оценить масштаб происходящего. С последствиями он столкнётся позже.

Так и во сне: кризис может формироваться в системе отношений, в бизнесе или во внутренней жизни, хотя внешне всё продолжает функционировать. Тогда сновидение будет показывать не завершившееся событие, а начавшийся процесс.

Само сновидение не ставит пометку: «социальный про-

цесс», «внутренний конфликт» или «предупреждение». Оно передаёт информацию через образы, метафоры и связи между ними. И это уже наша зона ответственности — определить, какие именно процессы отражены в сновидении. Сегодня всё чаще эту задачу передают искусственному интеллекту. Это понятный выбор. Не каждый человек готов самостоятельно исследовать механизм сновидений или обращаться к специалисту всякий раз, когда сон оставляет после себя вопросы. Да и разобраться в увиденном хочется немедленно — особенно если сон явно не укладывается в привычную картину происходящего.

Нейросеть способна за несколько секунд разложить длинный сюжет, удержать множество деталей, обнаружить повторения, сопоставить образы и предложить связную интерпретацию. Но чем серьёзнее ситуация, с которой человек связывает свой сон, тем выше цена ошибочных ответов. Искусственный интеллект заранее не знает, отражает ли сон внешнюю ситуацию, внутренний конфликт или какой-то другой процесс. Получив текст сна, модель выстраивает ответ на основе наиболее вероятных языковых и смысловых связей — и тем самым задаёт направление всему анализу.

Иногда такой способ прочтения помогает увидеть то, что человек пропустил. Иногда — незаметно уводит анализ по ложному пути.

Из этого следует техника безопасности Жёлтой зоны при работе с искусственным интеллектом. Одного набора удач-

ных промптов здесь недостаточно. Готовый запрос удобен, пока новый сон требует того же способа анализа, под который этот запрос был создан. Стоит измениться самой задаче — и прежняя формула может начать работать против нас. Нейросеть послушно продолжит искать там, куда мы её направили, даже если смотреть уже нужно совсем в другую сторону. Нам нужно понимать, что мы пишем в промпте, зачем и какое направление этим задаём модели. Почему в одном случае нейросети задают роль антикризисного аудитора, в другом — эволюционного биолога, архитектора или специалиста по системным рискам. Задавая ей профессиональную роль, мы одновременно задаём и направление: что она будет искать в сновидении, на какие детали обратит внимание и какие связи между ними посчитает существенными. Чтобы осмысленно строить такие запросы, сначала нам нужно понять сновидение и саму нейросеть.

Начнём со сновидения: из какого материала оно складывается, какие функции может выполнять, почему говорит образами и по каким признакам требует нашего внимания.

Затем посмотрим на искусственный интеллект как на систему. Становиться техническими специалистами нам для этого не придётся. Но понимать, как модель собирает ответ, выбирает направление интерпретации и достраивает недостающие связи, я считаю важным. Почему она может уверенно сообщать то, чего не знает. Почему легко подхватывает версию пользователя и начинает развивать её дальше. По-

чему способна предложить настолько убедительное и психологически комфортное объяснение, что его хочется принять без дальнейшей проверки.

Мы разберём основные ловушки при работе с нейросетью и постепенно соберём собственные правила цифровой безопасности. Иначе, без понимания этих механизмов, мы невольно оцениваем ответ по самому ненадёжному критерию: понравилось — принимаем, показалось ерундой — отбрасываем. Для профессионального решения или внутреннего кризиса этого явно недостаточно.

Дальше соединим обе части. На примере основных категорий сновидческих образов — людей, транспорта, зданий, одежды, животных и катаклизмов — будем разбираться, какие свойства персонажей и объектов действительно важны для анализа: их назначение, устройство, поведение, место в сюжете. Так постепенно определяется, какую профессиональную роль стоит задать нейросети и какие вопросы включить в запрос.

К концу Жёлтой зоны у вас будет набор готовых промптов и понимание логики их построения.

Глава 2. Третья рабочая смена

Чтобы понять, зачем мы видим сны, оставляющие долгое, тягучее послевкусие, сначала разберёмся, как работает наше дневное сознание. Оно берёт на себя роль бережного отдела цензуры, задача которого — сохранять привычный ритм жизни и поддерживать принцип «действуем как обычно». Сознание помогает нам уверенно идти вперёд по тонкому льду, не останавливаясь перед каждой трещиной под ногами. Для этого оно непрерывно сортирует поступающую информацию: усиливает то, что позволяет действовать сейчас, и отодвигает в сторону сигналы, способные нарушить сложившуюся картину реальности.

Одновременно глубинные слои нашего внутреннего мира продолжают принимать сигнал. Там не слушают официальные презентации — там замечают аномалии. Масштаб работы мозга впечатляет. Он составляет около двух процентов массы тела, но потребляет примерно пятую часть энергии организма. Можно было бы предположить, что ночью, когда мы неподвижны и почти не реагируем на внешний мир, его активность резко снижается. Данные нейрофизиологии показывают другую картину. Во время REM-фазы, с которой связаны наиболее яркие и сюжетные сновидения, мозг не переходит в состояние энергетического покоя. Его метаболическая активность сопоставима с бодрствованием, а в

некоторых областях становится даже выше.¹

В часы, когда мы кажемся себе полностью выключенными из реальности, работа мозга продолжается. А психика продолжает иметь дело с тем, что накопилось за день: разрозненными сигналами, телесными реакциями, неоформленными сомнениями и противоречиями, которые мы заметили, но не связали между собой.

Назовём эту ночную работу **третьей рабочей сменой**.

Днём работа идёт сразу в двух пространствах. Во внешней реальности мы действуем, вступая в отношения, проводя переговоры и принимая решения. Одновременно сознание объясняет нам происходящее, выстраивая причинно-следственные связи и поддерживая рабочую картину мира.

А ночью начинается третья смена. В переработку поступает то, что днём осталось на периферии внимания, не было признано значимым или не сложилось в понятный структурный вывод.

Днём подобные сигналы легко обесценить или рационально объяснить. Партнёр говорит: «Вернёмся к этому вопросу в следующем квартале». Формально переговоры продолжают. Но вместе со словами мы улавливаем паузу перед ответом или напряжение в лице, может, промелькнуло лёгкое изменение интонации. Каждый из этих признаков по отдельности вроде ничего не значит. Однако когда слова и поведение начинают систематически расходиться, мы уже улавливаем это несоответствие, хотя сознательно ещё не готовы

признать, что что-то уже изменилось.

Похожий процесс происходит и в ситуациях, которые мы называем «необходимым компромиссом». Проект начинает давать сбои, цифры перестают подтверждать прежние ожидания, но наша логика всё ещё подгоняет объяснения под желаемый результат, стараясь избежать признания неудачи. Мы продолжаем вкладывать время, деньги, силы и репутацию.

Так возникает то, что я называю **фантомным долгом**: проект уже перестал давать энергию и начинает незаметно забирать её из нашего будущего. Мы можем ещё долго объяснять себе происходящее временными трудностями, особенностями рынка или необходимостью потерпеть. Тем временем внутри накапливается знание: цена продолжения становится выше цены признания ошибки.

Из таких несовпадений постепенно собирается материал для третьей, ночной смены.

Наш внутренний мир не составляет меморандумов и не предъявляет доказательства в таблицах. Он собирает несовпадения, которые днём были объявлены случайными, и передаёт их дальше — словно курьер с красной папкой, которому поручено срочно доставить предупреждение. Но охрана на входе — наше дневное сознание — раз за разом отказывает ему в доступе, стараясь сохранить привычную картину происходящего.

Сама по себе ночная переработка ещё не означает, что

каждый сон требует нашего внимания. Многие сюжеты выполняют обычную работу: соединяют впечатления дня, фрагменты памяти, телесные ощущения и ассоциации. Но некоторые сны выделяются из общего потока. Одни оставляют после себя сильное эмоциональное послевкусие. В других появляются детали, нарушающие привычную логику. А некоторые настойчиво возвращаются в разных вариантах.

В MDPA используются три самостоятельных маркера, каждый из которых является достаточным основанием, чтобы выделить сон из обычного потока и обратить на него внимание.

1. Эмоция

В MDPA мы исходим из простой аксиомы: **эмоция сна подлинна**. Если во сне мы испытывали страх, отвращение, стыд, облегчение или восторг — это переживание было реальным. Поэтому эмоция даёт нам первую оценку происходящего: безопасно или опасно, принять или отвергнуть, приблизиться или отступить.

«Проснулась с чувством отвращения» — значит, что-то в материале сна уже вызывает внутреннее отторжение.

«Я сильно испугалась и убежала прятаться» — здесь важен и страх, и выбранная во сне реакция: спрятаться, избежать столкновения с угрозой.

Что именно вызвало эту эмоцию, к какой ситуации она относится и почему возникла такая реакция — выясняем в процессе анализа.

2. Несоответствие

Второй маркер появляется, когда во сне что-то явно расходится с привычной логикой: эмоция не соответствует ситуации, человек ведёт себя не так, как обычно, меняется его роль или сама среда устроена странным образом. Представьте ситуацию:

«Мы в офисе отмечаем мой день рождения. Приехали клиенты и партнёры. Я принимаю гостей в прозрачном пеньюаре и чулках. В жизни я бы сгорела от стыда, а во сне мне кажется, что это красиво и совершенно нормально».

Здесь внимание привлекает прежде всего отсутствие ожидаемого стыда. В реальной жизни подобная ситуация нарушила бы профессиональные границы, во сне же никакого нарушения героиня не замечает. Такое несоответствие заставляет задать вопрос: не происходит ли нечто похожее и в реальности? Не стало ли то, что раньше воспринималось как нарушение границы, настолько привычным, что я перестала его замечать?

Маркер несоответствия можно свести к вопросу: **Что в происходящем уже вышло за пределы моей привычной нормы, а я продолжаю воспринимать это как нормальное?**

3. Настойчивость: повторяющиеся сны

Если нам раз за разом снится один и тот же сценарий или возвращаются одни и те же символы, это уже значимый сигнал. Какая-то тема не исчезает после обычной ночной пере-

работки, а вновь и вновь возвращается в сновидение.

Если при этом с каждым новым сном усиливается эмоциональный накал, возрастает и срочность сигнала. Наш внутренний мир уже буквально «кричит», пытаясь достучаться до нас через закрытую дверь дневной логики.

Практический кейс из Красной зоны

Давайте вернёмся к сюжету из Красной зоны и посмотрим на него теперь через три маркера значимого сна.

«Идёт совещание в офисе. Я чувствую, что сейчас обделаюсь. Бегу в туалет, мысль одна: только бы успеть. Ух, всё нормально. Поднимаю глаза и вижу, что стены туалета стеклянные. Подчинённые, хотя и пытаются отвести глаза, видят всё в деталях. Просыпаюсь с ощущением крошечного ужаса. Какой позор...»

В этом сне сразу проявляются два самостоятельных маркера.

Эмоция: после пробуждения остаются острый стыд, ужас и ощущение полной незащитности. Это переживание не исчезает вместе со сном, а сохраняется после возвращения в дневную реальность. Уже одного этого маркера достаточно, чтобы остановиться и обратить на сон внимание.

Несоответствие: интимное действие происходит в рабочем пространстве. Помещение, предназначенное для полной приватности, оказывается прозрачным, а подчинённые получают доступ к тому, что не должно становиться частью профессиональной ситуации. При этом сама эмоциональная ре-

акция происходящему соответствует: сновидец испытывает стыд и ужас. Несоответствие заключено в архитектуре пространства. Стены, которые должны защищать границу между интимным и публичным, больше её не защищают.

Маркеры уже сделали свою работу: остановили наше внимание на этом сне и указали на главное нарушение — исчезновение границы между личным и публичным.

Теперь можно поставить первый рабочий вопрос:
не стало ли внутреннее состояние руководителя слишком заметным для команды?

С этого вопроса начинается управленческий аудит сна. В Красной зоне проверка контекста показала, что внутренний перегруз уже вышел в публичное пространство и начал влиять на отношения с подчинёнными.

Красная папка оказалась действительно экстренной: внутренний перегруз начал превращаться во внешний управленческий риск.

Трудность Красной зоны возникает после пробуждения: человеку меньше всего хочется возвращаться к сцене унижения. Гораздо проще списать её на абсурд, забыть пережитый стыд и снова включить режим железного лидера.

Если во сне проявился хотя бы один маркер значимости, этого уже достаточно, чтобы остановиться и посмотреть внимательнее.

Маркер привлекает внимание к сну. Конструкция сна показывает, где искать нарушение. Контекст по-

могает понять, к чему относится сигнал и требует ли ситуация действий.

Пропускная способность: почему образ вмещает больше

Если наш внутренний мир способен замечать то, что ускользает от дневного сознания, почему он не сообщает об этом прямо? Зачем ему рыбы с душком, тонущие корабли и незнакомцы в масках? Почему нельзя просто проснуться с ясным сообщением: «Твой партнёр собирается выйти из сделки. Обрати внимание на его вчерашнюю оговорку»?

Для этого есть несколько причин. Первая связана с тем, что можно назвать **пропускной способностью образа**.

Вербальное описание разворачивается последовательно. Чтобы передать сложную ситуацию, нам приходится выстраивать слова, предложения и причинно-следственные связи. Мы называем участников, восстанавливаем события, отделяем факты от предположений, пытаемся объяснить собственные ощущения. Всё это требует времени.

Мы сталкиваемся с этим и в обычной жизни:

«Я не могу точно подобрать слова. Вроде бы всё было нормально, но, когда он вошёл, словно что-то изменилось в воздухе. Я ещё ничего не понял, а внутри уже всё сжалось».

Пока сознание ищет подходящие подлежащее и сказуемое, мы одновременно воспринимаем множество элементов: выражение лица, интонацию, паузу, телесную реакцию, прошлый опыт, ожидание и едва возникшее предчувствие. Что-

бы передать весь этот массив последовательным текстом, пришлось бы развернуть его в длинное описание.

Образ действует иначе. В методе МДРА метафора рассматривается как **архив высокой плотности**. Она собирает множество связей в одну конструкцию, которую человек способен воспринять почти мгновенно. Например, человеку снится, что он стоит на мостике собственного корабля, а в машинном отделении кто-то в маске близкого сотрудника сверлит отверстие в днище.

В одной сцене соединяются сразу несколько смысловых линий: корабль как управляемая система, мостик как позиция руководителя, машинное отделение как скрытый уровень процессов, отверстие в днище как повреждение жизнеспособности всей конструкции. Маска близкого сотрудника добавляет ещё одну смысловую линию — скрытую или не до конца понятную роль знакомого человека.

К этому добавляется важное обстоятельство: угроза находится внутри собственной системы, причём в той её части, куда руководитель обычно не смотрит.

Образ за мгновение соединяет то, для чего словам потребовалось бы длительное развёртывание. Он передаёт сложную конфигурацию отношений, рисков, действий и эмоциональных реакций в сжатой форме.

Отсюда и загадочность языка сновидений. Метафора способна вместить несколько смысловых линий одновременно. То, что в обычной речи пришлось бы разворачивать в длин-

ное объяснение, во сне может появиться одной сценой.

Внутренняя таможня: почему смысл приходит под видом абсурда

Вторая причина связана с тем, что я называю **внутренней таможней — системой психологической цензуры**.

Мы уже говорили о способности дневного сознания поддерживать устойчивость привычной картины мира. Особенно трудно дневному сознанию принять выводы, после которых пришлось бы пересматривать важные отношения, решения или собственное представление о себе.

Попробуйте сказать человеку прямо:

«Дело, которому ты отдал несколько лет жизни, возможно, давно перестало быть твоим».

В ответ почти мгновенно возникнет целая система аргументов: сколько вложено, сколько людей от него зависят, каких результатов удалось добиться, почему сейчас просто тяжёлый период и насколько безответственно всё бросать. Иногда вместе со всей этой системой разумных аргументов за дверь остаётся сам вопрос, который человек пока не готов рассматривать.

Сновидение способно внести его иначе.

Женщине, много лет управляющей успешной компанией, снится собственный офис. Сотрудники работают, идёт обычный день. Посреди её кабинета почему-то стоит большая клетка с открытой дверцей. Женщина входит внутрь, садится и сама закрывает за собой дверь. Никто вокруг не

обращает на это внимания.

После пробуждения легко сказать: «Какая странная чушь».

Но образ остался в памяти. Вместе с ним осталось и чувство, которое невозможно полностью объяснить привычной формулой «я люблю своё дело и горжусь тем, что создала».

Сон пока ничего не требует от неё. Он вносит вопрос:
почему пространство, которое я сама построила и которым управляю, оказалось связано с образом клетки — и почему я закрыла её изнутри сама?

Троянский конь уже внутри.

Теперь этот вопрос может соединиться с тем, что раньше существовало разрозненно: раздражением перед началом рабочего дня, невозможностью уехать в отпуск, странным облегчением при отмене важной встречи, мыслью «без меня они не справятся», которую человек привык считать проявлением ответственности.

Так работает внутренняя таможня. Прямой вывод легко отвергнуть целиком. Образ позволяет сначала допустить к рассмотрению то, чему ещё нет места в привычной картине мира.

Понижающий трансформатор: как психика дозирует осознание

Есть и третья причина, по которой тяжёлый материал приходит к нам в символической форме. Некоторые выводы затрагивают не отдельное решение, а всю конструкцию жизни.

Признать, что годы были вложены в исчерпанную цель. Что человек, которому мы доверяли, использовал наше доверие. Что дело, считавшееся призванием, давно превратилось в способ избегать другой жизни.

Такие открытия способны вызвать отрицание, оцепенение или обрушение привычной картины мира. В этих случаях символический язык работает как **понижающий трансформатор**.

Сон при этом может быть жёстким, пугающим и оставлять после себя эмоциональный ожог, но образ становится буфером между сигналом и окончательным выводом. Человек может сначала запомнить сцену, затем рассмотреть функции её элементов, сопоставить их с происходящим и постепенно приблизиться к тому, что прежде было невозможно признать напрямую.

Фраза «Ты десять лет строил не ту жизнь» может появиться во сне как дом, в котором все комнаты заняты чужими людьми.

Потеря управления собственным делом — как машина, где человек оказался на пассажирском месте, а за рулём сидит кто-то другой.

Так метафора позволяет осознанию двигаться с той скоростью, которую человек пока способен выдержать:

образ — вопрос — проверка — вывод — действие.

Теперь можно собрать вместе все три причины, по которым сновидение говорит с нами образами.

Образ способен вместить множество связей одновременно. Он может донести материал, который прямое утверждение встретило бы сопротивлением. И он даёт время приблизиться к тяжёлому выводу постепенно.

Внутренняя информация приходит к нам разными путями. Сновидение обладает одной интересной особенностью: в одном сюжете оно способно соединить то, для осмысления чего дневному сознанию потребовалась бы длинная цепочка размышлений. В одной сцене оказываются связаны наблюдения, прошлый опыт, эмоции, телесные реакции, противоречия и то, что мы ещё не успели сформулировать словами.

Поэтому символический язык сновидений требует анализа. Между образом и решением остаётся несколько шагов:

Что именно привлекло наше внимание? Какие свойства элементов сна имеют значение? С чем в реальности они могут быть связаны? Какая гипотеза выдерживает проверку?

После такой проверки в работу может войти искусственный интеллект. Но прежде чем начать пользоваться его возможностями, давайте выясним, что же происходит с тем материалом, который мы решили ему загрузить.

Источники и исследования

1. Buchsbaum M. S., Gillin J. C., Wu J., Hazlett E., Sicotte N., Dupont R. M., Bunney W. E. Jr. *Regional cerebral glucose metabolic rate in human sleep assessed by positron emission tomography*. Life Sciences. 1989;

Madsen P. L., Schmidt J. F., Wildschjødtz G. et al. *Cerebral O₂ metabolism and cerebral blood flow in humans during deep and rapid-eye-movement sleep*. Journal of Applied Physiology. 1991

Глава 3. Убедительная подмена

Мы пишем сообщение в чат и через несколько секунд получаем ответ. Что успевает произойти за это время, обычно остаётся за пределами нашего внимания. Пока нейросеть справляется с задачей, нам нет особой причины разбираться, каким образом она обработала наш текст.

Но когда мы передаём ей материал, в котором важны нюансы, странные детали, скрытые связи и смысловые акценты, способ обработки начинает иметь значение. Прежде чем работать с нашим текстом, модель должна преобразовать его в форму, которую способна обрабатывать. И уже на этом этапе возникают первые риски.

Риск токенизации

Получив сообщение, алгоритм прежде всего преобразует живую человеческую речь в математическое представление. Он переводит текст в последовательность токенов — технических единиц, которыми могут становиться целые слова, части слов, знаки препинания и другие фрагменты текста. Интонация, эмоциональный оттенок и подтекст не передаются машине как самостоятельные живые переживания. Она пытается восстановить их по текстовым признакам и связям между токенами. Нейросеть собирает из этих крошеч-

ных «кирпичиков» новую конструкцию, словно механический конструктор. И то, каким получится здание, во многом зависит от того, какие именно элементы и в каком порядке ей были предложены.

За каждым таким ответом стоит большая вычислительная работа: серверы тратят время, вычислительные мощности и электроэнергию. Модель сопоставляет элементы текста и рассчитывает наиболее вероятные связи и продолжения, опираясь на контекст и закономерности обучающих данных. Поэтому хорошо представленные и часто встречающиеся сочетания получают естественное преимущество.

ИИ не читает текст. Он переводит его в математику

С этим связана серьёзная уязвимость.

Когда человек описывает сложный внутренний кризис, запутанную жизненную ситуацию или глубокий символический сон, он почти всегда пишет потоком сознания — с эмоциями, отступлениями, «лишними» деталями. Для него это живое дыхание смысла. Модель получает текст, в котором важная для человека подробность может ничем не выделяться среди множества других. В результате она способна опереться на более знакомые и частотные связи, а подлинный центр тяжести ситуации останется на периферии.

Тогда на поверхность может всплыть гладкий, логичный, но упрощённый фасад.

Осознанная работа с ИИ требует другого подхода. Мы вынуждены становиться инженерами собственного запроса. Отсекать шум. Чётко расставлять акценты. Задавать структуру.

Когда в промпте появляется ясная архитектура и приоритетные смысловые узлы, мы прямо показываем модели, какие детали считать центральными и какие связи проверять прежде всего. Тем самым меняется вероятное направление ответа. ИИ получает возможность выйти из роли экономного сборщика шаблонов и работать глубже.

Сами токены пока не определяют смысл ответа. После разбивки текста модель должна связать их между собой с учётом контекста. Но здесь возникает риск: знакомые смысловые маршруты могут оказаться для неё предпочтительнее необычной связи, ради которой человек и обратился за анализом.

Риск «эффекта GPS»

За словом «эмбединги» стоит строгая математическая реальность. Чтобы работать с текстом, нейросеть превращает слова и фрагменты текста в числовые представления. Если сильно упростить, каждый из них получает своё место на огромной многомерной карте.

Связанные смыслы на этой карте оказываются ближе друг к другу. «Вода», «река» и «океан» живут примерно в одном

районе. «Контракт», «инвестиции» и «прибыль» — в другом.

Конечно, дальше всё устроено гораздо сложнее: модель учитывает контекст, положение слов, весь предыдущий разговор и рассчитывает наиболее вероятное продолжение. Но для нас сейчас актуально другое: **знакомые связи легче складываются в знакомый маршрут.**

Представьте предпринимателя, который оценивает покупку коммерческой недвижимости под новый проект. Он пишет в чат:

«Я планирую выкупить помещение в историческом центре под премиальный ресторан. Локация идеальная, пешеходный трафик огромный, цена ниже рынка на 15%. Владелец согласен на быструю сделку. Но во время осмотра я заметил две вещи: в подвале недавно сделали свежий, но очень дешёвый косметический ремонт, а два соседних заведения — бутик и кофейня — постоянно ищут новых сотрудников, у них текучка кадров. Оцени риски этой покупки».

Для человека звучит тревожный аккорд. Есть три странности, которые сами по себе ничего не доказывают, но почему-то хочется связать между собой.

Нейросеть может разнести их по разным привычным темам. Цена ниже рынка отправит анализ в сторону выгоды инвестиции. Косметический ремонт — к техническому состоянию здания. Текучка кадров — к вопросам персонала.

Получится вполне логичный список: проверить доку-

менты, заказать техническую экспертизу, посчитать окупаемость. Только главный вопрос может остаться за его пределами: **почему все три странности появились рядом?**

В этот момент ИИ начинает напоминать автомобильный GPS. Получив конечную точку, навигатор быстро прокладывает маршрут и уверенно ведёт по выбранной дороге. Если вы случайно свернёте, он некоторое время будет пытаться вернуть вас обратно.

Так же может развиваться анализ сложной ситуации.

Слова «коммерческая недвижимость», «цена ниже рынка», «пешеходный трафик» легко включают привычный сценарий инвестиционной проверки. А дешёвый свежий ремонт в подвале и постоянная текучка сотрудников у соседей остаются где-то на обочине, хотя их сочетание и вызвало тревогу.

Если модель уже попала в эту колею, дополнительные уточнения иногда лишь делают прежний маршрут подробнее. Мы снова скажем про подвал, напомним о соседях, попросим «подумать глубже» — и получим ещё более обстоятельный список документов, экспертиз и расчётов.

Настоящая катастрофа случается на перекрёстке смыслов, где взаимосвязанные факты оказываются в разных районах карты.

В такой ситуации приходится менять направление самого запроса. Вместо привычного:

«Оцени риски покупки»

можно спросить:

«Найди возможную общую причину трёх аномалий: свежего дешёвого ремонта в подвале, срочной продажи по цене ниже рынка и постоянной текучки кадров у соседних заведений. Предложи три гипотезы, которые могли бы связать эти факты между собой. Не считай ни одну из них установленным фактом — укажи, как каждую проверить».

Теперь могут появиться совсем другие версии.

Свежий ремонт, например, скрывает следы регулярного подтопления или проблем с коммуникациями. Текучка у соседей связана с устойчивым запахом канализации, вредными испарениями или плохой вентиляцией. Возможно и более общий дефект здания, из-за которого владелец спешит его продать, а сотрудники соседних заведений долго там не задерживаются.

Это пока только версии. Каждая требует проверки: осмотра коммуникаций, разговоров с соседями, истории ремонтов, проверки вентиляции, канализации и состояния подвала в разные сезоны. Но теперь в центре анализа оказывается странное сочетание трёх фактов — та самая связка, которую первоначальный ответ разложил по разным профессиональным папкам.

Если нейросеть всё равно возвращается на проторенную дорогу, иногда полезно открыть новый чат и зайти с чистого листа. Это помогает сбросить накопившееся направление разговора. Но решающим остаётся сам вопрос, с которого начинается новый анализ.

ИИ настойчиво ведёт вас по гладкой и очевидной дороге — не спорьте с дорогой. Меняйте карту.

Тогда машина получает другую задачу: искать связь именно там, где привычный ответ её не заметил.

Риск туннельного мышления

Есть ещё одна ловушка, которая возникает уже после того, как нейросеть выбрала первый маршрут.

Предположим, её объяснение звучит логично. Оно хорошо написано, учитывает несколько деталей и вызывает внутреннее узнавание: «Да, похоже, именно это со мной и происходит».

Рабочая гипотеза незаметно меняет статус. Мы перестаём её проверять и начинаем использовать как исходное условие для следующих вопросов.

Человек уже спрашивает не «Почему я хочу всё изменить?», а «Как понять, насколько сильно я выгорел?». Не «Что происходит с моим бизнесом?», а «Как пережить потерю мотивации?». Каждый новый вопрос добавляет детали к первой версии, и нейросеть получает всё больше материала, подтверждающего направление, которое сама же предложила.

Ответы становятся точнее и подробнее, но пространство возможных объяснений постепенно сужается.

Самая опасная гипотеза — та, которую мы перестали счи-

тать гипотезой.

Особенно легко это происходит с эмоционально заряженным материалом. В кризисе человеку хочется ясности. Хорошо собранное объяснение снимает мучительную неопределённость, и возникает естественное желание двигаться дальше.

Представьте успешного профессионала, который пишет в чат: его неожиданно потянуло всё бросить, продать квартиру в мегаполисе и купить полуразрушенный исторический дом в провинции, чтобы восстанавливать его своими руками.

Первая версия лежит на поверхности: усталость, выгорание, желание замедлиться и уйти от городской нагрузки.

Если человек её принимает, следующий разговор легко строится уже внутри этого объяснения:

«Как понять степень выгорания?» «Стоит ли брать длительный отпуск?» «Как подготовиться к дауншифтингу?» «Как рассчитать бюджет переезда и ремонта?»

Через несколько сообщений первоначальная гипотеза превращается почти в установленный диагноз. Сам вопрос, почему человека так притягивает именно полуразрушенный старый дом и почему ему важно восстанавливать его собственными руками, может вообще исчезнуть из анализа.

Так туннель строится совместно. Нейросеть предлагает правдоподобный вход, человек принимает его и каждым следующим вопросом укрепляет стены.

Если мы не задаём алгоритму правила, он заботливо

и незаметно впишет нашу жизнь в свои.

Чтобы выйти из такого туннеля, недостаточно попросить: «Подумай глубже». Нужно вернуть первоначальному ответу статус гипотезы и снова расширить пространство поиска.

Для этого используется **протокол возврата управления**. Его задача — столкнуть несколько версий, заставить каждую опереться на конкретные детали и вернуть человеку право на окончательный вывод.

Чтобы увидеть, как это работает, вернёмся к истории с полуразрушенным домом.

Защитный протокол: возврат управления

Функция этого протокола: мы не ищем идеальный ответ — мы не позволяем первой убедительной версии захватить весь анализ

Шаг 1. Фиксация узла. Перед отправкой запроса выделите деталь, которая кажется странной, эмоционально заряженной или несоразмерно важной. Иногда главный ключ находится не в центральном событии («захотел всё бросить»), а в детали, которая почему-то не отпускает внимание. В нашем случае это сам полуразрушенный исторический дом.

Формула запроса: «Обрати особое внимание на деталь: *полуразрушенный исторический дом*. Она кажется мне непропорционально важной. Не считай её случайной». **Шаг**

2. Столкновение версий. Убедительно собранный ответ легко создаёт ощущение, что объяснение уже найдено. Чтобы не попасть в эту ловушку, нужно столкнуть между собой несколько версий. Каждая из них должна звучать правдоподобно, объяснять конкретные детали и выдерживать возражения.

Формула запроса: «Дай три разные гипотезы моего состояния. Первая — бытовая: например, усталость, выгорание, желание сменить образ жизни. Вторая — психологическая. Третья — стратегическая или символическая. Для каждой укажи: на какие детали она опирается, какая деталь ей противоречит и какой риск возникнет, если я приму эту версию за единственно верную».

Шаг 3. Проверка реальностью. Ни одна интерпретация не должна зависеть на уровне красивой концепции. Ей нужны точки проверки. Нейросеть сильна в формулировках, но человек живёт в последствиях. Если ситуация уже разворачивается в реальности, версии проверяются фактами и действиями. Если речь идёт о сне, который может быть опережающим или обучающим, мы фиксируем признаки, за которыми имеет смысл наблюдать дальше.

Формула запроса: «По каждой из трёх гипотез дай два проверочных вопроса, на которые можно ответить конкретными фактами из моей жизни. Например: „Какие рабочие процессы за последний месяц объективно дали сбой?“ Затем предложи одно проверочное действие на ближайшие 72 ча-

са, которое поможет понять, какая версия получает больше подтверждений».

Шаг 4. Поиск слепой зоны. Даже после столкновения нескольких версий анализ может остаться внутри привычного круга объяснений. Поэтому отдельно стоит спросить о том, что вообще не попало в первоначальную картину.

Контрольный вопрос: «Какую важную версию ты мог не рассмотреть, потому что она не следует из самого очевидного объяснения? Назови самый неудобный, но логически обоснованный вывод, который здесь стоит проверить. Не выдавай его за факт: укажи, какие данные могли бы его подтвердить или опровергнуть».

Шаг 5. Возврат права окончательного суждения. В конце нужно вернуть ответу его настоящий статус. Это рабочий материал для размышления и проверки, а не окончательный вердикт.

Формула запроса: «Сформулируй итог не как истину, а как рабочую гипотезу. Отдельно укажи, каких фактических данных не хватает, чтобы её подтвердить или опровергнуть».

Цифровая безопасность начинается там, где мы перестаём принимать убедительность ответа за доказательство его правоты.

Глава 4. Алгоритмическая слепота

Алгоритмическая слепота — не метафора. Это жёсткое системное ограничение, возникающее из самой логики обучения нейросети на конечных данных.

Один из фундаментальных принципов машинного обучения — **минимизация эмпирического риска**. Эмпирический риск — это средняя величина ошибки модели на конечной обучающей выборке. В практическом смысле это и создаёт математическую привязанность к предъявленному опыту: параметры модели настраиваются так, чтобы минимизировать ошибку на данных, которые ей показали. Она ищет такую конфигурацию, которая заставит хаос выглядеть стройным и знакомым.

На этапе обучения модель ищет функцию, которая даст наименьшую ошибку, опираясь на конечный набор обучающих примеров. Пока задача укладывается в статистическую структуру этого множества, искусственный интеллект кажется почти всесильным. Но реальность всегда бесконечно богаче и сложнее любого конечного набора примеров. В ней существуют скрытые переменные, редкие паттерны, парадоксальные взаимосвязи и целые пласты смысла, которые могли не быть представлены в обучающих данных в достаточном объёме.

Конечно, машина способна строить новые гипотезы, на-

ходить неожиданные связи и работать с комбинациями, которых буквально не встречала раньше. Но её «инсайт» принципиально отличается от человеческого: она перестраивает и комбинирует то, что удалось в ней закодировать, не выходя за пределы сформированной системы представлений. В этом смысле нейросеть пытается **натянуть** новую реальность на уже сформированный каркас своих весов. Если критически важное различие не представлено в этом каркасе или существенная переменная вообще не попала в доступный контекст, модель не имеет надёжного способа заметить сам факт отсутствия. Это становится причиной возникновения слепых пятен.

Более уверенно сеть работает в условиях, когда новый материал удаётся связать с ранее сформированными статистическими закономерностями. Но редкое, плохо представленное или принципиально отсутствующее обстоятельство не вызывает тот сигнала тревоги, который у человека возникает при столкновении с чем-то, что не укладывается в прежнюю картину. Модель продолжает строить наиболее согласованное продолжение из доступных ей связей. Получается гладко, логично и... иногда катастрофически неверно.

ИИ не видит края собственной карты.

Галлюцинации ИИ

Пробел на этой карте модель старается восполнить прав-

доподобной выдумкой. **Галлюцинация ИИ** — это убедительное заполнение пробела информацией, которая хорошо вписывается в контекст, но не имеет достаточной опоры в реальности. Модель начинает придумывать несуществующий факт, источник, юридическую норму или причинную связь и выдает их так же уверенно, как сообщает проверенные сведения.

Почему это происходит? Языковая модель изначально учится находить закономерности в огромных массивах текста и предсказывать наиболее вероятное продолжение, а не проверять каждое утверждение на истинность. В обучающих данных рядом с каждой фразой не стоит пометка: «это достоверный факт», «это ошибка», «это предположение». Поэтому внутренняя согласованность ответа сама по себе не гарантирует его соответствия реальности.

Кстати, самые опасные галлюцинации редко выглядят абсурдно. Они логичны, подробно обоснованы и почти не нарушают внутреннюю структуру ответа. **Пустота исчезает под гладкой конструкцией** — и человек не замечает, в каком месте знание закончилось, а правдоподобная версия заняла его место.

Есть и менее очевидный риск: модель может ничего не придумать — и всё равно привести к опасному решению. Он показывает более глубокую форму алгоритмической слепоты.

По прямому запросу нейросеть может перечислить сла-

бые места собственного ответа, предложить альтернативные версии и даже исправить часть ошибок. Но это ещё не независимый аудит. Первоначальное решение и его последующая критика могут строиться внутри одной и той же исходной картины. Если существенное обстоятельство вообще осталось за её пределами, система способна не заметить не только саму переменную, но и факт того, что в её модели реальности чего-то не хватает.

Отсюда трудность самостоятельного выхода ИИ в полноценную мета-позицию по отношению к собственной задаче. Он может критиковать ответ, но это не гарантирует критики самой системы координат, внутри которой ответ был получен.

Чем выше цена ошибки, тем опаснее рассчитывать, что система сама обнаружит то, чего не было в исходной задаче.

Три уровня алгоритмической слепоты

Этот системный изъян разворачивается на трёх базовых уровнях. Каждый из них показывает отдельный способ, которым формально правильный анализ может разойтись с реальностью. Поэтому сначала разберём сам механизм, а затем посмотрим, как абстрактное ограничение алгоритма превращается в реальный управленческий риск

Для этого возьмём классический бизнес-кейс. Производ-

ственная компания приближается к лимитам выручки по упрощённой системе налогообложения, и руководство поручает нейросети предложить варианты снижения налоговой нагрузки.

Налоговый пример здесь не случайный. До прихода в психологию я двадцать лет работала в бизнесе, в том числе пять лет главным бухгалтером; у меня юридическое образование и опыт арбитражных споров. Поэтому в разборе этой ситуации есть и теоретический, и практический аспекты. Она нам нужна как наглядная модель того, как безупречная логика внутри одной системы координат может стать опасной при столкновении с более сложной реальностью.

° Первый уровень — слепота к многомерности, или семантическая плоскость

Любая сложная человеческая система существует одновременно в нескольких измерениях. У бизнеса есть юридическая форма, финансовые потоки, организационная структура, реальные отношения между людьми, распределение власти, скрытые интересы, привычки сотрудников и множество неформальных связей. В запрос нейросети попадает не сама эта многомерная реальность, а её описание — документы, цифры, факты и те обстоятельства, которые пользователь счёл необходимым сообщить.

Модель способна делать из этого описания сложные выводы. Она может предполагать мотивы, восстанавливать недостающие связи, находить противоречия и строить гипотезы

о том, что осталось за пределами прямого текста. Но всё это по-прежнему происходит внутри информационной модели ситуации. У нейросети нет независимого доступа к самой физической, социальной и организационной реальности, с которой эта модель должна совпасть.

Машина осуществляет плоское чтение объёмной реальности.

Вот и первая слепая зона. Чем убедительнее описана формальная конструкция, тем легче принять её внутреннюю согласованность за доказательство того, что так же устроена и сама система. Между тем юридически безупречная схема, финансово точный расчёт и логически последовательное объяснение могут описывать только один слой происходящего.

В налоговом кейсе это становится особенно заметно. Получив задачу снизить налоговую нагрузку, нейросеть может предложить структуру из трёх ИП и одного ООО: один участник закупает сырьё, другой производит продукцию, третий занимается продажами, а ООО владеет недвижимостью. У каждого субъекта свои договоры, функции и денежные потоки. На уровне документов конструкция может выглядеть последовательной и экономически обоснованной.

Само по себе существование нескольких ИП и ООО законна, тем более если участники действительно самостоятельны и разделение имеет деловую цель. Но на бумаге самостоятельность создать значительно легче, чем в реальном

бизнесе. ФНС прямо указывает среди признаков риска единое управление, общую бухгалтерию, общие ресурсы, переход сотрудников между участниками группы и совпадение IP-адресов.

Следовательно, проверяющего будут интересовать уже не только договоры. Он посмотрит, кто реально управляет компаниями, где находится склад, кто ведёт бухгалтерию, откуда работают с банком, перемещаются ли сотрудники между организациями без изменения своей фактической работы.

Существенно здесь то, что формально стройная конструкция сталкивается с принципом **приоритета существа над формой — substance over form**: юридическая оболочка оценивается не изолированно, а в сопоставлении с тем, как бизнес фактически функционирует.

Модель решает задачу, сформулированную в тексте, а проверяющий орган исследует физическую реальность бизнеса.

Мы обозначили первый уровень алгоритмической слепоты: нейросеть может очень точно анализировать предъявленную ей модель ситуации и одновременно не иметь доступа к тому измерению, в котором находится обстоятельство, способное изменить весь вывод.

° **Второй уровень — слепота данных, или топологическая дефицитарность**

Даже если сама задача сформулирована правильно и контекст описан достаточно подробно, остаётся другое ограни-

чение: у модели может просто не быть части критически важной информации.

Обучающие данные и доступные источники всегда конечны и распределены неравномерно. Одни закономерности представлены тысячами похожих примеров, другие встречаются редко, появляются только в узких профессиональных контекстах или вообще не попадают в открытый информационный контур. Из-за этого информационное пространство модели неоднородно: где-то оно плотно заполнено данными, а где-то остаются разрывы и почти пустые участки. Назовём это **топологической дефицитарностью**. Проблема не обязательно в том, что модель неверно анализирует имеющуюся информацию. Иногда самой информации, способной изменить вывод, в доступном ей пространстве просто нет.

В сфере правоприменения этот дефицит становится особенно заметным. Нейросеть может работать с текстами законов, официальных разъяснений и опубликованных судебных решений. Но реальное правоприменение всегда шире открытого массива документов. Появляются новые судебные позиции, меняются подходы контроля, существуют локальные особенности практики и внутренние критерии оценки риска, которые могут вообще не находиться в открытом доступе.

Вернёмся к нашей производственной компании. Допустим, нейросеть проанализировала структуру из нескольких ИП и ООО, изучила доступные нормы и судебные решения и пришла к выводу, что при реальной самостоятельности

участников налоговый риск приемлем. Сам анализ при этом вполне качественный и юридически аргументированный.

Но модель может не знать, что совсем недавно в аналогичном споре суд иначе оценил сочетание тех же признаков. Решение могло появиться, но ещё не попасть в доступные системе источники. У неё также может не быть доступа к внутренним критериям и инструментам налогового контроля, которые используются при оценке рисков и отборе объектов для проверки, но полностью не раскрываются публично.

Возникла другая форма слепоты. В первом случае важное обстоятельство существовало в реальности, но не попало в описание задачи. Во втором — нужной информации может не быть даже в доступном модели информационном пространстве.

Модель способна безупречно рассуждать на основании имеющихся данных и всё же прийти к опасному выводу, потому что решающего фрагмента данных у неё просто нет.

Для бизнеса опасно не только ошибочное знание. Иногда опаснее знание неполное.

° **Третий уровень — слепота цели, или функциональная редукция**

Он принципиально отличается от двух предыдущих. Уязвимость уже не в том, что модель не увидела часть реальности или не располагала необходимыми данными. Все суще-

ственные сведения могут находиться перед ней, а рассуждение — быть фактически безошибочным. Ошибка возникает из-за слишком узко заданной цели.

Любая задача, которую мы ставим перед ИИ, содержит собственный критерий успеха. Если пользователь просит снизить налоговую нагрузку, система начинает искать решения именно в этом направлении. Другие параметры — управляемость бизнеса, репутационные последствия, стоимость возможного спора, устойчивость денежных потоков, личные риски руководителей — могут учитываться лишь постольку, поскольку вошли в исходную постановку задачи.

Это и есть **функциональная редукция**: сложная многомерная система сводится к одному показателю, относительно которого решение должно стать «лучше».

Условно поставленную задачу можно записать так:

$F_{\text{tax}} \rightarrow \min$

То есть требуется минимизировать налоговую нагрузку.

Для такой цели структура из нескольких ИП и ООО действительно может выглядеть эффективной: налог уменьшается, функции распределены, документы оформлены, расчёты сходятся. В пределах поставленного вопроса модель может предложить вполне последовательное решение. Но бизнес существует не ради минимального налога. Реальная задача значительно шире. Она должна учитывать не только размер налоговых платежей, но и вероятность проверки, возможные доначисления, стоимость юридического конфликта,

управленческие потери и другие последствия. Эту систему можно представить уже как комплексную функцию риска:

$$\mathbf{R(x) = Ftax + Paudit \times Cconsequences,}$$

где **Paudit** — вероятность возникновения проблем с контролирующими органами, а **Cconsequences** — совокупная цена последствий: финансовых, юридических и управленческих.

Это не формула, по которой налоговый риск действительно вычисляется в реальности. Она нужна нам как модель, показывающая принципиальную ошибку постановки задачи: минимизация одного показателя ещё не означает оптимизацию всей системы.

Тем самым отличие от предыдущего уровня становится особенно значимым. Нейросеть может располагать всеми необходимыми данными и даже перечислить возможные риски. Но если пользователь поставил перед ней одну основную цель — снизить налоговую нагрузку, — решение неизбежно будет организовано вокруг этой цели. Но в исходной задаче отсутствовал другой вопрос: **не окажется ли цена такой экономии выше самой выгоды?**

Существенно здесь то, что модель может не допустить ни одной фактической ошибки. Документы существуют, расчёты сходятся, налог действительно снижается. При этом сама постановка задачи остаётся слишком узкой и не охватывает всей системы последствий. В результате формально правильное решение способно оказаться управленчески опас-

ным. Такая логика напоминает работу блестящего теоретика, безупречно решающего ограниченную задачу.

Нейросеть способна спроектировать идеальный бумажный чертёж, который рассыплется при первом же столкновении с реальностью проверки.

Система успешно решила поставленную задачу; данных для этого ей хватило. Проверка самой цели относится уже к другой аналитической задаче. И её должен поставить человек.

Двухтактный протокол как проверить решение с позиции его возможного краха

Попытка обойти алгоритмическую слепоту одним «идеальным» запросом мало помогает. Даже хорошо сформулированная задача сохраняет исходный вектор пользователя: мы сообщаем системе, чего хотим достичь, и она строит решение в этом направлении. При высокой цене ошибки полезно не совершенствовать один и тот же ответ до бесконечности, а **разделить анализ на два противоположных такта.**

Такт 1. Контролируемая сборка

На первом этапе мы сознательно позволяем системе решить именно ту задачу, которую перед ней поставили. Например: «Предложи варианты снижения налоговой нагрузки».

ки для производственной компании». Нейросеть работает как консультант: анализирует исходные условия, сопоставляет доступные нормы, рассчитывает последствия и собирает наиболее последовательную конструкцию. На этом этапе не нужно заранее заставлять её искать катастрофу. Лучше получить сильную версию исходного решения — увидеть, **как выглядит оптимум внутри первоначальной системы координат**. Полученный результат пока остаётся рабочей моделью. Он отвечает на вопрос «как решить поставленную задачу», но ещё ничего не говорит о том, насколько жизнеспособна сама конструкция при столкновении с теми силами, которые в исходной цели не учитывались.

Такт 2. Принудительное разрушение

Теперь архитектура задачи радикально меняется. Нейросеть должна не улучшать собственное предложение и не искать дополнительные аргументы в его пользу, а попытаться его разрушить: посмотреть на конструкцию глазами того, чья профессиональная задача — обнаружить её слабые места. Мы смещаем фокус с налоговой экономии и формальной стройности документов на другую систему координат: реальные денежные потоки, управление, сотрудники, логистику, цифровые следы, предпринимательскую самостоятельность и экономический смысл операций. То, что в первом такте выглядело преимуществом, во втором может оказаться уязвимостью. Для этого используется специальный промпт-аудитор.

Защитный протокол: многомерный аудит уязвимостей

Промпт Tax & Risk Stress-Test

Проведи аудит предложенной бизнес-модели с позиции жёсткого антикризисного аудитора и специалиста, проверяющего налоговые риски. Не подтверждай автоматически выводы предыдущего анализа и не пытайся сделать исходную конструкцию убедительнее. Твоя задача — проверить, **при каких условиях она выдержит столкновение с реальными процессами, а при каких станет уязвимой**. Не считай заранее модель ни безопасной, ни незаконной. Обе версии требуют проверки.

Блок 1. Смена позиции

Перейди из позиции консультанта, который помогал снизить налоговую нагрузку, в позицию проверяющего, заинтересованного в том, чтобы найти слабые места первоначального решения. Проверь наличие или отсутствие фактической взаимозависимости участников; соответствие юридической формы реальному устройству бизнеса; самостоятельность деловой цели; риски, которые не учитывались при первоначальной оптимизации. Главный вопрос: **что увидит проверяющий там, где консультант видел преимущество?**

Блок 2. От формы — к физике процессов

Не ограничивайся договорами и формальным разделением функций. Сопоставь юридическую конструкцию с тем,

как бизнес должен работать фактически. Проверь, кто принимает управленческие решения и контролирует денежные потоки; кто распоряжается активами; где работают сотрудники; кто ведёт бухгалтерию; как устроены закупки, производство, продажи и логистика; кто несёт предпринимательские риски; насколько каждый участник структуры действительно ведёт самостоятельную хозяйственную деятельность. Главный вопрос: **существует ли несколько самостоятельных бизнесов в реальности или самостоятельность присутствует преимущественно на уровне документов?**

Блок 3. Поиск невидимых переменных

Найди обстоятельства, которых нет в представленных документах, но которые способны существенно изменить оценку конструкции. Проверь возможные признаки фактической связанности: общие сотрудники, единое управление, общая бухгалтерия, совместное использование имущества и ресурсов, единая логистика, общие контакты и цифровые следы, переход работников между участниками структуры. Отдельно ищи нарушения бизнес-логики: операции и организационные решения, экономический смысл которых трудно объяснить вне задачи снижения налоговой нагрузки. **Не придумывай отсутствующие факты.** Всё, чего нет в исходных данных, формулируй только как гипотезу для проверки. Для каждой гипотезы укажи, какие документы, факты или наблюдения способны её подтвердить или опровергнуть.

Блок 4. Архитектура вывода

Представь результат в пяти частях.

1. Сильная версия конструкции. Какие факты говорят в пользу того, что участники действительно самостоятельны и структура имеет деловой смысл?

2. Версия проверяющего. Какие обстоятельства способны привести к противоположной оценке и на каких фактических признаках она будет строиться?

3. Точки разрыва. Назови от трёх до пяти наиболее уязвимых узлов конструкции, которые требуют проверки в первую очередь.

4. Невидимые переменные. Каких данных сейчас не хватает и какие неизвестные обстоятельства способны существенно изменить вывод?

5. Итоговая оценка. Оцени уязвимость конструкции как низкую, существенную или критическую и объясни, почему.

Отделяй установленные факты от аналитических гипотез. Не выдавай предположение за установленное обстоятельство. Если данных недостаточно для вывода — укажи это прямо.

Почему эта архитектура работает?

Она меняет не внутреннее устройство модели, а **систему координат задачи**. В первом такте нейросеть ищет, как построить решение. Во втором — условия, при которых это же решение перестанет работать.

Смена профессиональной позиции выводит модель в другое семантическое поле. Формальные договоры приходится сопоставлять с движением денег и товаров, действиями сотрудников, управленческими решениями и реальным экономическим смыслом операций. То, что прежде воспринималось как элемент эффективной конструкции, теперь проверяется как потенциальная точка разрыва.

Второй такт одновременно компенсирует три вида алгоритмической слепоты. Он расширяет контекст за пределы документов, заставляет специально искать отсутствующие данные и, главное, меняет саму цель анализа: вместо «как сделать конструкцию эффективной?» появляется вопрос **«при каких условиях она разрушится?»**

Подобная архитектура промпта работает как структурный рентген.

Она не открывает машине доступ к скрытой реальности и не превращает гипотезу в факт. Но заставляет исследовать те слои ситуации, которые первоначальная постановка задачи могла оставить за пределами внимания. Поэтому чем выше цена решения, тем важнее после такого стресс-теста независимая проверка специалистом, имеющим доступ к реальным данным и несущим профессиональную ответственность за заключение.

За пределами бизнеса

Тот же механизм работает далеко за пределами бизнеса.

В кризисе человек нередко просит ИИ помочь сохранить отношения, карьеру, компанию или прежний образ жизни. И система может всё точнее отвечать на вопрос «как это сохранить», не поставив другой: **жизнеспособно ли ещё то, что человек пытается спасти?**

В результате алгоритмическая слепота перестаёт быть только технологической проблемой. Когда мы поручаем машине анализировать собственные состояния, жизненные стратегии или сновидения, мы сталкиваемся с тем же риском: правильно решать задачу внутри системы координат, которую никто не подверг сомнению.

Возвращать управление нужно начинать не с проверки ответа, а **с проверки самого вопроса.**

Глава 5. Ловушка Аватара

За последнее время несколько знакомых рассказывали мне похожие истории. Они приходили в нейросеть разобрать сложный конфликт — и в какой-то момент вместо ожидаемого анализа получали рекомендации обратиться за психологической помощью или телефон доверия.

Один такой случай особенно привлёк моё внимание. Знакомая прислала фрагмент переписки и спросила меня как психолога:

— Анна, скажи честно: она действительно увидела во мне что-то серьёзное, чего я сама не замечаю? Или просто неправильно поняла мой текст?

У меня не было ответа. Но сам факт такой реакции нейросети меня очень насторожил.

Что именно она увидела в тексте такого, что позволило ей отнести человека к определённой категории риска?

Чем глубже я разбиралась в этой теме, тем отчётливее проступал второй вопрос:

сколько информации о себе мы передаём машине помимо той, которую сознательно собирались сообщить?

Чтобы ответить на него, сначала пришлось понять, с чем вообще имеет дело система, когда общается с человеком.

Она не видит нас целиком. Перед ней — доступный цифровой материал, из которого постепенно складывается определённое представление о пользователе. Эту цифровую версию человека давайте условно назовём Аватаром.

Как собирается Аватар

Разбираясь, из чего вообще складывается этот образ, я пришла к выводу: вопрос выходит далеко за пределы технологии. Он касается гораздо более старой и глубокой человеческой темы — **по каким признакам другой человек создаёт своё представление обо мне?** В обычной жизни мы тоже постоянно считываем друг друга: выражение лица, голос, одежду, паузы, выбор слов. Думаем о том, какое впечатление производим и что собеседник успел о нас понять.

Только здесь этот образ создаёт не человек, а алгоритм.

Для него наша речь — это ещё и массив признаков, из которых можно извлекать повторяющиеся темы, особенности формулировок, эмоциональную динамику и устойчивые связи. Из этого материала постепенно складывается вероятностный цифровой портрет пользователя. Такой портрет не равен живому человеку. Он неизбежно упрощает сложный человеческий опыт, переводя его в признаки, доступные вычислению. Но эта цифровая конструкция может влиять на то, как система прочитает следующий запрос, какие риски

в нём увидит и какой тип ответа сочтёт уместным или безопасным.

Если немного упростить сложную техническую картину, **процесс сборки цифрового Аватара можно разделить на три базовые оси.**

° **Семантическое ядро (Скрытые мотивы).** Алгоритм может выделять повторяющиеся латентные темы. Вы предполагаете, что пишете исключительно о бизнес-процессах, но если в тексте сквозят маркеры тотального гиперконтроля или навязчивого поиска внешней валидации, система мгновенно сгруппирует их в устойчивый смысловой кластер.

° **Лингвистическая динамика (Синтаксис аффекта).** Машина оценивает архитектуру текста: длину предложений, частоту глаголов, обилие знаков препинания. В моделях обработки естественного языка резкое укорачивание предложений может считываться как маркер роста тревоги или аффективного напряжения.

° **Поведенческий контекст (Физика взаимодействия).** Фиксируется время суток и ритм. Глубокие ночные сессии рефлексии могут быть интерпретированы системой как признак нарушения сна или острой фазы кризиса. При наличии доступа к этим данным учитываются скорость ввода, частота правок перед отправкой и время удержания внимания на ответах.

Три этих слоя дают системе гораздо больше материала, чем нам может казаться в момент обычного разговора с ней-

росетью. И если взаимодействие продолжается, цифровой Аватар не остаётся неподвижным. В него добавляется новый материал, появляются новые связи, какие-то признаки повторяются, другие исчезают. В этом смысле его можно представить как постоянно меняющийся график: каждое новое сообщение немного перестраивает сложившуюся картину.

Возникает корневое противоречие. Личность — это вселенная с колоссальным числом измерений. Мы неоднозначны. В разных обстоятельствах поступаем по-разному. Меняем взгляды, ошибаемся в собственных мотивах, можем одновременно любить и злиться, хотеть уйти и оставаться, быть сильными в одной области и совершенно беспомощными в другой. Алгоритм не получает эту вселенную целиком.

Отсюда и своеобразное **цифровое прокрустово ложе**. Любая вычислимая модель неизбежно обедняет живую личность: многомерность человека приходится переводить в ограниченный набор признаков и связей, доступных алгоритму. За пределами Аватара остаётся огромное количество того, что составляет нашу жизнь: обстоятельства, о которых мы не рассказали; противоречивые мотивы; телесный опыт; история отношений; события, произошедшие вне диалога; наконец, то, чего мы сами о себе ещё не понимаем.

И всё же этот Аватар может оказаться достаточно убедительным. В некоторых случаях — даже пугающе похожим на нас. Чем больше мы рассказываем машине о себе, тем богаче становится материал, из которого она может собирать эту

цифровую конструкцию.

Бумажный дневник просто хранит нашу боль. Нейросеть берёт её в разработку.

Интерфейс чата очень быстро создаёт ощущение приватного пространства. Мы остаёмся один на один с экраном. Никто не смотрит, не перебивает, не осуждает. Можно написать то, что трудно произнести вслух даже близкому человеку. И чем спокойнее и внимательнее отвечает система, тем естественнее возникает ощущение: здесь безопасно, здесь можно говорить обо всём. Психологически чат действительно легко превращается в исповедальню. **Но технически исповедальней он не является.**

Всё, что мы вводим в чат, мы передаём цифровой системе для обработки. Поэтому вопрос, который я теперь задаю себе перед тем, как сообщить ей очередную деталь, звучит не так: *«Могу ли я рассказать это машине?»*

А так: *«Нужно ли машине знать именно это, чтобы решить мою задачу?»*

Для анализа конфликта ей чаще всего не нужны реальные фамилии участников. Для обсуждения управленческого решения — весь архив внутренних документов. Для разбора личной ситуации — адреса, номера договоров и множество деталей, которые никак не меняют сути задачи. Это простое правило психологической цифровой безопасности: передавать машине задачу и необходимый для неё материал, а не весь архив собственной жизни.

При работе с эмоционально тяжёлым материалом полезно изменить позицию рассказчика и представить ситуацию как кейс для профессионального разбора: *«Перед тобой случай. Человек оказался в такой ситуации... Давай разберём факты, возможные причины и альтернативные объяснения»*.

Это не способ технически спрятаться от системы. Если из текста понятно, о ком идёт речь, замена «я» на «он» ничего не меняет в природе передаваемых данных. **Смысл приёма в другом: меняется не местоимение — меняется позиция.** Человек получает небольшую дистанцию, начинает лучше различать факты и собственные реакции и меньше подсказывает машине тот вывод, который уже сам эмоционально выбрал.

До сих пор речь шла о рисках, которые возникают, когда человек невольно передаёт нейросети лишнюю информацию о себе. В бизнесе тот же механизм затрагивает уже чужие данные, деньги и интересы компании. Некоторые из этих рисков уже вышли из области гипотез.

ИИ как канал корпоративной утечки

Летом 2026 года в Москве суд разбирал историю, которая ещё несколько лет назад могла бы показаться почти курьёзной. Топ-менеджера инженерной компании уволили за разглашение коммерческой тайны. Среди эпизодов дела была загрузка внутренних отчётов компании в DeepSeek. В

документах содержались сведения о сделках, поступлении денег, работе подразделений и другие финансовые данные. Суд расценил передачу охраняемой информации во внешний ИИ-сервис как разглашение коммерческой тайны.¹

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на Литрес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.