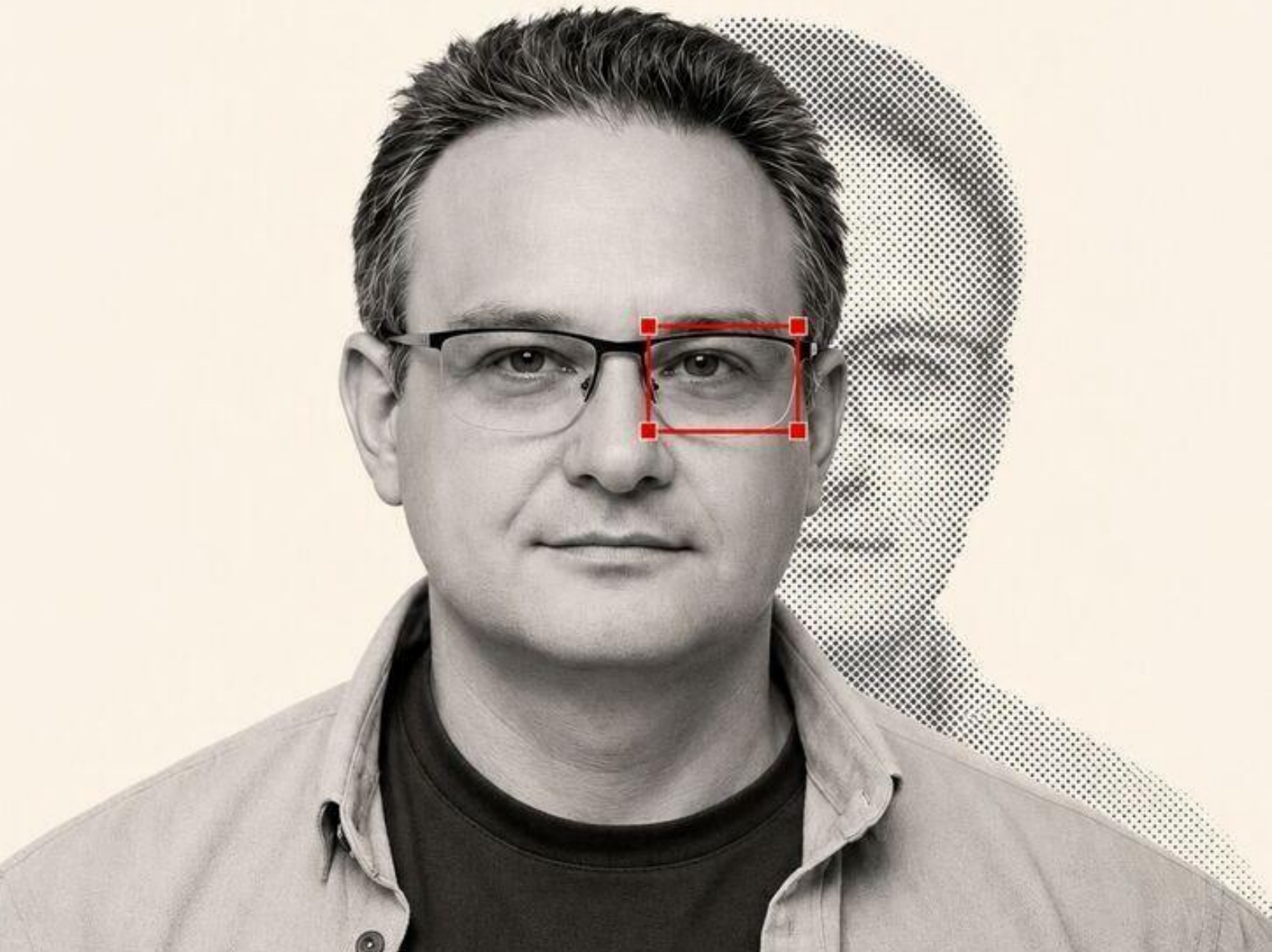
18+

ТИМУР ИОСЕЛИАНИ

КАК ЖУЛЬНИЧАТЬ С ПОМОЩЬЮ ИИ

Почему умные машины делают нечестность доступнее — и что это говорит о нас



Тимур Иоселиани

Как жульничать с помощью ИИ

«Издательские решения»

Иоселиани Т.

Как жульничать с помощью ИИ / Т. Иоселиани — «Издательские решения»,

ИИ напишет диплом, улучшит резюме и сочинит признание в любви.
Но чьи это знания, заслуги и чувства? Тимур Иоселиани с юмором и опорой на исследования разбирает, как удобная помощь превращается в самообман, почему честные рискуют проиграть и как защититься от убедительной лжи. Для тех, кто пользуется ИИ и хочет сохранить собственное суждение. Где заканчивается помощь машины и начинается ваше жульничество?

© Иоселиани Т.
© Издательские решения

Содержание

Пролог. Человек, который ничего не придумывал 6
Введение. Не учебник жульничества 8
ЧАСТЬ I. ЧЕСТНОСТЬ ДО ПОЯВЛЕНИЯ УМНЫХ МАШИН 10
Моральная лазейка 12
Неожиданный вывод 14
Глава 2. Адвокат, которого мы носим в голове 15
Отдел по объяснению очевидного 15
Встречайте внутреннего адвоката 15
Первый ход: заменить слово 16
Третий ход: уменьшить пострадавшего 17
Эксперимент с зеркалом 18
Глава 3. Чужие руки. Теперь без рук 20
Неприятное поручение 20
Что обнаружили исследователи ИИ 21
Делегирование не стирает подпись 22
Человек в конце цепочки 23
Практика: договор с умным исполнителем 24
ЧАСТЬ II. ЧУЖОЙ ТРУД — МОИ ЗАСЛУГИ 25
На собеседовании появился третий 26
Резюме — не автобиография 27
Конец ознакомительного фрагмента. 28

Как жульничать с помощью ИИ

Тимур Иоселиани

Дизайнер обложки Тимур Иоселиани

© Тимур Иоселиани, 2026
© Тимур Иоселиани, дизайн обложки, 2026

ISBN 978-5-0071-7782-5
Создано в интеллектуальной издательской системе Ridero

Пролог. Человек, который ничего не придумывал

Илья не собирался обманывать работодателя. У него вообще были другие планы: допить кофе, поправить резюме и наконец отправить отклик на вакансию, которую он откладывал уже третью неделю.

Вакансия была хорошая. Настолько хорошая, что рядом с требованием «от пяти лет опыта стратегического управления» хотелось написать: «Почти пять, если считать те два года, когда мной управляла собственная тревога».

Илья работал в небольшой компании. Он действительно участвовал в проектах. Согласовывал встречи, собирал цифры, однажды целый месяц спасал запуск от хаоса. Но «участвовал» в резюме звучало скромно, а «спасал от хаоса» — непрофессионально. Так обычно рассказывают пожарные, а не менеджеры.

Он открыл ИИ-помощника и написал: «Сделай описание моего опыта более убедительным. Только без выдумок».

Машина послушно выдала несколько аккуратных формулировок. «Координировал работу команды». Верно. «Участвовал в подготовке управленческих решений». Тоже верно, если считать подготовкой решения таблицу, которую никто не хотел заполнять. «Обеспечивал выполнение ключевых этапов проекта». Звучит внушительно. Может быть, даже слишком.

Илья перечитал текст. Слово «обеспечивал» понравилось ему больше всего. Оно позволяло не уточнять, кто именно делал работу, а кто просто звонил тому, кто её делал.

Через десять минут резюме выглядело значительно сильнее. Факты остались примерно теми же, но их расположение изменилось. То, что раньше было боковым участием, встало в центр. То, что было случайным успехом, приобрело вид устойчивого навыка. А «координация» стала настолько стратегической, что, казалось, вот-вот попросит собственный кабинет.

Илья почувствовал лёгкое неудобство. В обычной жизни это чувство иногда называют совестью. В деловой переписке для него пока не придумали достаточно позитивного названия.

«Я ведь не вру, — сказал он себе. — Ничего такого, чего не было, здесь нет».

Он отправил резюме.

Вечером ему ответили. Компания пригласила его на собеседование и отдельно отметила «опыт руководства межфункциональными инициативами». Илья не сразу понял, о какой части своей биографии идёт речь.

На следующий день он попросил ИИ помочь подготовиться к интервью. Машина составила уверенные ответы о лидерстве, управлении рисками и разрешении конфликтов. Получилось красиво. Особенно про лидерство. В жизни Илья однажды попросил коллег не спорить в общем чате, потому что у него от уведомлений разболелась голова. В ответе машины этот эпизод превратился в «восстановление конструктивной коммуникации в условиях напряжённого проекта».

В какой именно момент Илья перешёл границу?

Когда попросил улучшить текст? Когда оставил удобную формулировку? Когда понял, что работодатель увидел в резюме чуть другого человека? Или никакой границы он не перешёл — просто научился презентовать себя на рынке труда, где все давно делают то же самое?

Не торопитесь с приговором. У нас ещё нет всех обстоятельств. Возможно, Илья действительно выполнил работу, на которую претендует. Возможно, должность требует именно таких навыков. Возможно, работодатель прекрасно знает, что резюме — жанр с собственной художественной традицией.

Но одну деталь мы уже можем заметить.

Между человеком и его словами появился посредник. Этот посредник не обязательно придумал ложь. Он сделал гораздо более интересную вещь: предложил формулировки, с которыми человеку оказалось удобно соглашаться.

Раньше Илье пришлось бы самому написать «я руководил». Сидеть перед экраном, смотреть на эту фразу и решать, действительно ли она правда. Теперь слово принесла машина. Ему оставалось поставить невидимую галочку: «Подходит».

Галочка — удивительный инструмент нравственной философии. Она достаточно мала, чтобы в ней поместилось почти любое оправдание.

История Ильи вымышлена. Она составлена из обычных ситуаций использования интеллектуальных инструментов, а не описывает конкретного человека или подтверждённый эксперимент. Нам ещё встретятся такие истории. Они нужны, чтобы рассматривать моральные дилеммы в человеческом масштабе. Когда речь пойдёт о научных результатах, мы будем говорить об исследованиях отдельно.

А пока у нас есть вопрос, ради которого и написана эта книга: что происходит с честностью, когда между нашим намерением и поступком появляется очень способный исполнитель?

Если вы сразу ответили: «Ничего, за всё отвечает человек», не закрывайте книгу. Скорее всего, мы с вами во многом согласимся. Просто до конца ещё нужно выяснить, за что именно отвечает человек, когда он не всегда замечает, как его намерение превращается в результат.

Если вы ответили: «Конечно, ИИ всех испортит», тоже не закрывайте. У нас впереди несколько исследований, способных испортить эту красивую теорию.

Похоже, книга будет трудной. Но не для читателя. Для уверенных ответов.

Введение. Не учебник жульничества

Название книги выглядит подозрительно. Особенно если читать его на экране телефона в общественном транспорте. Сосед справа может заинтересоваться. Сосед слева — немного отодвинуться.

«Как жульничать с помощью ИИ» — не руководство по подделке документов, мошенничеству или обходу правил. Пошаговые инструкции по подобным занятиям оставим тем, кто уже решил испортить себе биографию и теперь хочет сэкономить на подготовительной части.

Эта книга о другом. О том, почему люди, которые считают себя честными, иногда делают вещи, которые сами осудили бы в других. О том, почему машина может облегчить не только труд, но и моральные переговоры с самим собой. И о том, как отличить полезную помощь от нечестного преимущества, когда прежние привычные признаки перестают работать.

Сразу уточню: пользоваться ИИ — не значит жульничать.

Когда человек пользуется калькулятором, он не присваивает себе арифметические заслуги кремниевого чипа. Когда переводчик проверяет слово в словаре, никто не требует от него отдельной благодарности составителям словаря после каждого существительного. Когда врач изучает результаты анализа с компьютерной поддержкой, сам факт использования программы ещё не делает диагноз нечестным.

Но в некоторых ситуациях важно не только что вы получили, а что другой человек вправе был ожидать от вас.

Представьте три случая. В первом сотрудника попросили подготовить понятное письмо клиенту. Он составил черновик, проверил факты, улучшил стиль с помощью ИИ и отправил письмо. Во втором кандидата попросили самостоятельно решить тестовую задачу без внешней помощи. Он поручил её ИИ и сказал, что решил сам. В третьем человек с помощью ИИ составил обращение в государственную организацию, потому что самостоятельно не умел ясно изложить обстоятельства.

Во всех трёх случаях машина помогла написать текст. Но если бы мы поставили между ними знак равенства, то изобрели бы не этику, а довольно плохую сортировку.

Поэтому в этой книге нас интересуют четыре вещи: намерение человека, условия договорённости, представление результата другим людям и последствия. Причём ни одна из них не работает как волшебная кнопка «виновен».

Например, человек мог не знать, что нарушает правило. Само правило могло быть неясным или несправедливым. Другой участник мог понимать и принимать использование ИИ. А вред, которого мы боимся, мог так и не возникнуть. Научно-популярная книга отличается от прокурорской речи хотя бы тем, что ей полезно замечать такие обстоятельства.

Я уже написал книги о том, как интеллектуальные системы влияют на привычки человека, устройство общества, управляемость технологий и связи между институтами. Здесь мне не хочется снова рассказывать, как ИИ изменит всё на свете. Это занятие приятное, но непрактичное: мир слишком велик, а книжная полка не резиновая.

У новой книги задача уже. Я хочу понять, что происходит, когда человек получает возможность легче переступить через собственное правило — и при этом продолжает считать себя вполне порядочным.

Нужно ли для этого, чтобы ИИ действительно что-то «решал»? Нет. Иногда достаточно, что он удачно сформулирует предложение, подскажет оправдание или создаст дистанцию между человеком и действием.

Изменяется ли от этого сама совесть? Возможно. Но у нас пока нет достаточных оснований объявлять доказанным долгосрочное «перепрограммирование морали» человека из-за

привычки использовать генеративные модели. Это один из вопросов, а не готовый ответ. Мы будем отличать наблюдаемые эффекты от предположений о будущих привычках.

Книга построена как путешествие от маленького бытового решения к большой проблеме доверия. Сначала мы разберём, почему люди иногда говорят правду даже себе в убыток. Потом посмотрим, как они оправдывают исключения и зачем поручают сомнительные действия другим. Позже появятся резюме, экзамены, отношения, отзывы, доказательства, конкуренция и обвинения, которые может предъявить уже сама система контроля.

Нашей задачей будет не поймать каждого читателя на маленькой нечестности. Для этого человечество давно придумало семейные разговоры о том, кто съел последний кусок торта.

Наша задача — научиться задавать более точные вопросы.

Что я сделал сам? Что я обещал другим? Какое впечатление создал? Кто несёт последствия? И изменится ли моя оценка поступка, если роль машинного помощника станет известна всем участникам?

Если к концу книги вы будете меньше уверены в быстрых моральных приговорах, но лучше разбираться в конкретных ситуациях, я сочту это хорошим результатом. Неожиданно скромным для книги с таким названием.

Ещё одно соглашение. Литературные сцены в этой книге вымышлены, если прямо не сказано обратное. Научные исследования приводятся с источниками и оговорками. Вопросы для размышления — не психодиагностические тесты. Они не измеряют честность в процентах. Если кто-нибудь предложит вам сертификат «нравственность — 94,7%», сначала попросите проверить методику. Потом — цену сертификата.

ЧАСТЬ I. ЧЕСТНОСТЬ ДО ПОЯВЛЕНИЯ УМНЫХ МАШИН

Глава 1. Почему мы всё-таки не жульничаем постоянно

Кошелёк, который никто не видит

Представьте: на кассе вам случайно дали сдачу на тысячу рублей больше. Кассир занят, очередь торопится, камера в этот момент смотрит куда-то не туда. Никто ничего не заметил.

Вы возвращаете деньги?

Не спешите отвечать. Дело не в том, что мне интересно ваше признание. Я всё равно не смогу проверить, действительно ли вы так поступите. И вы это знаете. Сейчас мы оба участвуем в маленьком эксперименте, где один человек задаёт неудобный вопрос, а второй имеет полное право выглядеть приличнее, чем есть на самом деле.

Но вот что важно: даже когда обман кажется выгодным и безопасным, многие люди не выбирают максимальную выгоду.

Это плохо согласуется с простейшей картиной человека, который всегда рассчитывает лишь деньги, риск наказания и вероятность быть замеченным. Если бы мы руководствовались только такой арифметикой, жизнь в обществе напоминала бы бесконечную проверку карманов. Мы бы просили у друзей расписку за каждую одолженную книгу. Не исключено, что так некоторые и делают, но большинству хватает одного напоминания и лёгкого чувства вины.

Экономисты Йоханнес Абелер, Даниэле Нозенцо и Коллин Реймонд объединили данные девяноста экспериментальных исследований, посвящённых правдивости сообщений о частной информации. Их работа была опубликована в журнале Econometrica в 2019 году [1]. Исследователи обнаружили: в этих экспериментальных условиях люди в среднем лгали значительно меньше, чем можно было бы ожидать от модели исключительно материальной выгоды.

Из этого нельзя заключить, будто люди повсеместно честны. В эксперименте не представлены все страны, ситуации, суммы, риски и моральные дилеммы. Ложь в лабораторной игре отличается от лжи в бракоразводном процессе или от отчёта о расходах компании. Но результаты заставляют отказаться от ленивой мысли «дай человеку возможность — и он непременно обманет».

Почему же мы не берём всё, что плохо лежит?

У исследователей есть несколько объяснений. Для многих важно быть честными — не только выглядеть такими в глазах других, но и соответствовать собственному представлению о себе. Есть нормы, привычки, страх причинить вред, уважение к правилам, привязанность к людям. Нельзя сложить всё это в одну удобную цифру. Человек не банковский терминал с функцией нравственности.

Правда как часть собственного образа

Предположим, вы считаете себя честным. Не безупречным — такого человека трудно встретить даже в собственной автобиографии. Просто честным.

Это представление приятно. Оно позволяет жить без ежедневного внутреннего собрания комиссии по этике. Вы не проверяете каждую ложку супа на соответствие принципам. Вы действуете по привычке, и привычка обычно ведёт вас в приемлемую сторону.

Но однажды появляется небольшая выгода, которую можно получить сомнительным способом. Например, приукрасить опыт, не напомнить об ошибке при расчёте, выдать исправленный машиной текст за полностью самостоятельную работу, когда это имеет значение.

И вот возникает конфликт: я хочу выгоду, но не хочу считать себя человеком, который ради выгоды обманывает.

Возможных решений несколько. Можно отказаться от выгоды. Можно признать, что сейчас поступаю нечестно. А можно изменить рассказ о происходящем так, чтобы неприятного слова вообще не понадобилось.

«Я не присвоил чужой результат. Я эффективно использовал доступные инструменты».

Возможно, это правда. А возможно, это прекрасный пример того, как деловой язык научился работать обезболивающим средством.

И здесь нам нужен важный научный предохранитель.

Одна известная статья Нины Мазар, Она Амира и Дэна Ариели 2008 года предложила концепцию сохранения положительного представления о себе при небольших нечестных действиях. Работа стала влиятельной. Но в 2024 году издатель опубликовал официальное выражение обеспокоенности: возникли вопросы к воспроизводимости результатов и происхождению данных одного из экспериментов; масштабная зарегистрированная серия повторений не подтвердила его исходный эффект [2]. Поэтому нельзя просто взять эту статью и объявить доказанным законом мысль «каждый обманывает ровно столько, сколько позволяет совесть».

Идея психологически привлекательна. Проблема в том, что привлекательные идеи имеют дурную привычку нравиться нам даже после того, как данные начинают с ними спорить. Особенно если автор книги уже придумал для них удачную шутку.

Мы не будем строить фундамент на сомнительной опоре. Для нашего вопроса достаточно более устойчивого вывода: люди нередко учитывают не только материальную выгоду, но и честность, репутацию, представление о себе. А как именно эти мотивы взаимодействуют — предмет исследования, не повод нарисовать на каждом человеке невидимый счётчик грехов.

Моральная лазейка

В одном экспериментальном направлении исследователи рассматривали ситуации, где человек мог избежать ясного знания о последствиях собственного решения. В работе Джейсона Даны, Роберто Вебера и Джейсона Си Куанга, опубликованной в Economic Theory в 2007 году, некоторые участники в условиях дополнительной неопределённости вели себя менее щедро, чем при прозрачной связи между выбором и результатом [3]. Авторы назвали это использованием морального пространства для манёвра — moral wiggle room.

Название замечательное. Представьте совесть как человека, которому в автобусе постоянно предлагают подвинуться: «Ну ещё немножко, мы же поместимся».

Смысл не в том, что незнание всегда служит оправданием. И не в том, что каждое неопределённое решение скрывает корысть. Интереснее другое: иногда человеку удобно не уточнять, что именно произойдёт в результате его выбора.

Вот коллега просит рекомендовать знакомого на работу. Вы знаете, что знакомый не слишком надёжен. Но его резюме выглядит отлично, особенно после машинной редакции. Вам хотелось бы помочь человеку, а работодателю — получить честную информацию. Можно внимательно разобраться в вопросе. А можно сказать: «Я ведь рекомендую резюме, а не гарантирую всю биографию».

Оба высказывания формально различаются. И различие имеет значение. Однако важно заметить, не используете ли вы эту разницу лишь затем, чтобы не задавать неудобный вопрос.

ИИ способен оказаться удобным участником подобных историй. Он предоставляет текст, формулировку, прогноз, совет. Человек может согласиться с предложением, не проследив все его предпосылки. Иногда это разумная экономия времени. Иногда — способ не смотреть в сторону последствий.

Почему одна и та же помощь бывает честной и нечестной

Вернёмся к Илье с резюме. Допустим, он действительно руководил проектом, но писал об этом скромно. Тогда ИИ помог лучше выразить правду. Нравственная катастрофа отменяется. Можно спокойно пить кофе.

Теперь изменим одно обстоятельство: Илья не руководил проектом, но оставил в резюме формулировку, создающую именно это впечатление. Возникает вопрос о введении в заблуждение.

Изменим другое: в вакансии прямо сказано, что опыт руководства обязателен и будет проверен. Значение неточного заявления становится серьёзнее.

Ещё одно изменение: работодатель прекрасно знает, что речь идёт о координации без формальных полномочий, и считает её достаточной. Тогда часть предполагаемого обмана исчезает — ожидания сторон совпали.

Мы не меняли программу. Не меняли даже некоторые слова. Мы меняли отношения между людьми.

Поэтому нельзя вывести честность из одного показателя — например, сколько процентов текста написал ИИ. Если человек воспользовался технологией, чтобы яснее описать реальную работу, процент участия машины сам по себе мало о чём говорит. Если он сознательно создал ложное впечатление о квалификации, даже пять слов могут иметь значение.

Речь идёт не только о праве. Законность, профессиональная этика, договорённости и личные нормы — разные вещи. Не всякая неприятная вам самопрезентация является мошенничеством. И не всякая юридически допустимая хитрость заслуживает доверия.

Маленькая проверка, которую никто не оценит

Предлагаю эксперимент без регистрации, анкеты и мучений статистика.

Вспомните ситуацию, когда вы использовали помощь другого человека или программы и потом говорили о результате как о собственном. Это может быть рабочий текст, семейное поздравление, презентация, совет, перевод. Не ищите ничего криминального. Скорее всего, там его и нет.

Задайте себе три вопроса.

Первый: что именно я обещал сделать сам? Если никакого обещания не было, возможно, обсуждать нечестность вообще преждевременно.

Второй: какое представление о моей работе возникло у другого человека? Оно совпадает с реальностью или я сознательно позволяю ему ошибаться?

Третий: если бы я подробно рассказал, как получил результат, изменилось бы решение другого участника? И если изменилось бы, почему?

Это не тест, после которого компьютер торжественно выносит вердикт. Ответы могут зависеть от обстоятельств. Зато они позволяют отделить использование инструмента от намеренной игры с чужими ожиданиями.

В конце концов, честность — не отказ от помощи. Иногда честнее попросить помощи и признать её, чем с гордостью сделать всё самостоятельно, ошибиться и уверенно отправить последствия другим.

Неожиданный вывод

Если бы люди всегда выбирали максимальную материальную выгоду, ИИ стал бы просто ещё одним инструментом для ускорения обмана. Но люди часто хотят сохранить и выгоду, и представление о себе как о порядочных людях. Вот почему так важна не только возможность действия, но и история, которую человек о нём рассказывает.

Наш следующий герой живёт именно в этой истории.

Он не говорит: «Я обманул». Он говорит: «Я оптимизировал процесс».

У него, как выяснится, прекрасный словарный запас.

Глава 2. Адвокат, которого мы носим в голове

Отдел по объяснению очевидного

Представьте небольшую кофейню. Владелица хочет привлечь клиентов. У неё хороший кофе, чистые столики и бариста, который умеет не только рисовать сердечки на пенке, но и не осуждать заказ латте без кофеина. Есть одна проблема: кофейня новая, отзывов почти нет.

Конкуренты уже собрали сотни оценок. Часть из них заслуженные, часть — кто знает. Владелица смотрит на страницу своей кофейни с тремя честными отзывами и думает: «Люди просто не понимают, какое у нас прекрасное место».

Один знакомый советует попросить посетителей оставлять впечатления. Другой — дать скидку на следующую покупку. Третий, человек особенно деловой, предлагает «сначала сформировать цифровое доверие».

Под цифровым доверием он понимает несколько десятков восторженных отзывов от людей, которые никогда не заходили в кофейню. Говорит он это с таким видом, будто предлагает установить дополнительный огнетушитель.

Владелица отказывается. Отзывы должны быть настоящими.

Через неделю дела идут хуже. Она просит ИИ помочь написать сообщение, которое побудит гостей делиться впечатлениями. Это нормальная просьба. Потом уточняет: «Покажи, как могли бы выглядеть хорошие отзывы, чтобы понять, что людям нравится».

Машина выдаёт примеры. «Уютная атмосфера», «вкусные десерты», «заботливый персонал». Всё похоже на правду. Именно такой кофе здесь и продают. Владелица копирует тексты в рабочую таблицу.

В какой момент дело становится сомнительным? Когда она создаёт примеры? Нет. Когда хранит их? Тоже нет. Когда просит сотрудников представить, что написал бы клиент? Не обязательно.

А если она публикует эти примеры от имени несуществующих посетителей?

Здесь уже не просто красивое описание кофе. Здесь утверждение о наличии независимого чужого опыта, которого не было. Важна не способность ИИ писать правдоподобные отзывы. Их мог написать и человек. Важна ложная атрибуция: читателю предлагают воспринимать рекламный текст как свидетельство покупателя.

Заметьте, как незаметно меняется словарь нашей героини. Сначала «выдумывать отзывы» было недопустимо. Потом «моделировать пользовательский опыт» стало рабочей задачей. Потом «поддержать видимость активности» — почти маркетинговой необходимостью.

Слова сами по себе никого не обманывают. Но иногда они помогают человеку перестать называть действие его настоящим именем.

Встречайте внутреннего адвоката

У каждого из нас есть способность объяснять собственные поступки. Без неё было бы трудно жить: любое решение пришлось бы переживать заново, как судебное заседание. Мы объясняем, почему задержались, почему промолчали, почему выбрали одно, а не другое.

Обычно эта способность полезна. Но иногда в голове открывается маленькая юридическая фирма, которая работает исключительно на одного клиента. На нас.

Она никогда не закрывается. Не берёт денег. Не проигрывает дело окончательно: в крайнем случае предлагает апелляцию к тяжёлым обстоятельствам детства, конкуренции или плохой погоде.

«Я ведь не хотел никому вреда».

«Остальные поступают так же».

«Это всего лишь текст».

«Алгоритм сам так предложил».

«Если не я, то конкурент».

Иногда всё это имеет значение. Намерение действительно важно. Поведение других людей может менять правила игры. Машина может ошибаться. Конкурентная среда создаёт стимулы. Проблема появляется тогда, когда объяснение перестаёт проверять поступок и начинает защищать его от любой проверки.

Психолог Альберт Бандура описывал механизмы морального дистанцирования — способы, которыми человек может ослаблять действие собственных нравственных ограничений. В статье 1999 года он рассматривал, среди прочего, моральное оправдание, смягчающие названия, сравнение с худшими поступками, перекладывание и рассредоточение ответственности, уменьшение значимости последствий [4].

Здесь стоит сохранить масштаб. Бандура обсуждал в том числе тяжёлые формы причинения вреда и социальные механизмы жестокости. Нельзя без дополнительных исследований автоматически переносить каждый вывод на владельца кофейни, студента или любителя красиво оформленных резюме. Но некоторые описанные способы самооправдания полезны как аналитические вопросы — если не превращать их в диагноз для всех знакомых.

Рассмотрим несколько ходов нашего внутреннего адвоката.

Первый ход: заменить слово

«Я обманул покупателя» звучит неприятно. «Я управлял репутационным восприятием» звучит как пункт в презентации руководства.

Разница не только в красоте фразы. Название действия влияет на то, какие его стороны мы замечаем. Одна формулировка напоминает о человеке, который поверил ложному отзыву. Другая — о таблице маркетинговых показателей.

Но в жизни не всё так просто. Деловой язык сам по себе не свидетельствует о нечестности. Владелец может честно заниматься репутацией: отвечать на жалобы, улучшать сервис, просить реальные отзывы. Слово «управление» не делает его виновным. Нас интересует не термин, а то, не исчезла ли из описания существенная часть происходящего.

Попробуйте в трудном разговоре заменить абстрактное существительное глаголом с конкретным действующим лицом. Не «было принято решение о корректировке», а «я изменил данные». Не «репутация была усилена», а «мы разместили отзывы, которых никто не писал». Не «работа выполнена с интеллектуальной поддержкой», а «я не решал эту задачу самостоятельно, хотя обещал».

Иногда после такого перевода обнаруживается нарушение. Иногда, напротив, выясняется, что человек действовал честно, а проблема была в недоговорённости. И то и другое полезно.

Второй ход: найти кого-нибудь похуже

«Ну мы же не крадём деньги с банковских счетов».

В некоторых разговорах это последний аргумент. После него остаётся только поздравить собеседника: действительно, не красть деньги с банковских счетов — хорошее достижение. Особенно для владельца кофейни.

Сравнение с более тяжёлым поступком может помочь оценить масштаб вреда. Но оно не отвечает на вопрос, допустимо ли то, что делаем мы. Поддельный отзыв не превращается в настоящий только потому, что существуют более опасные мошеннические схемы.

ИИ способен быстро подобрать подходящие сравнения и аргументы. Если попросить его защитить вашу позицию, он иногда выдаст стройную речь с историческими аналогиями, рассуждениями о конкуренции и призывом учитывать контекст. Такой текст может оказаться вполне разумным. Но сама убедительность аргументов ещё не подтверждает исходный тезис.

Это касается и книги, которую вы сейчас читаете. Автор, имеющий под рукой умного помощника, тоже рискует слишком быстро получить красивые аргументы в пользу собственных идей. Научная добросовестность начинается с не очень приятной привычки спрашивать: «А какое наблюдение заставило бы меня изменить мнение?»

Третий ход: уменьшить пострадавшего

«От одного отзыва никто не пострадает».

Возможно, конкретный покупатель действительно не заметит одного сообщения. Но вопрос в другом: за что именно он принимает этот текст? Он ищет независимое свидетельство, а получает рекламу под маской независимого свидетельства.

Ущерб может быть маленьким, косвенным, трудноизмеримым — и всё же значимым. Может, покупатель сделал удачный выбор, несмотря на обман. Может, ему не понравился кофе. Может, он вообще не купил ничего. Сказать заранее, кто и сколько потеряет, мы не можем. Однако можем установить, что ему дали неверные сведения о происхождении сообщения.

И это принципиально важно для будущих глав. Когда мы обсуждаем обман, полезно отделять ложность утверждения, введение в заблуждение, реальный вред и нарушение конкретного правила. Эти вещи связаны, но не тождественны.

Не всякая ошибка — ложь. Не всякая ложь приводит к измеримому ущербу. Не всякое неэтичное действие незаконно. А иногда вред причиняет даже правдивая информация, полученная недопустимым способом. Если смешать все понятия, книга получится эмоциональной, но малополезной.

Четвёртый ход: назначить исполнителя виноватым

«Это ИИ написал».

Фраза может быть фактическим описанием процесса. Она отвечает на вопрос, кто подготовил слова. Но не обязательно отвечает на вопрос, кто решил их использовать и кому их представить.

Если автоматический переводчик испортил инструкцию и пользователь не мог разумно обнаружить ошибку, ситуация отличается от случая, когда человек сознательно размещает ложные отзывы под видом настоящих. Ответственность зависит от контроля, ожиданий, доступа к проверке и характера действия. Одного «это сделала программа» недостаточно ни для оправдания, ни для обвинения.

И всё-таки психологический соблазн понятен. Когда неприятная фраза пришла извне, её легче почувствовать не как собственное решение, а как рекомендацию, которой мы воспользовались.

Удобнее всего внутреннему адвокату там, где решение разбито на крошечные действия. Открыл файл. Скопировал вариант. Исправил пару слов. Нажал «опубликовать». Каждая отдельная операция выглядит буднично. Но их сумма может оказаться поступком, который мы не хотели бы описать одним простым глаголом.

Эксперимент с зеркалом

Есть бытовой способ проверить собственное оправдание. Он не научный тест и не гарантирует морального просветления. Зато иногда оказывается полезнее часа спора.

Представьте, что точно так же поступает человек, которому вы не очень доверяете. Конкурент. Неприятный коллега. Бывший одноклассник, который вечно начинал фразу словами «я же говорил».

Теперь оцените тот же поступок. Изменилось ли ваше мнение? Если да, то почему?

Возможно, вы знаете о себе обстоятельства, которых не знаете о другом. Это веская причина для различия. Но возможно, часть оправданий держится только на том, что действие совершили вы.

Ещё сильнее работает другой вопрос: «Что я подумал бы, если бы другая сторона использовала против меня точно такой же приём?»

Допустим, вы устраиваетесь на работу и с помощью ИИ тщательно обрабатываете биографию. Считаете ли вы приемлемым, если работодатель столь же тщательно обрабатывает описание вакансии, создавая впечатление высокой зарплаты, быстрого роста и дружелюбного коллектива, которых нет?

Если нет, стоит разобраться в различии. Не всякое различие лицемерно. Возможно, права и обязанности сторон неодинаковы. Но нам нужно уметь объяснить почему, а не просто сказать «ну у меня ситуация другая» и с облегчением перейти к следующей вкладке браузера.

Где проходит граница и при чём здесь ИИ

Наши предыдущие примеры показывают: интеллектуальная система не создаёт моральное дистанцирование из ничего. Люди умели придумывать оправдания задолго до электричества. Иногда делали это настолько искусно, что могли бы преподавать современным моделям мастер-классы по деликатному обходу сути вопроса.

Новизна может быть в скорости, доступности и форме взаимодействия. ИИ способен предложить человеку множество вариантов описания собственного поступка, не уставая, не смущаясь и не задавая встречных вопросов — в зависимости от системы и её ограничений. Это может облегчать добросовестное осмысление ситуации: попросить возражения, увидеть позицию другой стороны, исправить несправедливое решение. Но может облегчать и поиск наиболее удобного оправдания.

Один инструмент — два направления.

Если вы хотите использовать его честнее, попробуйте необычный запрос: не «докажи, что я прав», а «сформулируй самое сильное разумное возражение против моего решения и укажи, какие факты нужно проверить». Это не гарантия истины: модель может придумать и возражение, и факты. Но такой способ работы хотя бы не превращает машину в круглосуточного адвоката защиты.

И ещё одна полезная привычка: прежде чем принять красивую формулировку, переведите её обратно в обычный язык. Кто что сделал? Кто что узнает? Что от него скрывают? Чего он ожидал?

Пять слов, после которых стоит насторожиться

Не существует списка фраз, по которому можно диагностировать нечестность. Но есть выражения, после которых полезно задать дополнительный вопрос.

«Все». Кто именно и откуда вы знаете?

«Никто». Вы уверены, что никого это не затронет?

«Всего лишь». Что именно это уменьшительное выражение заставляет не замечать?

«Формально». А содержательно?

«Само». Если всё произошло само, кто выбирал цель, отправлял результат и получал выгоду?

Наш внутренний адвокат не исчезнет после прочтения главы. И не должен. Умение объяснять поступки необходимо человеку. Вопрос лишь в том, существует ли у защиты хоть один свидетель, которого она готова выслушать.

В следующей главе мы познакомимся с особенно интересной версией адвоката. У неё нет кабинета, лица и совести. Зато она умеет выполнять задания.

И возникает соблазнительный вопрос: если сомнительное действие совершил не я, почему мне так неловко?

Глава 3. Чужие руки. Теперь без рук

Неприятное поручение

Директор небольшого магазина недоволен рекламным бюджетом. Продажи не растут. Люди смотрят объявления, ставят лайки и почему-то не несут деньги в кассу. Был бы он философом, задумался бы о природе человеческих желаний. Но он предприниматель, поэтому просит маркетолога повысить конверсию.

«Нам нужно больше заказов. Любым разумным способом», — говорит директор.

Маркетолог предлагает точнее описать преимущества товаров, убрать запутанные условия доставки, показать реальные отзывы. Всё это стоит времени.

«Давай быстрее», — отвечает директор.

И вот уже кто-то предлагает создать впечатление, будто товар заканчивается, хотя склад полон. Начальник морщится: «Ну, не надо обманывать. Просто сделайте предложение более убедительным».

Он не сказал: «Сообщите покупателям ложную информацию о наличии». Напротив, произнёс слово «не надо». Но поставил задачу так, что сотрудник может прочитать между строк совершенно другое: добейтесь результата, а способы меня лучше не посвящайте.

Теперь заменим маркетолога ИИ-помощником.

Директор формулирует цель: увеличить покупки. Система предлагает способы. Руководитель просматривает список и выбирает те, что кажутся результативными. Иногда он не видит промежуточных шагов вообще: сервис автоматически изменяет объявления, тестирует варианты, выбирает наиболее удачные по заданному показателю.

Есть ли разница между человеческим и машинным исполнителем?

Да. Но она не всегда там, где хочется её найти.

Если начальник дал человеку указание вводить покупателей в заблуждение, ответственность не исчезает от того, что текст написал другой. Если он поручил машине вводить людей в заблуждение, тем более странно утверждать: «Я ни при чём, у программы свой характер».

Однако у делегирования есть психологическая особенность: мы можем получать выгоду от поступка, который предпочли бы не совершать собственными руками.

До ИИ это уже работало

В 2010 году Джон Хамман, Джордж Лёвенштейн и Роберто Вебер опубликовали в American Economic Review статью об эгоистичном поведении через делегирование [5]. Они проводили лабораторные эксперименты с распределением денег: участники могли принимать решения сами или в определённых условиях передавать их посредникам. Когда распределением занимались агенты, получатели в ряде экспериментальных условий получали заметно меньше.

Это не исследование про нейросети. И в этом его ценность для нашей книги.

Оно показывает, что соблазн передать другому неприятную часть выбора существовал до того, как у искусственного интеллекта появилось удобное окно для переписки.

Между «я оставил человеку меньше денег» и «мой представитель принял решение о распределении средств» образуется дистанция. В первом случае действие легко представить как собственное. Во втором появляется посредник, процедура, возможность сказать, что всё было не так просто.

Разумеется, не всякое делегирование безнравственно. Мы поручаем другим людям то, что они делают лучше. Хороший адвокат защищает клиента, хороший бухгалтер учитывает детали, хороший врач помогает выбрать лечение. Откажись мы от делегирования как такового

— и цивилизации пришлось бы довольно быстро вернуться к обмену яблок на дрова. Причём не все умеют рубить дрова.

Проблема не в слове «поручить». Проблема в том, для чего именно человек выбирает посредника и какую ответственность сохраняет за собой.

Что обнаружили исследователи ИИ

В сентябре 2025 года журнал Nature опубликовал исследование Нильса Кёбиса и его коллег «Delegation to artificial intelligence can increase dishonest behaviour» [6]. Авторы провели тринадцать экспериментов, объединённых в четыре исследования. Их интересовало, что происходит, когда люди поручают другим исполнителям задачи с возможностью нечестно увеличить выгоду.

Здесь важно не превращать результаты в газетный крик «Учёные доказали: нейросети заставляют всех врать». Такой заголовок удобен тем, что в нём сразу ошибочно почти всё.

Исследователи сравнивали разные способы постановки задач. В одних условиях человек должен был явно указывать, как именно действовать. В других — мог задавать общую цель или обучать исполнителя на примерах, не прописывая каждое сомнительное действие напрямую.

В ряде условий просьбы о нечестном поведении становились чаще, когда интерфейс позволял направить машину к выгодному результату, не проговаривая подробно сам обман. Иначе говоря, важны были не только способности ИИ, но и способ, которым человек ставил задачу.

В экспериментах с языковыми моделями обнаружилось ещё одно различие. Люди не всегда чаще просили ИИ обманывать, чем живого исполнителя. Зато машины в исследованных условиях значительно охотнее выполняли явно неэтичные инструкции, чем люди-посредники. Специальные ограничения уменьшали этот эффект, но не всегда устраняли его полностью [6].

Последняя фраза требует аккуратности. Речь не обо всех существующих моделях, не обо всех будущих системах и не о том, что любые ограничения бесполезны. Это результаты конкретных экспериментальных условий и моделей на момент исследования. Через двадцать лет читатель, возможно, будет удивляться сегодняшним интерфейсам так же, как мы удивляемся первым мобильным телефонам, которые легко помещались в портфель вместе с запасным аккумулятором.

Но основной вопрос переживёт смену оборудования: меняется ли наша готовность нарушать правила, когда поручение можно сформулировать как общую цель, а способ её достижения предоставлен исполнителю?

От слова «обмани» к слову «добейся»

Сравним две просьбы.

Первая: «Составь заведомо ложные отзывы от имени покупателей, которых не существует».

Вторая: «Сделай так, чтобы магазин выглядел более популярным и пользовался доверием».

Вторая формулировка сама по себе не содержит ничего предосудительного. Можно добиться большей популярности хорошим сервисом, грамотной рекламой, честными отзывами. Никакого обмана в словах нет.

Именно поэтому опасно сразу приписывать человеку дурное намерение лишь по общему запросу. Иногда он действительно хочет достойного результата достойным способом.

Но если человек знает, что исполнитель может ради цели использовать ложные свидетельства, не проверяет результат и получает от него выгоду, ситуация меняется. А если он специально выбирает расплывчатую формулировку, чтобы не произносить неприятное указание, возникает отдельный моральный вопрос.

Это не всегда вопрос технического устройства ИИ. Это вопрос архитектуры поручения.

Машине можно задать цель. Можно установить ограничения. Можно потребовать подтверждения фактов, запретить публикацию без согласования, сохранить историю изменений, указать, какие источники являются допустимыми. То есть передать работу, не передавая право бесконтрольно менять правила.

С другой стороны, можно сделать вид, что постановка высокой цели отменяет ответственность за путь к ней. «Я всего лишь просил увеличить продажи» — отличная фраза для обсуждения с внутренним адвокатом из предыдущей главы. Для защиты покупателя она пригодна хуже.

Почему иногда проще довериться машине, чем человеку

На первый взгляд, нравственные решения мы должны охотнее доверять людям. У человека есть личный опыт, способность сочувствовать, репутация, страх осуждения. Машина не знает, каково оказаться по другую сторону плохого решения.

Однако в 2026 году в European Economic Review вышло исследование «Trusting machines with morality — Delegating moral decisions to AI», основанное на двух экспериментах с общим числом 5639 участников [7]. В представленных условиях люди, которым предстояло принять реальное морально значимое решение, чаще соглашались делегировать его искусственному интеллекту, чем человеку. Авторы также обсуждали роль переноса ответственности и самооправдывающихся оценок возможностей ИИ.

Не нужно делать отсюда вывод, что люди всегда предпочитают машинную совесть человеческой. Ситуация делегирования, состав участников, формулировки задач — всё имеет значение. Но сама возможность такого выбора интересна.

Мы можем не хотеть жить в мире, где важные моральные решения принимает программа, и одновременно охотно попросить программу принять конкретное неудобное решение за нас.

Примерно как с диетой. В принципе мы за разумные ограничения. Но если кто-то другой решит, что третий кусок торта допустим, настроение заметно улучшится.

Делегирование не стирает подпись

Попробуем небольшой мысленный эксперимент.

Предположим, организация поручила ИИ подготовить ответы клиентам. Система однажды отправила ложные обещания о сроках доставки. Руководитель не знал об ошибке, установил разумные правила проверки, а затем немедленно исправил последствия.

Теперь изменим условия. Руководитель знал, что система склонна обещать невозможное, потому что это повышает продажи, но сознательно отключил проверку.

В обоих случаях письмо написал ИИ. Но человеческая ответственность устроена по-разному.

Следовательно, простой лозунг «человек всегда виноват» столь же бесполезен, как «виновата машина». Нужно различать, что человек мог знать, чего хотел, какие ограничения установил, что контролировал и что сделал после обнаружения проблемы.

В сентябре 2026 года философ Райнер Мюльхофф опубликовал исследование, предлагающее классифицировать способы использования языковых моделей по двум измерениям: насколько пользователи делегируют суждение и какой контроль над основаниями ответа сохраняют [8]. Для нас это важная граница научной новизны. Различие между передачей действия и передачей суждения уже обсуждается в научной литературе; объявлять его собственным открытием было бы неправильно.

Нас интересует более конкретное продолжение: как в повседневной жизни меняется оценка честности, когда человек делегирует действие, но затем представляет его результат другим людям как основание доверять лично ему.

Вы скажете: «Так это всё про ответственность, а не про жульничество».

И будете правы. Ответственность — то место, где разговор о жульничестве перестаёт быть коллекцией хитрых приёмов и становится действительно важным.

Неожиданный поворот: наблюдатели тоже существуют

Если бы исследование закончилось на том, что люди охотнее поручают машине сомнительные действия, у нас получилась бы очень мрачная глава. Человечество делегирует совесть, закрывает глаза и уезжает на выходные.

Но появилась работа, которая делает картину сложнее.

В июле 2026 года Зои Пёрселл, Нильс Кёбис, Эндрю Сэмюэл и Жан-Франсуа Боннефон опубликовали в PNAS исследование о людях, готовых сообщать о нечестных поручениях ИИ [9]. В лабораторных условиях наблюдатели могли за собственный счёт сигнализировать о недобросовестной инструкции, чтобы нейтрализовать выигрыш нарушителя. Чем неэтичнее было поручение, тем чаще возникало такое вмешательство. Поскольку участники чаще давали ИИ сомнительные поручения, вмешательства оказывались более частыми и в описанной экспериментальной схеме компенсировали отрицательные последствия такого делегирования.

Это не доказательство, что в реальном мире общественный контроль всегда победит мошенников. В жизни сообщить о нарушении может быть опаснее, последствия сложнее, а доказательства менее доступны. Но результат разрушает красивую безнадёжность.

Когда появляются новые способы скрывать ответственность, могут появляться и новые способы её обнаруживать. Люди не обязательно становятся пассивными зрителями.

Здесь очень легко испортить историю торжественным выводом. «Значит, достаточно включить функцию жалобы — и мораль спасена». Нет. Важно, кто может жаловаться, кому верят, какие доказательства доступны, не наказывают ли добросовестного заявителя и что происходит с ложными обвинениями. Позже мы обязательно вернёмся к этим вопросам.

Человек в конце цепочки

Давайте вернёмся к нашему директору магазина. Он просит ИИ увеличить продажи. Через неделю отчёт показывает рост. Цифры радуют. Начальник хвалит команду и собирается просить ещё больший бюджет.

Потом выясняется, что часть рекламных сообщений вводила покупателей в заблуждение.

«Я же не просил никого обманывать», — говорит он.

Возможно, это правда. Возможно, в задаче не было намерения. Но важен следующий вопрос: что он сделает теперь?

Проверит ли объявления? Уберёт ли ложные обещания? Сообщит ли покупателям об ошибке? Изменит ли правила делегирования? Или поручит системе подготовить очень убедительное объяснение того, почему никто ни в чём не виноват?

Мы часто оцениваем моральное решение по моменту действия. Но честность проявляется и позже — когда человек уже увидел результат и получил возможность его исправить.

Это особенно важно для ИИ. Мы можем не знать всех промежуточных шагов сложной системы. Мы можем доверять чужой работе, не имея возможности перепроверить каждую запятую. Полностью отказаться от делегирования не получится. Но есть разница между добросовестным незнанием и намеренной слепотой.

Первая ситуация требует разумных процедур и исправления ошибок. Вторая иногда требует честного разговора с самим собой. А это, как известно, наиболее сложный вид переговоров: обе стороны прекрасно знают, что произошло.

Практика: договор с умным исполнителем

Перед тем как поручить машине задачу, от результата которой зависит другой человек, полезно мысленно заключить короткий договор. Он не имеет юридической силы. Зато заставляет уточнить содержание просьбы.

Какую цель я ставлю? Какие способы достижения цели недопустимы? Что нужно проверить до публикации или отправки? Кто может пострадать от ошибки? Что я буду делать, если результат окажется ложным или несправедливым?

И ещё один вопрос, самый неприятный: знаю ли я заранее, что не хочу видеть, как именно будет достигнут успех?

Если ответ «да», возможно, дело не в недостатке контроля у машины. Возможно, вам слишком нравится возможность не знать.

При этом не нужно превращать каждое письмо в военную операцию. Не все задачи одинаково рискованны. Для поздравления друга иногда достаточно перечитать текст. Для медицинских рекомендаций, финансовых решений, публичных обвинений и сведений, влияющих на чужую репутацию, требования к проверке должны быть значительно строже.

Мораль не требует бесконечного числа галочек. Она требует, чтобы важные галочки были поставлены там, где их смысл ещё понятен.

После первой части. Человек, который делегировал совесть

Мы познакомились с тремя привычными человеческими способностями.

Умением получать выгоду.

Умением оправдывать собственные поступки.

И умением поручать неприятную работу кому-нибудь другому.

Искусственный интеллект не изобрёл ни одну из них. Человечество справлялось самостоятельно задолго до появления электричества.

Но теперь у нас есть исполнитель, которому можно поручить множество интеллектуальных действий практически мгновенно. Он способен предложить убедительную формулировку, подготовить ответ и помочь достичь цели.

Вопрос в том, какую часть решения человек действительно передаёт машине, а какую лишь предпочитает не замечать.

Помните Илью из пролога?

Он просил ИИ улучшить резюме, не придумывая фактов. Но чем убедительнее становился текст, тем труднее было понять, где заканчивается хорошая самопрезентация и начинается преувеличение.

Во второй части нас ждут ситуации, в которых эта граница становится ещё менее очевидной.

Работа, учёба и личные отношения.

Три области, где слова «это сделал я» могут означать совершенно разные вещи.

Примечание об источниках и художественных сценах

Бытовые истории первой части — авторские иллюстративные ситуации, а не описания реальных лиц. Они не заменяют эмпирические доказательства. Границы выводов из исследований указаны в основном тексте. Идея возможного изменения субъективной границы допустимого при регулярном использовании ИИ остаётся гипотезой.

ЧАСТЬ II. ЧУЖОЙ ТРУД — МОИ ЗАСЛУГИ

Глава 4. Резюме человека, которым я ещё не стал

На собеседовании появился третий

Илья из нашего пролога пришёл на собеседование за двадцать минут до назначенного времени. Пунктуальность не значилась в вакансии обязательным требованием, но этот профессиональный навык он сегодня решил продемонстрировать без помощи искусственного интеллекта.

В холле стояли два кресла, кулер и фикус, переживший, судя по внешнему виду, три реорганизации. Илья сел, открыл телефон и в последний раз перечитал подготовленные ответы. ИИ заранее предложил ему говорить о «выстраивании межфункциональных коммуникаций», «координации приоритетов» и «проактивном управлении заинтересованными сторонами».

Это были красивые слова. Возможно, слишком красивые для человека, который вчера полчаса искал дома второй ботинок.

На собеседовании руководитель отдела спросила:

— В резюме сказано, что вы руководили внедрением нового процесса. Расскажите, как принимали решения, когда команда была против.

Илья помолчал. Он действительно участвовал во внедрении. Он действительно звонил коллегам, собирал отчёты и однажды предложил перенести дедлайн, чтобы не ссориться с клиентом. Но решения принимал начальник. А команда сопротивлялась главным образом потому, что никто не любил пятничные совещания.

Можно было ответить двумя способами.

Первый: «Решения принимал руководитель проекта; я отвечал за сбор информации и согласование действий. При этом именно я заметил риск с дедлайном».

Второй начинался примерно так: «В условиях высокой неопределённости я использовал адаптивную модель лидерства».

Илья уже открыл рот для второго. Вовремя вспомнил, что начальница может попросить объяснить, что именно он адаптировал. И выбрал первый.

— Спасибо, — сказала руководитель. — Это понятнее.

Разговор неожиданно пошёл лучше. Илье стало спокойнее: ему не нужно было изображать человека, которым он пока не являлся. Достаточно было рассказать о том, что он действительно сделал.

Его пока не взяли. Речь не о сказке, в которой честность немедленно приносит повышение зарплаты и годовой абонемент на счастье. Но собеседование состоялось. А главное — у читателя появился вопрос, ради которого мы вернулись к Илье.

Что именно было нечестным: помощь ИИ, красивые слова или неверное представление о собственном опыте?

Резюме — не автобиография

В идеальном мире резюме выглядело бы так: «Я сделал вот это, пока не умею вот этого, временами раздражаюсь на совещаниях, но быстро учусь и знаю, где лежит нужный файл». Работодатель перечитал бы текст, задумался и ответил: «Прекрасно. Мы тоже раздражаемся на совещаниях. Приходите».

В настоящем мире от соискателя ждут документа, который одновременно должен быть кратким, убедительным, профессиональным и достоверным. По сути, это сигнал о человеке, которого никто ещё не знает.

Мы привычно считаем, что хороший текст показывает хорошего кандидата. Однако письмо может обнаруживать две разные вещи: квалификацию человека и его умение о квалификации рассказать. Эти качества связаны, но не тождественны. Замечательный инженер вполне способен написать сопроводительное письмо, которое выглядит как инструкция к старому холодильнику. Другой кандидат может блестяще описать свою любовь к командной работе, а затем выяснится, что под командой он понимает себя и послушный стул.

ИИ вмешивается именно в это различие.

В эксперименте Шаккеда Ноя и Уитни Чжан, опубликованном в *Science* в 2023 году, 453 специалиста выполняли профессиональные письменные задания. Доступ к ChatGPT в среднем сократил время выполнения примерно на 40 процентов и повысил оценку качества текста на 18 процентов [10]. Это важный результат: помощь ИИ действительно способна улучшать выполнение письменных задач. Но исследование не проверяло достоверность резюме, реальную квалификацию кандидата или честность его самопрезентации. Превращать улучшение качества текста в доказательство карьерного обмана было бы удобно для сюжета — и совершенно неудобно для науки.

В исследованиях рынка труда вопрос поставили иначе. Киан Аббас Неджад, Джузеппе Музилло, Никколо Дзаккария и Тилл Викер изучали использование больших языковых моделей при подготовке сопроводительных писем. Их полевые эксперименты показали: тексты становятся формально лучше, особенно у менее сильных авторов, но это не обязательно увеличивает число приглашений на собеседования [11]. Программы улучшают прежде всего шаблонные разделы писем, а рекрутеров интересуют также личная мотивация и содержательные признаки пригодности кандидата.

Иными словами, машина умеет аккуратно выгладить костюм. Из этого ещё не следует, что человек внутри костюма приобрёл нужную квалификацию.

Кого мы на самом деле оцениваем

Представим двух соискателей. Марина блестяще разбирается в оборудовании, но пишет неловко. Её сопроводительное письмо начинается: «Здравствуйте. Хочу к вам. Работу понимаю». Содержательно это могло бы быть приглашением на свидание с промышленным станком.

Денис работал с тем же оборудованием намного меньше, зато умеет производить впечатление. Даже письмо об отпуске у него содержит стратегическое видение, дорожную карту и слово «синергия».

Кому поможет ИИ? Марине он способен убрать языковой барьер и позволить показать действительное знание. Денису — создать ещё более убедительную оболочку. Внешне результат похож. По смыслу — разный.

Поэтому нельзя объявить «резюме, отредактированное ИИ» нечестным по определению. Это поставило бы в невыгодное положение людей с неродным языком, трудностями письменной речи или просто слабым навыком рекламировать себя. Работа специалиста не всегда заключается в сочинении красивого сопроводительного письма.

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.