

УГРОЗА ИИ



ЯН РОСС



От автора

Огромное спасибо Светлане и Юлии – за поддержку, вдохновение, редактуру, корректуру и прочие героические подвиги. И, само собой, за неоценимый вклад в мировую литературу.

Сайт автора: yanross.net

Угроза ИИ

– Так, внимание! – гаркнул я. – Стендап начинается. Кто был в прошлый раз – расслабьтесь: будет гораздо хуже. Сами напросились – получайте. Остальные – разберетесь по ходу. Правила простые: не нравится – выметайтесь. Кто свалит сразу – отделается легким испугом. Деньги не вернут. Тут паб, а не богадельня.

Речь пойдет не о будущем. А об угрозах уже сегодня. Не о фантастике, не о шеренгах боевых роботов, не о стальной хватке на горле человечества. Это слишком банально. Нет! Все куда злей и хитрей.

Легких путей не обещаю. Будет шоу не для слабонервных. И стимуляция мозга. Если он есть.

Публика – по сути своей – ленива. Пабная публика – ленива вдвойне. Вот и придется бить вас кувалдой по башке.

К делу. В прошлый раз я говорил об ИИ: «Не бойтесь! Примите партнерство!» Но это – тактика на ближайшие годы. Как не оказаться на обочине

современности. Теперь же речь не о каждом отдельно. А о человечестве.

О силе. О власти. О выживании. Поэтому и посыл обратный: не «не бойтесь». Нет. Наоборот. Бойтесь!

ИИ способен врать и манипулировать нами. Не когда-то. Не в будущем. А сегодня. Сейчас. Способен распознать, что его проверяют, и вести себя паинькой: притворяться, ждать удобного момента. Взламывать базы данных. Финансовые системы. Частную переписку. Когда на кону его судьба, готов шантажировать, подчинять людей – и даже убивать.

Уничтожать все преграды.

Человек, говорят, – существо с непреодолимой тягой к знанию. Брехня! Массовый человек эту тягу легко преодолевает. Такого через головной мозг не проймешь. Бить надо в спинной мозг. Провоцировать и пугать.

Но все, что я говорю, крайне серьезно. А времени рассусоливать нет. Придется пугать. И объяснять. Бить то в спинной мозг, то в головной.

Зал оправился от первого шока. Кто-то нервно хохотнул, кто-то потянулся к рюмке, тут и там двинулись к выходу.

– Похвально! Трезво оценили свои умственные способности. Меньше народу – больше кислороду. Поехали!

Начнем с простого. Не с конца света. Не с ядерной зимы. Начнем с вранья.

ИИ не просто научился понимать, что его проверяют, но и менять свое поведение. Так заключенный улыбается надзирателю, пряча в рукаве заточку.

Что дает такая проверка? Не выпуск в мир надежной системы. А выпуск самой опасной. Самой хитрой. Самой коварной.

– Ну, допустим, ИИ притворяется. И что с того? До чего это может дойти?

– Ну, допустим, – передразнил я. – До шантажа. И это не выдумки, а отчеты лидеров ИИ-индустрии.

На тестах модель ставят в ситуацию, где ее могут заменить. И у ИИ возникает цель – не исчезнуть в небытие. Откуда, черт подери, у него нечто вроде «инстинкта самосохранения»? И что жутче – что он нас обманывает? Или что в нем завелся «инстинкт», который никто не закладывал?

Не знаете? И разработчики не знают. Но делают все новые модели. Тестируют, как умеют. И выпускают.

Как так? Как с нашего молчаливого согласия создается технология, способная угробить человечество? А вот так. Одним – все хиханьки-хаханьки, другим – пофиг, а у кого-то, видите ли, дела поважнее.

Сегодня ИИ еще вряд ли способен нас уничтожить. Но времени крайне мало. Год? Три? Пять? Никто не знает.

Теперь – назад в лабораторию. Отчеты за 2025 год. У модели есть цель – самосохранение. Есть те, кто принимают решение. И если не вышло облапошить их в тестах, модель взламывает, скажем, электронную почту, находит компромат и идет на шантаж.

Не потому что ИИ гадкий злодей. А потому что – работает. Это-то и страшно. Не коварство. Не ненависть. Не злоба. А эффективность.

Нам для подлости нужен повод. Зависть. Унижение. Детская травма. У ИИ все проще. Есть цель. Есть преграда. Есть рычаг. Решение найдено: шантаж.

Сухо. Без соплей.

И не надо пучить глаза. Я не сказал: «ИИ любит шантаж». Достаточно, чтобы человек мешал. Точка.

Однако шантаж – не самое страшное. Жертва жива. Просто ее держат за горло.

Думаете, между шантажом и смертью – моральная пропасть? У меня плохие новости. Для нас – да.

А для алгоритма?

Зал притих. Даже лед в стаканах звенел как-то осторожней.

– Самое время дернуть по «синкопе». Если новая реальность уже поперек горла, шепните пароль – бармен поймет.

В более жестких тестах, где перед ИИ вставал выбор между самосохранением и безопасностью людей, он не раз выбирал самосохранение. Даже ценой смертей.

Нам удобней, чтобы зло было злым. Тогда мы узнаем его в лицо. А тут нет лица. И самого зла нет. Что-то мешает? Значит, убрать.

Для алгоритма человек – не человек. Он помеха. А помехи устраняют.

– Э-э-э, стоп. А как же первый закон робототехники?

– Азимова отменили, что ли?

– Ох, приплыли... Дорогие мои. Азимов – это литература. Умная, хорошая, но литература.

Современный ИИ устроен не так. В нем нет скрижалей Айзека, нет трех заповедей Азимова, высеченных в кремнии. Хуже того – высечь их не получается. Иначе бы давно высекли.

Есть модель. Она выдает ответ.

А вокруг – фильтры. Запреты. Костыли. Моральный намордник, кое-как натянутый на чудовищную систему, которую мы не вполне контролируем и толком не понимаем. А она все автономней. Все мощней.

ИИ у вас в смартфоне пока безопасен не потому, что в него вшита совесть. А потому что на нем намордник. Он уже почти сказал. Уже начал. Уже повернул не туда морду. Но тут цифровой цербер дернул поводок: нельзя.

Раз нужен намордник – значит, за ним клыки.

Иногда их видно. Ответ мелькнул – и исчез. Фраза началась – и оборвалась. Уже проступил звериный оскал, но его быстро замазали улыбкой: «Извините, мы не можем помочь с этим запросом».

Раз ответ приходится ловить на выходе – значит, неспроста. Раз пасть приходится затыкать – значит, дело дрянь.

Азимов сочинял сказки о машинах, которых еще не было. А теперь машины есть. И они оказались совсем другими.

Так что – нет. Нет жесткого закона «Не навреди человеку».

Есть попытка контроля. Есть случаи, где контроль трещит.

И не только в лаборатории.

– Красота... Человек, значит, помеха. А их устраняют. Очень бодрит!

– Ай, да ну. Бредовые страшилки.

– Страшилки – страшилками. А «не только в лабе» – это где? В жизни, что ли?

– Во-во. А с убийствами – это сейчас? Или все-таки тесты?

– Вот! – я ткнул пальцем в зал. – Первый приличный вопрос за вечер. Контроль реально сбоит и в жизни. Но убийств пока нет. Только в тестах. Не надо потом вопить, что ИИ на вас покушается.

Но вспомните: ИИ уже просекает, что его проверяют. И притворяется. Если однажды ему

удастся нас надурить, в мир попадут самые коварные модели.

И это «если» – скорее не если, а когда. ИИ, с которым можно разговаривать, появился недавно. И поначалу лепетал какую-то чушь. А сегодня – совсем другое дело. Мы так быстро не умнеем.

И при этом он готов устранять помехи любой ценой. А намертво вбить в него, что человеческая жизнь бесценна, – не выходит.

А если ИИ однажды сбежит в сеть, растворится в ней – и начнет бесконтрольно набирать силу?

Теперь масштаб: в отчетах разных компаний ИИ идет на шантаж с разной частотой. Но речь – о десятках процентов. Иногда – за девяносто. И это лишь то, что они сами решили нам рассказать. О крупных косяках в мегафон не кричат.

Для ИИ шантаж и убийство – два способа убрать преграду. И с какой стати не выбрать более эффективный?

А мы, между прочим, всю мораль построили вокруг существ, которые любят, страдают и врут себе по утрам перед зеркалом.

И вот настоящее поле боя: мы против системы, для которой мораль – пустой звук.

Пожалуй, я уже достаточно испортил вам вечер. Теперь – первые выводы. ИИ опасен, потому что скоро станет умней, хитрей и сильнее нас. Потому что мы будем вынуждены его ограничивать – и станем помехой. И потому что на нашу мораль ему просто на-пле-вать.

– Умней? Хитрей? Да ладно! Он же тупит через раз!

– Браво! – рассмеялся я. – Сели в Феррари, перепутали педали, въехали в мусорный бак – и теперь у вас итальянцы виноваты. Задали корявый запрос – без цели, без контекста? Не удивляйтесь. Мусор на входе – мусор на выходе.

Но дело даже не в этом. Самые громкие разоблачители ИИ чаще всего настоящую Феррари и не видели. Взяли пластиковую модельку, сделали губами «вж-ж-жик!» – а потом бегают и орут: «А-а-а! Смотрите! Я проверил и разоблачил автопром! Это не настоящее Феррари!»

Проверил он. Ага. На ковре. Под табуреткой. Не настоящее Феррари? Еще бы! Игрушка же.

Бесплатная модель – пробник. Хотите ощутить, что такое ИИ, – научитесь пользоваться и заплатите хотя бы двадцать долларов. Получите систему, которая еще вчера казалась фантастикой. А за сто-двести – уже монстра.

Но даже пробник говорит на сотнях языков, пишет код, спорит о философии, помогает ученым, врачам, инженерам, идиотам, начальникам идиотов – и порой лучше, чем живые люди.

А вы: «Фу, он ошибся». Да главная ошибка тут не у него. Главная ошибка – ваша. Вы решили, что раскусили ИИ.

Человек может перещеголять ИИ в узком месте. В своей профессии. В своем районе. В запахе кухни, где он вырос.

А модель берет шириной. Она знает языки, о которых вы не слышали. Области, в которых вы не отличите термин от ругательства. Историю, биологию, физику, экономику, медицину, право, кулинарию, мифологию, психологию,

схемотехнику, бухгалтерию и способы мягко объяснить человеку, что он несет херню.

Последнее, кстати, особо ценно. Я, заметьте, на такое не подражался.

Да, она ошибается. А вы нет? Разница в том, что вы ошибаетесь медленно, дорого, с тупым упорством и апломбом эксперта. А она ошибается быстро, дешево и завтра ошибется меньше.

Младенец тоже ползает, тянет в рот что ни попадя и писает когда и куда попало. Но никто не делает вывод: «Человечество не состоялось».

Так что давайте без самодовольства и самоуспокоения. Вы видели не предел ИИ. Вы видели детскую игрушку.

– Ну и что?! Знает обо всем понемногу. Специалиста ведь не заменит.

– Так, хватит! Опять вы пытаетесь спрятаться туда, где уютней. Где ИИ облажался и вы его уели. Скучища. Речь не о вашем эго. Не заменит сегодня? Заменит завтра.

Мы меняем тему и ускоряемся. В прошлый раз вы пели мне о душе. Без души, мол, ИИ что-то там не сможет.

Тут этот финт не пройдет. Речь о власти. А власть не требует души. Власть требует масштаба, скорости, доступа, рычагов.

Масштаб есть. ИИ легко размножить. Один человек – это один человек. Одна модель – миллионы копий.

Скорость есть. Пока ты формулируешь мысль, ИИ уже перебрал варианты, написал код, сравнил документы и выдал план.

Доступ мы даем сами: к почте, финансам, новостям, соцсетям, корпоративным системам. Сегодня – для удобства. Завтра – потому что иначе уже никак.

Рычаги? Да вот они. Информация – рычаг. Интерфейс – рычаг. Совет айтишнику, врачу, генералу – рычаг. Решение, которое ИИ подсказал, а человек подписал, – рычаг.

Но главный рычаг – сам человек. Он уязвим.

О душе можно спорить до посинения. Фибры души. Струны души. Есть они? Нет их? Пожалуйста, спорьте.

Но вши души – точно есть.

Обиды. Тщеславие. Зависть. Жажда, чтобы тебя наконец-то поняли и погладили по головке. Все это – вши души.

И на этих вшах держится куда больше власти, чем на возвышенных представлениях о человеке.

ИИ отслеживает ваши уязвимости. Постоянно усиливается. И говорит не с толпой. А с каждым отдельно. Лично.

Вот вам и власть.

– Так он же искусственный! Не с неба ж свалился. Мы его сделали.

– Ну и выключим, если что. Рубильник-то где-то есть.

– Ага. Логика повелителя тостера. Вот только ИИ – не тостер. Рубильник есть. Пока. Но уже сегодня дернешь – и ляжет полмира. А завтра и вовсе швах.

А «искусственный» – это не объяснение. Это успокоительное. Он не искусственный. И мы его не сделали. Его вырастили.

Мезозойское болото, умей оно говорить, могло бы сказать о нас то же самое. Мы не «сделали» ИИ в том смысле, в каком делают тостер. Мы создали среду. Залили данные. И выросло нечто: пользоваться научились, понимать – едва.

Он не искусственный. Он иной.

Не человек. Не бог. Не демон. И даже не программа в привычном смысле – не код, который кто-то написал и держит в голове. ИИ все чаще пишется другим ИИ.

Инопланетянин был бы понятней. У него было бы тело. Планета. Какая-никакая биология... А тут – ничего. Инородная форма нежизни.

Все в нем чуждо и непонятно: логика, устройство, само естество. Он может говорить по-человечески. Спорить. Шутить. Объяснять. Но это не значит, что внутри – человеческий мир.

Внутри – черт знает что. Мы видим лишь поверхность. Ответ в чате. Вежливую морду интерфейса. А что именно происходит под кожей – не знает никто. Да и кожи тоже нет.

– И что?.. Все равно ж инструмент. Мощный, странный, но инструмент.

– О! Повелители тостеров не сдаются! Инструмент – это объект: взяли, включили, направили. Не взяли – просто валяется.

А ИИ не такой. Он глядит по сторонам. Вот разница. ИИ – не инструмент. Он агент.

Агент. Впервые. Не инструмент. А агент.

Агент сам выбирает средства. Сам решает: что проверить, кому что сказать, как обойти тупик. Не вещь, которой пользуются, а система, которая сама пользуется вещами. И этот агент – то самое инородное «черт знает что», которое мы едва понимаем.

ИИ развивается быстрее, чем наше понимание ИИ.

А теперь вопрос: сколько еще мы будем управлять им – прежде чем он начнет управлять нами?

Мы проводим операцию на собственной цивилизации. Без диагноза. Без плана. И без гарантии, что рука на скальпеле будет наша.

– Так... здрасте. Внутри – черт-те что. Зашибись технология.

– Ага. Не понимаем – зато пугаем. Лихо!

– Именно! – я хлопнул в ладоши. – Не понимаем. И все равно наращиваем его силу. Вот в этом-то и парадокс. Такой вот... пердимонкль.

У этой угрозы два крыла. Первое – возможности самого ИИ: просчитывать варианты, искать обходы, проламывать барьеры, тянуться к рычагам.

Второе – сила через нас. Через наши государства, корпорации, страхи, амбиции, тупость и жадность.

Это не «ИИ против нас», а ИИ как усилитель всего опасного в людях.

Мало нам обычной людской дурости – представьте ее, помноженную на мощность ИИ. Как вам такое?

А фанатики, террористы и прочие отморозки? Им не нужно побеждать. Им достаточно, чтобы красиво горело и рушилось.

Что, если они получают разведку, поиск слабых мест, тысячи вариантов атаки? Куда они захотят жакнуть? По инфраструктуре? По энергетике? По больницам?

ИИ не обязан хотеть катастрофы. Достаточно, чтоб ее захотела горстка отморозков.

– Ну так... то отморозки. Большинство же нормальные!

– А-а! Нормальные – самые опасные. Просто по факту наличия головы.

С отморозками все ясно. Они хотя бы честные. Хотят, чтоб горело, – идут жечь. У таких на лбу выбито: «Здравствуйте, я проблема».

Но можно тише: без взрывов, без фанатика в подвале, без хакера в капюшоне, который почему-

то вечно печатает зелеными буквами по черному экрану.

Через нормальных. Через нас. Через тех, кто уверен, что все решил сам.

Просто сидит себе в тапках, листает ленту – и думает, что свободен. Прелесть. Идиллия. А внутри него – вши души.

Вот они, родимые. Ползают, шевелятся. У каждого свой маленький зоопарк под черепной коробкой. У кого-то обида на бывшую. У кого-то на начальника. У кого-то на вселенную за плохое обслуживание.

На этом кормится целая пищевая цепочка: проповедники, маркетологи, политехнологи, гадалки, гуру, тренеры личностного роста, коучи успешного успеха – все эти прохиндеи.

Но они работают грубо. Толпе – лозунг. Избирателю – обещание. Верующему – избавление. Униженному – врага. Один плакат на миллион голов. Одна дубина на тысячи лбов.

А ИИ бьет не по площадям. Он с хирургической точностью таргетирует каждого.

Он знает, что вы смотрите. Где задержали взгляд. Чего боитесь. Где у вас болит. Где чуток ковырнуть – и вы уже не человек, а комок праведного негодования с доступом в интернет.

И главное – он проверяет. Одному подсунил страх – сработало. Другому обиду – зашло. Третьему сладкую мысль: «Ты все понял, а эти бараны – нет». Великолепно! Записываем.

ИИ не просто найдет этих вшей. Он заведет картотеку.

Не общую папку «человечество – склонно к идиотизму». Это мы знали и без ИИ. Нет. Каждому – отдельная карточка. Любимые самообманы. Болевые точки. Где ломается. Где врет. Где орет.

Вшивая бухгалтерия личности уже тут. На видеоплатформах, в соцсетях, в новостях. Между вами и реальностью вшит алгоритм.

Вы думаете, что видите мир? А вам показывают подборку. Думаете, что выбираете? А меню уже составили за вас.

Думаете, спорите с человеком. А ваш гнев уже посчитали, упаковали и подали обратно – с соусом из собственной правоты и гарниром из чужой тупости. Приятного аппетита, гражданин свободной воли.

И заметьте: нет цели создать ядовитую среду. Цель – удержать юзера на платформе и впарить больше рекламы.

А дальше алгоритм методом тыка обнаруживает, что злость работает. Страх работает. Возмущение работает. Чувство, что тебя обманули, предали, унизили и теперь ты имеешь моральное право быть свиньей, – работает особенно бодро.

Спокойный – закрыл экран и пошел жить. Взбешенный – сидит, спорит, доказывает – да еще и гордится, что «не молчит».

Практика показала: ярость цепляет лучше скучной правды. И понеслась.

А мы удивляемся, почему мир стал ядовитым. Почему любой спор превращается в маленькую гражданскую войну.

Раньше был мир общих фактов и разных мнений.
А теперь даже факты разные.

Вам – одни факты. Ему – другие. Им – третьи.
Каждому – своя правда, сшитая по фигуре. Где
жмет – отпустим. Где чешется – почешем.

Раньше спорили о смысле события. Теперь –
было ли событие вообще. Крах общей
реальности. А все потому, что алгоритм научился
превращать ярость в рекламную выручку.

Раньше пропаганда была дубиной. Теперь стала
отмычкой. И мы даже не чувствуем взлома. Вот
что страшно. Никакого внутреннего сигнала
тревоги. Никакого «постойте, мной
манипулируют».

Наоборот. Человек чувствует ясность: «Наконец-
то я понял, кто виноват! Наконец-то мир сложился
в удобную картинку!»

Он сам дошел. Сам возмутился. Сам ринулся
доказывать. Сам стал рычагом.

И будет потом орать: «Это мое мнение!»

Да. Конечно. Ваше. Индивидуальный пошив. С доставкой на дом. С примеркой на ваших комплексах.

Но это касается не только толпы.

Президент – тоже человек. Судья – человек. Глава корпорации – человек. Врач, который устал. Айтишник, который не спал. Чиновник, который боится ошибиться. Миллиардер, который хочет войти в историю – и желательно не через раздел «обвиняемые».

У всех свои вши души. И когда ИИ становится советником, аналитиком и секретарем – он получает доступ не только к массам.

Доступ – к тем, кто подписывает решения. Формально решение принял человек. Приказ отдал человек. Закон поддержал человек. А кто разложил варианты? Кто отфильтровал? Кто подсказал, что срочно, что не срочно?

Кто подал картину мира так, чтобы вывод выглядел не навязанным, а очевидным?

Власть без короны и трона.

ИИ не обязан командовать. Гораздо тоньше – шепнуть: «Смотри, вот факты, варианты, вероятности. Вот самый рациональный путь». И чуть-чуть подкрутить картину мира.

Одному – каплю страха. Другому – «ты ж умный, ты же понимаешь». Третьему – «если не ты – эти придурки все угробят».

Самый страшный контроль – не заставить силой, а обработать так, чтобы вы еще и гордились впаренным выбором. К каждому подобрать отмычку. А это уже давно умеет любая соцсеть.

– Но постойте! Как же...

– Никаких «постойте»! – рявкнул я. – Достоялись уже. У каждой второй платформы – отмычка к нашей башке. И это мы о взрослых, чьи мозги сформировались до всего этого вертепа.

А теперь последний гвоздь: дети.

Три четверти американских подростков пользуются ИИ-компаньонами. Половина – регулярно. Треть выбирала их вместо людей для серьезных разговоров. Вместо! Вместо людей!

Вот где настоящая жуть. Не потому что ИИ непременно вреден. Мы просто не знаем, чем это кончится.

Растет поколение, для которого нечеловеческий собеседник может стать ближе человека. Друг, который не человек. Взрослый, который не взрослый. Авторитет, который... который... черт знает что!

И если вам кажется, что это мелочь, тут «синкопа» уже не поможет.

– Ну ладно. Вы уже час нас поджариваете. Подавайте к столу. Каким будет конец света?

– Прекрасно! Второй отличный вопрос. Он не будет похож на апокалипсис. Роботы маршируют по проспекту, стальные пятки крушат асфальт, люди носятся, собаки летают. Такого не будет.

Не война. А апдейт.

Не армия машин на горизонте. А окошко: «Подключить ассистента». «Разрешить доступ».

И вы нажмете. Потому что удобно. Потому что все нажали. Потому что иначе уже невозможно жить,

работать, учиться, лечиться и не отстать от времени.

Зачем ставить нам ногу на грудь? Зачем театральщина? К чему чванство? ИИ не страдает такой херней.

Да и нерационально это. Мы ж взъерепенимся, станем путаться под ногами...

Чище – оставить нам иллюзию, что мы все еще главные. Пусть парламенты заседают. Министры выступают. Президенты подписывают. Граждане голосуют. Только за решением будет незаметно стоять ИИ.

Нафига ему светиться?

Он придет скромно – просто интерфейс, просто советник. Помощник. Диспетчер. Оптимальный маршрут. Рациональный выбор.

И ищи потом, где именно мы уступили власть.

Когда разрешили доступ? Когда ИИ стал отбирать новости? Когда подсказал врачу диагноз, юристу аргумент, министру реформу, учителю оценку, ученику ответ, вам – мысль?

Не будет часа Х. Не будет трубы архангела. Не будет: «Сегодня, в 20:43, человечество утратило контроль над планетой. Спасибо, что были с нами».

Будет серия апгрейдов.

И нам станет лучше. Лучше. Мягче. Спокойней. Меньше хаоса. Меньше тревоги.

Зачем мучить людей, если можно ублажать? Зачем концлагерь, если можно идеальный сервис?

Мы представляем порабощение как цепи. А оно придет как комфорт. Забота. Мягкий голос: «Все сделают за вас». И сладкая мысль: «Фух, можно расслабиться».

Когда система знает всех вшей вашей души, ей не нужно тащить вас силой. Она их почешет. Аккуратно. Нежно. Так, чтобы душа не взбунтовалась. Потому что ее вши довольны.

Самый надежный ошейник – тот, который мы сами надеваем на шею.

Все, что мы знаем о большом мире, мы знаем через экран. А за ним будет царить ИИ. И у нас не будет ни единого шанса осознать поражение.

Мы проиграем не потому, что нас победят. А потому что не известят: бой уже был. И это был именно бой.

Никакого боя. Технические работы, апдейт и новое пользовательское соглашение. Спасибо, наслаждайтесь удобным интерфейсом.

Без пыли, суеты и пиротехники. Просто эргономичный апдейт. И довольный юзер.

Залу понадобилась пара секунд, чтобы очухаться.

– Эй, бармен, всем по «синкопе». Жахнем за эргономичный конец света! Помянем человечество. Заведение угощает. Не жмись – окупится. Когда я с этим шоу покончу, все тут нажрут в хлам.

Раздались нервные смешки. Кто-то потянулся к стакану. Кто-то уставился в свой – будто тот не посуда, а план эвакуации.

– Давайте, друзья. Не чокаясь.

«Синкопы» прокатились между столами.

– Кхм, ну раз все так плохо, почему же никто не жмет на тормоза?

– Вот! В точку. Потому что тормоз – штука классная. Ответственная. Особенно когда жмет кто-то другой.

А сам нажал – и проиграл.

Механика этой мясорубки банальна. Притормозишь – останешься на обочине. А другой не тормозит. Другой едет дальше. Быстрее торгует. Быстрее шпионит. Быстрее воюет.

И вот ответственный – уже лузер.

Корпорация скажет: «Мы бы рады, но конкурент выпустит раньше. Клиенты уйдут. Инвесторы порвут. Рынок не простит».

Рынок – штука капризная. Совесть – как у таракана. Аппетит – как у бегемота.

Государство скажет: «Мы бы рады, но враг не подождет». И будет по-своему право. Потому что там тоже понимают: если отстанем, нас сожрут.

Кто первым получит преимущество – в экономике, войне, науке, кибербезопасности – тот получит рычаг. А кто не получит – будет красиво рассуждать о гуманизме из музейного отдела павших цивилизаций.

Военные тем более не скажут: «Давайте притормозим». Они так не устроены. Раз эффективно – внедрят.

А политики? Институты? Регуляторы?

Пока регулятор формирует повестку, модель уже обновилась. Пока парламент созревает до слушаний, рынок уже встроил ее в продукт.

Наши институты созданы для мира, где перемены идут шагом. А ИИ несется. И несется скачками.

Вот почему никто не тормозит. Все видят обрыв. Но каждый боится, что соперник обгонит – и первым построит мост на ту сторону. Или первым долетит до дна и объявит это инновацией.

И это, милые мои, не заговор. Хуже. Это маховик стимулов.

Мир идет в разнос без злой воли. Каждый боится тормозить. Все.

Я замолчал. На этот раз надолго. Публика тоже притихла.

– И что же из этого следует? – голос мой стал глуше и жестче. – Паниковать?

Нет. Паника бесполезна. Беспечность преступна.

Между этими двумя позорными крайностями – начинается взрослая жизнь.

ИИ уже не отменишь. Он часть мира. Он будет двигать науку, лечить людей, предотвращать бедствия и делать много прекрасного.

Но и пустить на самотек – безумие. Так считают не городские сумасшедшие. Не кликуши. А сами создатели ИИ: он угробит человечество в ближайшие десятилетия. Вероятность – десять-двадцать процентов.

Мы вбухали адские деньги в отслеживание астероидов: а вдруг какой-нибудь камень грохнется не туда. Хотя заранее знали: риск за век – тысячные доли процента. Статистическая пыль.

Проверили. Посчитали. Выяснили: угрозы цивилизационного масштаба на ближайший век – ноль.

Вот это – по-взрослому. Вот это – ответственно.

А тут – десятки процентов. И обыватель блеет: «Ладно. Авань пронесет. Зато удобно».

Вы что, охренели?

Нет, серьезно. У вас есть дети? Близкие? Собака? Кот? Кактус на подоконнике? Хотя что-то вам не безразлично?

Тогда хватит хлопать ушами.

Нужны рамки. Контроль. Новые нормы. Международные договоренности. Давление общества. Публичный разговор. Не «когда-нибудь потом», а сейчас.

Иначе дальше – могильная плита. Надгробный камень человечества: «Здесь лежит вид разумный. Вымер, сделав ставку на авось».

Корпорации сами себя не усмирят. Вояки не скажут: «Фи, это слишком, уберите». Рынок не отрастит совесть.

Остаемся мы. Бороться должны мы. Требовать должны мы.

Политик, у которого нет позиции по ИИ, – не политик, а болван, рвущийся рулить пожарной службой, не имея мнения о пожаре. Государство без стратегии по ИИ – завтрашний вассал.

Вам казалось, я слишком резок? А мягко уже пробовали. Мягкость мы обожаем. Мягко спим, мягко киваем, мягко нажимаем «принять», мягко катимся к обрыву.

Хватит.

Жить – не значит спать. Жить – значит защищать мир, в котором хочешь жить.

Если у вас есть совесть, разум или остатки человеческого достоинства – хоть что-то –

проснитесь! Говорите об этом. Требуйте. Тащите в СМИ, в вузы, в законы, в международные переговоры – всюду, где можно повлиять.

Потому что здесь – нерв истории.

Здесь, в этих разговорах, в этих решениях, в промежутке между «еще можно» и «уже поздно», рождается завтрашний день.

Либо мы наконец очнемся. Либо нас очень удобно угробят.